Published in final edited form as:
*J Speech Lang Hear Res*. 2007 August ; 50(4): 968–981. doi:10.1044/1092-4388(2007/068).

# The Discriminant Accuracy of a Grammatical Measure With Latino English-Speaking Children

**Vera F. Gutiérrez-Clellen**
San Diego State University

**Gabriela Simon-Cereijido**
San Diego State University and University of California, San Diego

## Abstract

**Purpose—**To evaluate the discriminant accuracy of a grammatical measure for the identification of language impairment (LI) in Latino English-speaking children. Specifically, the study examined the diagnostic accuracy of the Test of English Morphosyntax (E-MST; Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore (n.d.) to determine (a) whether use and exposure to Spanish had an effect on the performance of bilingual children compared with monolingual Latino children and (b) whether dialectal differences within Latino English speakers might result in performance differences and a greater incidence of misclassifications for children from Caribbean English backgrounds.

**Method—**One hundred and eleven children (i.e., 59 children with typical language development and 52 children with LI) were sampled from the Southwest and Northeast regions of the U.S. Southwestern children were of Mexican origin. Children from the Northeast were from Puerto Rican or Dominican backgrounds. Linear discriminant analyses evaluating group classifications on the basis of the E-MST were performed on exploratory and confirmatory data sets across 3 groups: Southwestern English-only proficient (SW EP) children, Southwestern English-dominant bilingual (SW EDB) children, and Northeastern (NE) children.

**Results—**Results of the exploratory discriminant analyses indicated good sensitivity for the SW EP children. The discriminant functions derived from the exploratory analysis were able to predict group membership in confirmatory discriminant analyses with fair sensitivity and good specificity for the SW EDB children and with fair sensitivity but poor specificity for the NE children. Children who were English-dominant bilingual were not more likely to be misclassified compared with their English-only proficient peers. However, nonmainstream English dialect differences appeared to affect classification accuracy and resulted in a greater number of misclassifications for the NE children with typical language development.

**Conclusion—**The measure seems to be suitable for identifying LI in SW children who are exposed to Spanish and/or who are English-dominant bilingual. Additional assessment tools will be needed to rule out the disorder in children who are exposed to African American or Caribbean English dialects.

### Keywords

language impairment; English-speaking bilingual children; discriminant accuracy



Contact author: Vera F. Gutiérrez-Clellen, School of Speech, Language, and Hearing Sciences, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-1518. vclellen@mail.sdsu.edu..

Research with mainstream English-speaking children has helped advance our understanding of the grammatical characteristics of specific language impairment (SLI) and the potential predictive value of grammatical measures in clinical assessments (for a detailed review, see Leonard, 1998). However, accurate identification of SLI in children from nonmainstream backgrounds may be difficult because these children may score lower than the norms on standardized language tests in spite of having typical language development (TLD; Norris, Juarez, & Perkins, 1989; Peña, Quinn, & Iglesias, 1992; Qi, 2005; Qi, Kaiser, Milan, Yzquierdo, & Hancock, 2003; Restrepo et al., 2006; Washington & Craig, 1992, 1999; Wilcox & Aasby, 1988). The differential test performance of nonmainstream children cannot be attributed to differences in socioeconomic status (SES). In an early study, 56% of 299 African American preschoolers attending a Head Start program received Test of Auditory Comprehension of Language (Carrow-Woolfolk, 1999) scores of more than 1 *SD* below the mean of the middle SES normative sample (Ramsey Musselwhite, 1983), yet their scores remained low when compared with the means of children from low SES.

These differences have been related to content bias, linguistic bias, normative bias (Laing & Kamhi, 2003), and poor psychometrics of current language tests (Spaulding, Plante, & Farinella, 2006). Available language tests may penalize nonmainstream English dialect use, and as a result, they may underestimate the language skills of children with TLD. For example, African American children received significantly lower scores than their White peers on the Test of Language Development–Primary (TOLD-P; Hammer, Pennock-Roman, Rzasa, & Tomblin, 2002; Newcomer & Hammill, 1991). The items that differentiated the two TLD groups involved not only vocabulary but also grammatical targets (i.e., items included in Grammatic Understanding, Sentence Imitation, and Grammatic Completion subtests). Scoring adaptations to address the use of African American English (AAE) by some speakers did not improve the performance of African American children on this test. Because the majority of the test items are still included in the third edition of the TOLD-P (Newcomer & Hammill, 1997), these measures continue to be problematic for the assessment of nonmainstream children.

Hispanic English is probably the native dialect for many Latino children in the United States (Wolfram & Schilling-Etes, 1998a). Yet there is no available research on the potential effect of nonmainstream English dialect use on their test performance. This English variant includes use of postnoun modifiers for possessives (e.g., “homework of my brother”), nonobligatory use of plural (e.g., “the girl are playing”), nonobligatory use of regular past –*ed* (e.g., “I talk to her yesterday”), nonobligatory regular third-person present tense (e.g., “she eat too much”), use of “no” before the verb for negation (e.g, “she no eat candy”), omission of subject pronouns (e.g., “Father is happy. Bought a new car”), and lack of inversion and auxiliary verbs in questions (e.g., “Mary is going?”), among other characteristics (Owens, 1991; Wolfram & Schilling-Estes, 1998b; Zentella, 1997). Little is known regarding specific Hispanic English dialect features, frequency of use, or contexts (Wolfram & Schilling-Estes, 1998a). Within Latino speakers, nonmainstream dialect use is not stable. Many Puerto Rican children speak both mainstream and nonmainstream English varieties (Zentella, 1988). Nonmainstream dialect use may vary across Latino subgroups, as well. For example, Puerto Rican or Caribbean English (i.e., a dialect spoken in the Northeast region of the United States) and Chicano English (i.e., a dialect spoken by Mexican Americans in the Southwestern United States) may share some but not all of their features. Because these English varieties emerged in environments where both English and Spanish were used, they show influence of Spanish in phonology (i.e., high vowels, final consonant omissions), lexical choices, and morphosyntax (Wolfram, 1974). Children from the Northeast region may also exhibit dialect features much like their AAE-speaking peers (Wolfram, 1974). Some Caribbean English characteristics are *vestigial*—that is, they are transferred from Spanish, such as the use of “no” (e.g., “you no smell no nasty air,” “I no

used to it"). However, this use is sometimes replaced by "not" ("He not even missed one guy"). "No" is never used before an auxiliary (Wolfram, 1974). Other characteristics overlap with AAE but have a different origin. For example, deletion of past tense *–ed* in Caribbean English may be related to consonant cluster reduction.

In contrast, Southwestern Mexican American children are not expected to demonstrate features of AAE to the same extent as Northeastern Caribbean speakers (Wolfram & Schilling-Estes, 1998b). Some of their nonmainstream dialect features include multiple negation (e.g., "she didn't tell me nothing about it"; Fought, 2006; Ornstein-Galicia, 1981), use of *was* with plural subjects, and final consonant deletion in clusters (e.g., *tes*' for *test*). These features are also found in Native American English and AAE speakers (Wolfram & Schilling-Estes, 1998a).

Although many speakers of these nonmainstream dialects are not bilingual (Fought, 2006; Ornstein-Galicia, 1981), children who are bilingual may also exhibit differences related to their use and/or exposure to Spanish—that is, if children experience cross-linguistic influence, one may find a greater use of ungrammatical forms for a longer period of time (i.e., Döpke, 2000), in addition to nonmainstream Latino English features. These potential differences have not been addressed in the development of most available tests. Current language tests do not provide sufficient information about the bilingual characteristics of the Latino children included in the sample norm or about their performance for specific subtests. A review of commonly used English language tests, including the TOLD-P (Newcomer & Hammill, 1997), the Test of Language Development–Intermediate (Hammill & Newcomer, 1997), the Clinical Evaluation of Language Fundamentals–Preschool (Wiig, Secord, & Semel, 1992), the Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel, Wiig, & Secord, 2003), and the Preschool Language Scale–4 (PLS-4; Zimmerman, Steiner, & Pond, 2001) revealed that Hispanics accounted for 9%–17% of the standardization sample. Yet no characterization of bilingual status, bilingual proficiency, or exposure to Spanish was reported in the technical manuals. Only the PLS-4 indicated the number of speakers of other languages (52 Chinese, 1 Tagalog, 1 Korean, 50 Spanish, 6 other).

Attempts to adjust cutoff scores or eliminate problematic items from language tests have not been successful. The Hannah–Gardner Test of Verbal and Nonverbal Language Functioning (Hannah & Gardner, 1978) was administered to 540 preschool English-speaking children participating in a Head Start program, including 376 African American, 82 Hispanic, and 82 Anglo children. The items on the Expressive subtest were adapted to address the dialectal differences of speakers of African American and Hispanic English. Even with this adaptation, more than 68% of the children screened obtained scores at least 1 *SD* below the mean of the normative sample (Norris et al., 1989). The comparisons continued to reveal significant differences using different cutoff scores. About 32% of the Head Start children scored below 2 *SD*s from the mean. The chance of underestimating the language skills of these children was significant for both nonmainstream groups.

The potential for misdiagnosis using current language tests exists even when no statistical differences between groups are found. A recent validation study of the Structured Photographic Expressive Language Test–Third Edition (Dawson, Stout, & Eyer, 2003), a test that targets morphosyntactic structures, was conducted with 4- and 5-year old children of different backgrounds (Perona, Plante, & Vance, 2005). Differences between the scores of Latino and non-Latino children and between children exposed to a second language at home or not exposed to a second language at home were not significantly different ($p = .065$ and .22, respectively). However, reported effect sizes were moderate ($d = 0.61$ and 0.41, respectively), indicating that these differences may be clinically important. The lower scores of the Latino participants might be related to their use of a non-mainstream English dialect

or to cross-linguistic influence from Spanish, specifically when children are tested in their nondominant language. There is some evidence that bilinguals may differ from monolinguals depending on the particular language skills or tasks examined (e.g., use of complex syntax, morphosyntactic accuracy, grammaticality judgments) and the specific grade levels involved (Gathercole, 2002a; Pearson, 2002). For example, a comparison between 160 Spanish–English bilingual children and 80 monolinguals in Miami showed a main effect for morphosyntactic accuracy (i.e., grammatical well-formedness) in second and fifth grade (Pearson, 2002). Group differences were also found for complex syntax in second but not in fifth grade. Experiments comparing the patterns of acquisition of Spanish–English bilinguals and monolinguals using grammaticality judgments for mass/count distinctions in English also indicated that the bilinguals lagged behind monolinguals in second grade, with smaller group differences by fifth grade (Gathercole, 2002a). Grammaticality judgments for Spanish gender showed better performance for monolinguals than bilinguals as well (Gathercole, 2002b). For both count/mass distinctions and gender judgments, the bilinguals took longer to match the performance of monolingual children. However, when comparing these groups on their ability to judge English sentences involving the extraction of embedded subjects (e.g., “Who did you say that came to the party?”), the results varied depending on SES background and grade (Gathercole, 2002c). In contrast, we found no group differences when we evaluated the grammatical performance of bilingual Spanish-dominant and Spanish-only children using the Spanish Morphosyntax Test (Gutiérrez-Clellen, Restrepo, & Simon-Cereijido, 2006).

The present research was motivated by the need for test instruments that can accurately assess the language skills of Latino children who may speak nonmainstream English varieties. The study is part of a larger project, the Bilingual English Spanish Assessment (Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, n.d.), which is designed to develop and validate an assessment measure for Latino children 4–7 years of age with the aim of identifying Latino children with language impairment (LI). One aspect of that project (reported in Gutiérrez-Clellen et al., 2006) examined the validity of a Spanish Morphosyntax Test for the identification of LI in Latino Spanish speakers. The Spanish study included 160 children who were either Spanish-only or Spanish-dominant bilingual speakers. The present study was based on a different sample of children who spoke English only or who were English-dominant bilingual speakers (participant descriptions are provided later). Its purpose was to evaluate an English Morphosyntax Test (E-MST) for the diagnostic classification of Latino English speakers.

## Development of the E-MST

As was discussed earlier, the literature has not examined the potential role of nonmainstream dialect use or cross-linguistic influence on the grammatical performance of Latino children. Yet, many Latino children are expected to show variable nonmainstream (e.g., AAE and/or Puerto Rican English [PRE]) dialect use and/or differences related to cross-linguistic influence. In addition, because the characteristics and frequency of use of nonmainstream dialect use are not yet known, one cannot make direct predictions regarding which specific grammatical structures might have diagnostic accuracy. Therefore, the E-MST was conceptualized as a broad measure of grammatical ability.

Based on current research, the E-MSTwas designed to include verb forms (i.e., past tense –*ed*, third-person singular –*s*, copula and auxiliary *be*, auxiliary *do*) because they were found to be difficult for both AAE and mainstream English speakers (Oetting & McDonald, 2001). The E-MST also assesses noun morphology because previous research showed that a noun morphology composite that included correct production of possessive 's, plural 's, and articles correctly classified 79% of a group of children with SLI and 100% of the TLD

children in a discriminant function analysis study (Bedore & Leonard, 1998). When the study was replicated with a separate sample, all children (with and without LI) were correctly classified using the noun composite. Furthermore, because the goal was to examine broad grammatical skills, the measure also examines production of passives and complex syntax, such as relative clauses, adverbial clauses, and conditional clauses. There is evidence that children with LI demonstrate deficits with passives presumably due to (a) a general difficulty with complex movement operations affecting the passive transformation (Bishop & Butterworth, 1979; van der Lely, 1998), (b) processing limitations that affect nonsalient morphemes such as the past participle inflection *–ed*, and/or (c) rigid reliance on subject–verb–object word order (Leonard, Wong, Deevy, Stokes, & Fletcher, 2006). Passives were also found to help identify LI in nonmainstream English speakers with high sensitivity and specificity (Craig & Washington, 2000). Complex syntax was considered potentially useful because Latino children with limited language skills have limited syntactic complexity (Gutiérrez-Clellen, 1998; Gutiérrez-Clellen& Hofstetter, 1994). A sentence repetition task targeting complex syntax was an effective indicator of English LI (Conti-Ramsden, Botting, & Faragher, 2001). Thus, the E-MST included forms that were predicted to have potential for the clinical identification of nonmainstream English speakers with LI.

A series of two pilot studies were conducted during the development phase of the E-MST. The original E-MST included 127 items. This set of items was first piloted on 14 Mexican American children from the Southwest region of the United States (half with LI matched by age) who were English-only speakers, using a 25%-of-children cutoff criterion. That is, we examined items that elicited an alternative response in more than 25% of the children to decide if an item had to be eliminated due to a lack of an obligatory context. Using this approach, several items were eliminated, and the prompts were modified. Items that elicited inconsistent responses across speakers with TLD were also eliminated because they did not elicit targeted responses (e.g., a prompt of “This man has a moustache. And this man?” elicited the following responses: “Doesn't,” “Has no moustache,” “There was,” and “He don't”). These initial observations also indicated that children should not be penalized for use of double negatives because this dialectal form was used by a majority of typically developing speakers.

The E-MST includes actional passives with both animate and inanimate recipients. At the time of the study, no information was available on how to best assess passives reliably and efficiently. Thus, a second pilot study with 12 Mexican American children sampled from the Southwest region (3 English-only children with TLD, 5 English-only children with LI, 4 English-dominant children with LI) was conducted to examine performance on passives using an open-ended script to elicit the child's response (Gutiérrez-Clellen & Erickson, 2000). The responses were coded as correct when they contained the recipient, the auxiliary *is*/*was*, and the past tense *–ed* form of the verb. The use of *got* was considered acceptable on the basis of previous research with English-speaking children (Johnston, 1985). The children with TLD appeared to make errors by omitting the recipient and changing the sentence to active voice. Thus, they exhibited appropriate comprehension of passive constructions. None of the children with TLD used an incorrect recipient. They were also able to produce the active voice with the correct recipient and *–ed* marker. In contrast, the children with LI appeared to have limited comprehension of the passive construction. Only 3 of the children with LI used appropriate active voice. Five of the children with LI used the incorrect recipient, and the majority showed errors on the verbs (7 of 9 children with LI omitted the auxiliary; 8 of 9 children with LI omitted the *–ed*). On the basis of this preliminary work, and to facilitate the production of the target constructions, the format of the prompts was modified to elicit targets using a cloze task.

The E-MST items were also evaluated to determine if they had adequate levels of difficulty and discrimination. The goal of the item analyses was to retain only those items whose difficulty value showed a developmental progression across three distinct age groups and that had a fair discrimination value across children with different abilities. These item analyses were conducted on the total pool of participants available at this stage of development (i.e., 126 children with TLD; 72 children with LI) regardless of bilingual status (i.e., English only or English dominant), geographic origin, or other demographic variables. To evaluate item difficulty, the performance of children with TLD was compared across three age groups: 48–55 months (18 children), 61–71 months (54 children), and 76–84 months (54 children). Item difficulty was determined by dividing the number of children who passed each item by the total number of children at each age group. The item difficulty value (or *P*) for each item showed a developmental progression from more difficult for the 4-year-olds to less difficult for the older groups. These preliminary results indicated that the English items were developmentally appropriate and had sufficient difficulty. Item discrimination was based on 24 children with LI from 48 to 55 months, 27 children with LI from 61 to 71 months, and 21 children with LI from 76 to 84 months. The percentage of children with LI who passed each item was subtracted from the percentage of children with TLD 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 (Anastasi & Urbina, 1997; Friedenberg, 1995). The item analysis indicated that certain items had fair discrimination across the three age groups and others had better discrimination for the younger or the older groups. Items that had low *D* values (i.e. below 2) were eliminated. Based on this work, 63 items were retained.

The purpose of the present study was to determine the diagnostic accuracy of the E-MST not only across children with and without LI but also with children who might exhibit nonmainstream dialect differences. First, we evaluated the classification of children with TLD and LI with Southwestern children who could only speak English and had limited exposure to Spanish (participant groups are described in detail later). Second, we replicated the analysis with bilingual children from the same geographic region (i.e., Southwestern children who were English dominant and had high levels of Spanish exposure and use). We speculated that if cross-linguistic influence related to the child's use of Spanish had an effect on grammatical performance, the bilingual children would show lower grammatical accuracy rates than their age-matched monolingual peers, and they would be more likely to be misclassified as impaired. Finally, the E-MST was evaluated with a third group of children that was sampled from the Northeast region of the country (i.e., Philadelphia, PA) in order to determine if the measure could be used in clinical assessments with Caribbean English speakers.

## Method

### Participants

One hundred and eleven children and their families were sampled from school districts serving predominantly low-income families in the Southwest (California and Texas) and the Northeast (Pennsylvania) regions of the United States. The children from California and Texas were of Mexican American descent. The children from Pennsylvania were of Puerto Rican and Dominican backgrounds. Table 1 describes the parental education level, income, and ethnic characteristics of the TLD and LI 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. There were 59 children with TLD (*M* = 5;0 [years;months], *SD* = 0;6, age range = 4;0–5;10) and 52 children with LI, individually matched by age (*M* = 5;0, *SD* = 0;6, age

range = 4;0–5;10). The groups did not differ in parental level of education. There were three levels of education: (a) primary and some secondary education, (b) high school graduates and a few years of college (1–3 years), and (c) college graduates. Eleven parents reported some college education without graduation. As a group, these parents had 1.54 (0.69) years of college. Very few families reported college graduation in each group. A chi-square test comparing the number of families that reported primary and some secondary education, high school and some college, and college completion across the two language ability groups (TLD, LI) indicated no significant group differences in level of education, $\chi^2(2, N = 85) = 1.988$, $p = .3702$. The sections that follow describe the language characteristics of the participants (see also Tables 2, 3, and 4). Children were not matched for age across geographic regions because of limitations in the eligible pool of Northeastern (NE) children.

**Procedures and criteria to establish bilingual status and dominance—**The procedures and criteria for assigning bilingual status were based on parent and teacher questionnaires (see Gutiérrez-Clellen & Kreiter, 2003, for details on the validity of the questionnaires). These procedures are also described in Gutiérrez-Clellen et al.'s (2006) study. Children were determined to be English-only proficient (EP) if they had (a) minimal or limited exposure to Spanish; (b) a minimum parent rating of 3 for language use and proficiency in English based on a 5-point rating scale for each measure (0 = *no use or proficiency*, 4 = *use all the time and native-like proficiency*); and (c) a minimum teacher rating of 3 for language use and proficiency in English (based on a 5-point rating scale). Children were judged to be bilingual if they had (a) a minimum of 20% of time exposed to both English and Spanish; (b) a minimum parent rating of 3 for language use and proficiency in English; (c) a minimum teacher rating of 3 for language use and proficiency in English; and (d) the ability to speak Spanish, based on the analysis of their spontaneous language (see further details later in this section). The bilingual status of the group with LI was established using the same language exposure and language use ratings criteria. Proficiency ratings of the bilingual children with LI were lower than 3 in both languages. Table 2 presents the means and standard deviations of amount of language input and ratings of use and proficiency for the Southwestern English-only proficient (SW EP) participants with and without LI. Although the two groups met criteria with no significant differences for language use at home or at school, the group with LI appeared to have more exposure to Spanish at home than the group with TLD. To further establish whether the two groups were comparable, we examined the children's proportion of English and Spanish output. This information was collected in the parent questionnaires by asking parents about the child's language of choice when speaking to each member of the household. The comparison indicated no significant differences in the proportion of English output at home between the two groups ($p = .199$, $d = 0.78$), despite the TLD group having a significantly greater proportion of English input at home compared with that of the LI group. The TLD group had a mean proportion of English output at home of .95 ($d = 0.08$), and the LI group had a mean proportion of .88 ($d = 0.2$). These patterns of language use reflect a communicative style of many children from Spanish–English communities in the United States. Children may respond in English even when addressed in Spanish. Bilingual children may be dominant in English regardless of the language spoken by the family at home (Hakuta, 1986). Furthermore, although the two EP groups had some exposure to Spanish, they were unable to produce Spanish utterances even in informal interactions with the examiner.

Table 3 presents the means and standard deviations of amount of language input and ratings of use and proficiency for the Southwestern bilingual participants with and without LI. As with the English-only samples, there were no significant differences in English output across the two bilingual groups ($p = .867$, $d = 0.11$). The TLD group had a mean proportion of English output at home of .59 ($d = 0.18$), and the LI group had a mean proportion of .57 ($d = 0.18$; see Table 4 for further details on the language characteristics of the NE sample). In the

NE group, the TLD group had a mean proportion of English output at home of .81 ($d$ = 0.24), and the LI group had a mean proportion of .74 ($d$ = 0.24).

Because children who are bilingual may vary in their achievements in the two languages, and to ensure that these children were not tested in their weaker language, English dominance was determined using a direct measure of grammatical proficiency based on spontaneous narrative samples. The narrative samples were obtained using the following wordless picture books: *Frog, Where Are You?* (Mayer, 1969) and *One Frog Too Many* (Mayer,1975) for English; *Frog on His Own* (Mayer, 1973) and *Frog Goes to Dinner* (Mayer, 1974) for Spanish. The languages were tested in two separate administrations of the narrative procedure, one per language. Children were considered typically developing English-dominant bilingual (EDB) if they were able to produce English narrative samples with minimal grammatical errors (i.e., below 20% ungrammatical utterances) but with more difficulty in Spanish (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, Restrepo, Bedore, Peña, & Anderson, 2000) and was used successfully in previous research with bilingual children (Gutiérrez-Clellen et al., 2006).

**Criteria for identification of children with LI—**Children with LI were identified using procedures previously validated with English speakers (Bedore & Leonard, 1998) and with Latino children (Gutiérrez-Clellen et al., 2006). These criteria included evidence of (a) parent and/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, etc.) and (b) a significant percentage of ungrammatical utterances in their narrative samples (i.e., above 20% ungrammatical utterances; Gutiérrez-Clellen et al., 2000; Restrepo, 1998) in both their English (i.e., the child's “best” language) and their Spanish (if any was spoken). None of the children evidenced hearing impairments, mental retardation, emotional disturbance, motor difficulties, or neurological deficits, according to parent report and school records. At the time of the testing, 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 and/or teacher concerns. Both the children with LI and their typically developing peers were recruited from the same classrooms and schools. The children with TLD were judged as having TLD on the basis of parent and teacher reports as well as clinical observation. Upon verification that they met criteria, they were individually age matched with the children with LI.

## Experimental Procedures

The following sections describe the E-MST as well as the testing, scoring, and statistical procedures.

**E-MST—**The E-MST included 32 items presented in cloze format (e.g., Examiner: “Look, this is a girl, and this is the girl's hat. This is a boy, and this is the …” Answer: “Boy's hat.”) and 31 items presented in a sentence repetition format (e.g., Examiner: “Just say what I say. `The boy who broke the window is crying.” Answer: “The boy who broke the window is crying.”). Although in the sentence repetition task, children were asked to repeat verbatim the target sentence, not all the words in the sentence were scored as items. For example, the sentence above included only three items (*who, broke, is crying*). The cloze task targeted possessives, third-person singular, four regular and one irregular past tense, plurals, copula, auxiliary and negation, and passive voice. The sentence repetition task examined production of complex sentences with relative clauses, adverbial clauses, and conditionals. Every

section of the E-MST had a set of one or two demonstration items that were used depending on the child's need for training. Some children needed this practice to understand that the sentences targeted for repetition were supposed to be repeated, not commented on. However, most children were able to complete the two sections of the E-MST without difficulty.

**Testing and scoring procedures—**The children were tested at their schools by trained bilingual research assistants. During testing, the child was addressed in only one language: English. Children who were bilingual were tested in Spanish on a different day by a different examiner. About one half of all tests were independently scored by a second rater to achieve at least 90% item-by-item agreement. Reliability was based on audiotapes and responses written by the examiners. This was necessary because data collection spanned a few years and involved the work of several raters. To ensure that the scoring procedures were followed uniformly across data collection sites, all the raw data were coded and analyzed at the California site. Each item was scored as correct, incorrect, or not applicable to account for cases of unintelligible or code-switched responses to be excluded from the analyses. Although this procedure may have overestimated the scores of a few children, a conservative approach was deemed necessary, given the experimental nature of the instrument.

On the basis of the pilot studies, alternative responses were provided to maximize scoring efficiency. For passives, correct alternatives included *be*–passives (e.g., *is* + past participle, *is* + *being* + past participle, *was* + past participle, *was* + *being* + past participle) and *get*–passives (e.g., *is* + *getting* + past participle, *was* + *getting* + past participle, and *got* + past participle). Previous research has shown that young children (3;6–5;5) both use and comprehend these two types of auxiliaries (i.e., *be* and *get*) as verbal passives rather than adjectival passives (Guasti, 2002). *Get*–passives are usually used with action verbs and even when they are not accompanied by a *by*–phrase, they are interpreted as verbal passives rather than adjectival passives (Leonard et al., 2006). On the basis of the pilot work described earlier, responses to the past-tense items using simple past tense (i.e., washed) or past progressive (i.e., was washing) were scored as correct. Reduplication errors such as “washeded” and over-regularizations such as “broked” were coded as incorrect. An E-MSTscore was derived based on the number of correct items divided by the total correct and incorrect items.

**Analysis procedures—**Linear discriminant analyses evaluating group classifications (i.e., children with TLD; children with LI) were performed on the percentage correct scores. This was accomplished in two phases. In Phase 1 (exploratory discriminant analyses), the test scores of 56 children (i.e., 28 SW EP with TLD and 28 SW EP with LI, matched by age) were entered into a discriminant analysis. In Phase 2, two confirmatory discriminant analyses were conducted on the remaining 55 children. Using the discriminant coefficients derived from the exploratory analysis, the first confirmatory analysis was based on EDB children (i.e., 10 SW EDB children with TLD and 10 SW EDB children with LI, matched by age). The second confirmatory analysis was based on children sampled from the Northeast region of the United States (i.e., 21 NE children with TLD and 14 NE children with LI). The goal was to determine if the discriminant functions derived from the exploratory data sets were able to predict group membership with independent samples of children from these backgrounds. Thus, in the confirmatory analysis, a discriminant function score for each case is manually calculated by multiplying the original test score by the coefficient generated by the exploratory discriminant functions and then adding a constant (Tabachnick & Fidell, 1989). These discriminant function scores were used to classify cases into groups. A case was classified into the TLD group if its discriminant function score was above zero and into the LI group if the discriminant function score was below zero.

## Results

Table 5 displays the means and standard deviations of the TLD and LI children across SW EP, SW EDB, and NE groups. A comparison based on the EP children with TLD showed no E-MST score differences between the children sampled in California and Texas, $t(26) = 1.311$, $p = .201$, $d = 0.48$. Because it was reasonable to assume that these groups were equivalent, the two sites were combined. Before running the discriminant function analyses, however, it was important to ensure that homogeneity of variance could be assumed. A Box's Test of equality of variance–covariance matrices indicated a significant difference in equality of variance–covariance matrices (Box's $M = 4.438$, $p = .037$). Violations of equality of variance were corrected using arcsine transformations of the E-MST scores (Box's $M = 0.8671$, $p = .358$). The discriminant analysis was then run on arcsine scores. The exploratory analysis yielded a significant squared canonical correlation of .74, $p = .000$. These results indicate large and significant associations between E-MST– weighted scores and group membership. The test scores classified correctly 82.1% of the SW EP children with LI and 89.3% of their peers with TLD.

Using the derived discriminant coefficients (constant = −2.856; coefficient = 4.350), the cutoff score for this group was calculated to be .61. This cutoff score was used to evaluate the separate sample of Southwestern children who were English-dominant bilingual (SW EDB). The E-MST was able to classify correctly 80% of the SW EDB children with LI and 90% of their peers with TLD. However, when the cutoff was used with the NE sample, the measure classified correctly 85.7% of the children with LI but only 61.9% of the children with TLD. Although most of the LI children from this region were correctly classified, the E-MSTscore was not sufficiently accurate to rule out the disorder in children without LI (see Table 5). There was some overlap in the scores of the two groups. The scores of the TLD children who were misclassified ranged from .60 to .15, and the scores of the children with LI ranged from .52 to .02. Thus, even with a lower cutoff score of .53, specificity was still unacceptable (71.4%). Although the NE children as a group were older than the SW groups, their performance was significantly lower, and as a result, they were more likely to be misclassified as impaired.

There appeared to be specific forms responsible for the low specificity of the E-MST with the NE children. An error analysis comparing the 38 SW children with TLD and the 21 NE children with TLD indicated that the NE TLD children had significantly lower scores for possessives, $t(28.8) = 2.12$, $p = .043$, third-person singular, $t(29.3) = 2.13$, $p = .041$, negatives, $t(57) = 2.35$, $p = .022$, and passives, $t(57) = 3.69$, $p = .001$, with large effect sizes for all of these forms and the largest effect size for passives (see Table 6). Past tense, plurals, and copula were also lower, but they appeared to be less problematic than the other forms. These findings are consistent with previous reports describing optional use of possessive nouns and plurals (e.g., “mother shoes,” “two puppy”) as well as optional use of auxiliaries in negation (e.g., “no eat ice cream” for “don't eat ice cream”) in PRE speakers (Wolfram, 1974; Wolfram & Schilling-Estes, 1998a; Zentella, 1997). Thus, performance on these three grammatical targets might reflect dialectal differences. To address this possibility, a modified scoring system was attempted not to penalize errors on these grammatical forms. The test scores of the NE children were reviewed; all dialect-based responses to possessives, plurals, and negation items were re-scored as correct; and the E-MST score was recalculated on the basis of the new scores for these items. However, the procedure did not improve specificity to an acceptable level (i.e., only 15 of 21 children, or 71.4%, were correctly classified). For example, 50% of the misclassified TLD children omitted the third-person singular and past tense morphemes, and the remaining children were not consistent in the use of these features. In addition, the modified scoring procedure resulted in an unacceptable sensitivity rate (i.e., only 10 of 14, or 71.4%, of children were

correctly classified). Although the mean scores of both groups increased, the overlap between the groups remained, yielding inadequate classification accuracy.

## Discussion

The results based on the samples obtained with the Southwestern participants showed that the E-MST can be clinically useful in identifying LI in children who are exposed to Spanish or who are EDB speakers. On the basis of Plante and Vance's (1994) guidelines, measures with accuracy above 90% are considered to have good discrimination, and those with 80%–89% are considered fair. The E-MST was able to identify SW EP children with LI with fair sensitivity and good specificity. Although higher sensitivity and specificity rates have been reported on the basis of children who are in clinicians' caseloads at the time of the testing, it is important to note that sensitivity and specificity may be enhanced when treatment status (as opposed to diagnostic status) is used to identify children with LI. Children who are receiving treatment are more likely to obtain poor test scores compared with children who are not receiving treatment (Ellis Weismer et al., 2000). Many of the children with LI 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 may approximate the true diagnostic accuracy of the E-MST measure by examining the validity of the E-MST with a broad range of children.

The findings corroborate previous research with monolingual children with SLI. A grammatical measure that pays close attention to verb morphology in addition to other grammatical forms has good clinical value with English-speaking children, even if they are exposed to or use Spanish. The first confirmatory analysis applied the cutoff score derived from the samples of EP children to the sample of bilingual children whose English was their best language (i.e., English dominant). The results showed that the bilingual children were not more likely to be misclassified by the E-MST than their EP peers. Similar proportions of monolingual and bilingual children were correctly classified as TLD and LI with the application of the cutoff score. This finding also applied to the classification of the NE children. Based on parent and teacher questionnaires, all of the NE misclassified children had minimal or no Spanish use and proficiency. They were not able to produce a Spanish narrative and could not complete a Spanish morphosyntax task. These findings are consistent with previous research showing no significant differences between bilingual children and those with proficiency in only one language on language processing tasks (Gutiérrez-Clellen, Calderón, & Ellis Weismer, 2004). Again, these results are based on children who are dominant in the language tested (in this case, English). Differences between monolinguals and bilinguals are likely to exist if children are tested and compared in their weaker language.

The discriminant accuracy of the E-MST based on the SW children (i.e., children from Southern California and Texas) may apply to other language communities in the country. However, the study showed that for children who may be learning their English in contact with AAE speakers or who may exhibit features typical of PRE, such as those found in the NE children, the E-MST may not reach adequate specificity. The error analysis indicated that many of the same items that identified the disorder in the affected children penalized the children with TLD and that when these items were excluded, the measure could not classify children correctly. On the basis of these findings, one may conclude that the low specificity of the measure may not be related to dialectal variation but to the language status of the participants. However, the first interpretation is not supported by the results of the re-scoring. First, although some dialect features were re-scored as correct, not all structures that could be possibly affected by dialect could be excluded from the E-MST score because the measure needed to retain a minimum number of items to be usable. Second, the re-scoring

did not result in a distinct profile of AAE and/or PRE dialect that could help differentiate dialect. The analysis demonstrated that surface characteristics of nonmainstream dialects tend to overlap with Standard English SLI characteristics. These observations are consistent with findings reported by Oetting and McDonald (2001) based on comparisons between speakers of South AAE with and without SLI.

A second interpretation for the low specificity found with the NE sample is the possibility that some of the children classified as TLD actually had LI. This interpretation, however, is not supported by the stringent procedure used to rule out LI during the identification process. As was described earlier, children with TLD had no reported concerns by parent or teacher and met the criterion based on their use of mainstream English grammatical utterances in their spontaneous language. Children with parent or teacher concern who demonstrated no grammatical problems in their spontaneous language were excluded. Similarly, children without parent or teacher concerns but who exhibited a significant proportion of ungrammatical utterances were excluded, as well. On the basis of these results and observations, it is reasonable to conclude that the test items continued to penalize dialectal variation in children who were exposed to AAE or PRE, in spite of the use of re-scoring adaptations to account for certain dialect features. The E-MST may not be accurate to rule out the presence of the disorder in speakers of AAE or PRE. For these children, the use of a dialect screening such as the Diagnostic Evaluation of Language Variation–Screening Test (Seymour, Roeper, & de Villiers, 2003a) is recommended. If the child exhibits evidence of dialectal features, a measure such as the Diagnostic Evaluation of Language Variation–Criterion Referenced (Seymour, Roeper, & de Villiers, 2003b) in conjunction with other language measures may increase diagnostic accuracy in clinical assessments. For example, measures of spontaneous language such as mean length of communication units, use of complex syntax, and number of different words in addition to responses to *wh*- questions and probes of active and passive sentence constructions were found to have excellent sensitivity and specificity in a study of African American young children with and without LI (Craig & Washington, 2000). The Diagnostic Evaluation of Language Variation-Criterion Referenced has a passive sentences subtest, but it is a comprehension task. Therefore, it may not elicit the types of dialectal features found in the passives section of the E-MST. However, with modest correlations, there appears to be some association between *–ed* use and comprehension of complex passives in speakers of both Standard English and AAE dialects. Children who demonstrated mastery of complex passives were likely to be those who used *–ed* (Pearson & Roeper, 2004). However, future studies with Latino speakers will be needed to demonstrate the clinical accuracy of these measures with this population. If clinical observations, family interviews, and/or responses to the Diagnostic Evaluation of Language Variation–Screening Test indicate no evidence of dialect use, then the E-MST can be administered because of its fair sensitivity for English-only or English-dominant speakers.

The replication with the NE children was based on a small sample of Caribbean children. Further research is needed to examine the language performance of these and other Latino groups who are raised in AAE- and PRE-speaking communities. These children may exhibit the dialect features discussed above, even though they may not share the same country of origin or background. For example, data collected on one Mexican American child with TLD who resided in the same geographical area (but who was not included in this study) indicated a score below the cutoff score, even though the child was not of Caribbean origin. Because of the growing influx of migrant workers of Mexican origin to the NE area, it is likely that many children from these backgrounds will demonstrate AAE or PRE features in their English, as well. Thus, it is critical to consider the level of dialect use in the community when planning a language evaluation using the E-MST. The grammatical differences found for children who used AAE or PRE features in the present study and the limited specificity

of the grammatical measure with this dialect group are likely to be corroborated using other language instruments, such as the TOLD-P (Newcomer & Hammill, 1997) or the Test of Early Grammatical Impairment (Rice & Wexler, 2001), because these measures include many of the grammatical forms included in the E-MST. Thus, these grammatical measures should be used with caution and in combination with other measures because they may also result in inadequate rates of misclassification with nonmainstream English dialect speakers.

In summary, the present study is the first validation of an E-MST for the identification of children with LI in Latino English-speaking children who are exposed to Spanish or who are EDB. This measure was developed targeting forms that could differentiate children with TLD from children with LI while examining the potential impact of Spanish use and nonmainstream English dialect such as AAE or PRE. The results showed that the E-MST could be used with fair classification accuracy with SW children who are EDB. However, for NE children who may be using AAE or PRE dialect features, alternative assessment tools will be necessary to rule out the presence of the disorder. The present research is a preliminary step in establishing the accuracy of language measures for the identification of bilingual children with LI. Future studies with the E-MST in combination with language processing tasks (e.g., Gutiérrez-Clellen et al., 2004) and across a broader age range will be needed to gain a better understanding of the characteristics of the disorder and its manifestations in children who are learning two languages.

## Acknowledgments

This project was partially supported by Grant 1-DC-8-2100 from the National Institute on Deafness and Other Communication Disorders and by Grant 5 R25 GM58906-06 from the National Institute of General Medical Sciences San Diego State University Minority Biomedical Research Support Program. We are grateful to Lisa Bedore, Janet Calderón, Brian Goldstein, Aquiles Iglesias, Ellen Stubbe Kester, and Elizabeth Peña 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

Anastasi, A.; Urbina, S. Psychological testing. Prentice-Hall; Upper Saddle River, NJ: 1997.

Bedore LM, Leonard LB. Specific language impairment and grammatical morphology: A discriminant function analysis. Journal of Speech, Language, and Hearing Research. 1998; 41:1185–1192.

Bishop D, Butterworth G. A longitudinal study using the WPPSI and WISC-R with an English sample. British Journal of Educational Psychology. 1979; 49:156–168.

Carrow-Woolfolk, E. Test for Auditory Comprehension of Language. 3rd ed.. Pro-Ed; Austin, TX: 1999.

Conti-Ramsden G, Botting N, Faragher B. Psycholinguistic markers for specific language impairment (SLI). Journal of Child Psychology and Psychiatry. 2001; 42:741–748. [PubMed: 11583246]

Craig HK, Washington JA. An assessment battery for identifying language impairments in African American children. Journal of Speech, Language, and Hearing Research. 2000; 43:366–379.

Dawson, JI.; Stout, CE.; Eyer, JA. Structured Photographic Expressive Language Test. 3rd ed.. Janelle Publications; DeKalb, IL: 2003.

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

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

Fought, C. Talkin' with mi gente (Chicano English). In: Wolfram, W.; Ward, B., editors. American voices: How dialects differ from coast to coast. Blackwell Publishing; Malden, MA: 2006. p. 233-237.

Friedenberg, L. Psychological testing: Design, analysis and use. Allyn & Bacon; Needham Heights, MA: 1995.

Gathercole, VCM. Command of the mass/count distinction in bilingual and monolingual children: An English morphosyntactic distinction. In: Oller, DK.; Eilers, RE., editors. Language and literacy in bilingual children. Multilingual Matters; Tonawanda, NY: 2002a. p. 175-206.

Gathercole, VCM. Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In: Oller, DK.; Eilers, RE., editors. Language and literacy in bilingual children. Multilingual Matters; Tonawanda, NY: 2002b. p. 207-219.

Gathercole, VCM. Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomena in English and Spanish. In: Oller, DK.; Eilers, RE., editors. Language and literacy in bilingual children. Multilingual Matters; Tonawanda, NY: 2002c. p. 255-280.

Guasti, MT. Language acquisition: The growth of grammar. The MIT Press; Cambridge, MA: 2002.

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.

Gutiérrez-Clellen VF, Calderón J, Ellis Weismer S. Verbal working memory in bilingual chilldren. Journal of Speech, Language, and Hearing Research. 2004; 47:863–876.

Gutiérrez-Clellen, VF.; Erickson, A. “The chair kicked the cow”: Passive constructions in Latino children. Paper presented at the ASHA National Convention; Washington, DC. Nov. 2000

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

Gutiérrez-Clellen VF, Kreiter J. Understanding child bilingual acquisition using parent and teacher reports. Applied Psycholinguistics. 2003; 24:267–288.

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 V, Restrepo MA, Simon-Cereijido G. Evaluating the discriminant accuracy of a grammatical measure with Spanish-speaking children. Journal of Speech, Language, and Hearing Research. 2006; 49:1209–1223.

Hakuta, K. Mirror of language. Basic Books; New York: 1986.

Hammer CS, Pennock-Roman M, Rzasa S, Tomblin JB. An analysis of the Test of Language Development–Primary for item bias. American Journal of Speech-Language Pathology. 2002; 11:274–284.

Hammill, DD.; Newcomer, PL. Test of Language Development–Intermediate. 3rd ed.. Pro-Ed; Austin, TX: 1997.

Hannah, EP.; Gardner, JO. Hannah-Gardner Test of Verbal and Nonverbal Language Functioning. Lingua Press; Northridge, CA: 1978.

Johnston, JR., editor. Cognitive prerequisites: The evidence from children learning English. Erlbaum; Hillsdale, NJ: 1985.

Laing SP, Kamhi A. Alternative assessment of language and literacy in culturally and linguistically diverse populations. Language, Speech, and Hearing Services in Schools. 2003; 34:44–55.

Leonard, LB. Children with specific language impairment. The MIT Press; Cambridge, MA: 1998.

Leonard LB, Wong AM-Y, Deevy P, Stokes SF, Fletcher P. The production of passives by children with specific language impairment: Acquiring English or Cantonese. Applied Psycholinguistics. 2006; 27:267–299.

Mayer, M. Frog, where are you?. Dial Press; New York: 1969.

Mayer, M. Frog on his own. Dial Press; New York: 1973.

Mayer, M. Frog goes to dinner. Dial Press; New York: 1974.

Mayer, M. One frog too many. Dial Press; New York: 1975.

Newcomer, PL.; Hammill, DD. Test of Language Development–Primary. 2nd ed.. Pro-Ed; Austin, TX: 1991.

Newcomer, PL.; Hammill, DD. Test of Language Development–Primary. 3rd ed.. Pro-Ed; Austin, TX: 1997.

Norris MK, Juarez MJ, Perkins MN. Adaptation of a screening test for bilingual and bidialectal populations. Language, Speech, and Hearing Services in Schools. 1989; 20:381–390.

Oetting JB, McDonald JL. Nonmainstream dialect use and specific language impairment. Journal of Speech, Language, and Hearing Research. 2001; 44:207–223.

Ornstein-Galicia, J., editor. Form and function in Chicano English. Newbury House; Rowley, MA: 1981.

Owens, RE. Language disorders: A functional approach to assessment and intervention. Merrill/ Macmillan; Columbus, OH: 1991.

Pearson, BZ. Narrative competence among monolingual and bilingual school children in Miami. In: Oller, DK.; Eilers, RE., editors. Language and literacy in bilingual children. Multilingual Matters; Tonawanda, NY: 2002. p. 135-174.

Pearson, BZ.; Roeper, T. Learn-ability and triggers: Obligatory versus optional triggers for the passive in two dialects of English and in language impairment. Paper presented at the Linguistic Society of America Annual Meeting; Boston, MA. Jan. 2004

Peña, ED.; Gutiérrez-Clellen, VF.; Iglesias, A.; Goldstein, B.; Bedore, LM. Bilingual English Spanish Assessment (BESA). n.d.

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

Perona K, Plante E, Vance R. Diagnostic accuracy of the Structured Photographic Expressive Language Test: Third Edition (SPELT-3). Language, Speech, and Hearing Services in Schools. 2005; 36:103–115.

Plante E, Vance R. Selection of preschool language tests: A data-based approach. Language, Speech, and Hearing Services in the Schools. 1994; 25:15–24.

Qi, CH. Language performance of low-income, Hispanic preschool children on the Preschool Language Scale–4. Paper presented at the Symposium on Research in Child Language Disorders; Madison, WI. Jun. 2005

Qi CH, Kaiser AP, Milan SE, Yzquierdo Z, Hancock TB. The performance of low-income, African American children on the Preschool Language Scale–3. Journal of Speech, Language, and Hearing Research. 2003; 46:576–590.

Ramsey Musselwhite C. Pluralistic assessment in speech-language pathology: Use of dual norms in the placement process. Language, Speech, and Hearing Services in Schools. 1983; 14:29–37.

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

Restrepo MA, Schwanenflugel PJ, Blake J, Neuharth-Pritchett S, Cramer SE, Ruston HP. Performance on the PPVT-III and the EVT: Applicability of the measures with African American and European American preschool children. Language, Speech, and Hearing Services in Schools. 2006; 37:17–27.

Rice, M.; Wexler, K. Test of Early Grammatical Impairment. The Psychological Corporation; San Antonio, TX: 2001.

Semel, E.; Wiig, E.; Secord, W. Clinical Evaluation of Language Fundamentals. 4th ed.. The Psychological Corporation; San Antonio, TX: 2003.

Seymour, HN.; Roeper, TW.; de Villiers, J. Diagnostic Evaluation of Language Variation–Screening Test. Harcourt Assessment; San Antonio, TX: 2003a.

Seymour, HN.; Roeper, TW.; de Villiers, J. Diagnostic Evaluation of Language Variation (DELV-Criterion Referenced). Harcourt Assessment; San Antonio, TX: 2003b.

Spaulding TJ, Plante E, Farinella KA. Eligibility criteria for language impairment: Is the low end of normal always appropriate? Language, Speech, and Hearing Services in Schools. 2006; 37:61–72.

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

van der Lely HKJ. SLI in children: Movement, economy and deficits in the computational syntactic system. Language Acquisition. 1998; 7:211–224.

Washington JA, Craig HK. Performances of low-income, African American preschool and kindergarten children on the Peabody Picture Vocabulary Test–Revised. Language, Speech, and Hearing Services in Schools. 1992; 23:329–333.

Washington JA, Craig HK. Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. Language, Speech, and Hearing Services in Schools. 1999; 30:75–82.

Wiig, E.; Secord, W.; Semel, E. Clinical Evaluation of Language Fundamentals–Preschool. The Psychological Corporation; San Antonio, TX: 1992.

Wilcox KA, Aasby SM. The performance of monolingual and bilingual Mexican children on the TACL. Language, Speech, and Hearing Services in Schools. 1988; 19:34–40.

Wolfram, W. Sociolinguistic aspects of assimilation: Puerto Rican English in New York City. Center for Applied Linguistics; Arlington, VA: 1974.

Wolfram, W.; Schilling-Estes, N. American English: Dialects and variation. Blackwell Publishing; Malden, MA: 1998a.

Wolfram, W.; Schilling-Estes, N. Tri-ethnic dialect situations. In: Wolfram, W.; Schilling-Estes, N., editors. American English. Blackwell Publishing; Malden, MA: 1998b. p. 181-182.

Zentella, AC. The language situation of Puerto Ricans. In: McKay, SL.; Wong, S-LC., editors. Language diversity: Problem or resource?. Newbury House; New York: 1988. p. 140-165.

Zentella, AC. Growing up bilingual: Puerto Rican children in New York. Blackwell Publishing; Malden, MA: 1997.

Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool Language Scale-4. The Psychological Corporation; San Antonio, TX: 2001.

## Table 1

Percentage of families of the children with typical language development (TLD) and language impairment (LI) by educational level, region ($N$= 111), and eligibility for lunch program.

| Characteristic | TLD ($N$ = 59) | LI ($N$ = 52) |
|---|---|---|
| Age (years;months) | 5;0 | 5;0 |
| Educational level in the home | | |
| Primary education + some secondary | 17% (8 of 48) | 22% (8 of 37) |
| High school graduate + some college experience | 69% (33 of 48) | 73% (27 of 37) |
| College graduate | 14% (7 of 48) | 5% (2 of 37) |
| Region | | |
| Southwest (Mexican origin) | 64% (38 of 59) | 73% (38 of 52) |
| Northeast (Puerto Rican and Dominican origin) | 36% (21 of 59) | 27% (14 of 52) |
| Lunch eligibility | | |
| Regular | 37% | 36% |
| Free or reduced | 63% | 64% |

*Note*. Fifteen Southwest and 11 Northeast participants had incomplete questionnaire data.

## Table 2

Means and standard deviations of amount of language input, and ratings of use and proficiency for the Southwestern English-only proficient participants ($N = 56$).

| Characteristic | Language ability | | | |
|---|---|---|---|---|
| | Typical ($N = 28$) | Impaired ($N = 28$) | *p* | *d* |
| Age (years;months) | 4;10 | 4;10 | .91 | 0.03 |
| Proportion of English input at home | 0.86 (0.11) | 0.74 (0.19) | .02 | 0.78 |
| Proportion of Spanish input at home | 0.14 (0.11) | 0.26 (0.19) | .02 | 0.78 |
| Parents' rating of use of English | 4.0 (0.0) | 3.6 (0.94) | .07 | 0.62 |
| Parents' rating of use of Spanish | 1.6 (0.75) | 1.6 (1.3) | .99 | 0.00 |
| Parents' rating of proficiency in English | 3.8 (0.40) | 3.7 (0.57) | .48 | 0.20 |
| Parents' rating of proficiency in Spanish | 1.3 (0.79) | 1.4 (1.2) | .83 | 0.10 |
| Teachers' rating of use of English | 3.8 (0.39) | 3.6 (0.56) | .32 | 0.41 |
| Teachers' rating of use of Spanish | 0.5 (0.79) | 0.69 (0.85) | .63 | 0.78 |
| Teachers' rating of proficiency in English | 3.8 (0.44) | 3.2 (0.88) | .02 | 0.83 |
| Teachers' rating of proficiency in Spanish | 1.0 (1.5) | 0.86 (0.95) | .77 | 0.12 |
| Proportion of input at school in English | 0.63 (0.42) | 0.80 (0.27) | .25 | 0.49 |
| Proportion of input at school in Spanish | 0. 37 (0.42) | 0.20 (0.27) | .25 | 0.49 |

## Table 3

Means and standard deviations of amount of language input as well as ratings of use and proficiency for the Southwestern English-dominant bilingual participants ($N = 20$).

| Characteristic | Language ability | | *p* | *d* |
|---|---|---|---|---|
| | Typical ($N = 10$) | Impaired ($N = 10$) | | |
| Age (years;months) | 5;1 | 5;1 | .83 | 0.09 |
| Proportion of English input at home | 0.57 (0.15) | 0.45 (0.11) | .054 | 0.91 |
| Proportion of Spanish input at home | 0.43 (0.15) | 0.55 (0.11) | .054 | 0.91 |
| Parents' rating of use of English | 3.9 (0.32) | 4.0 (0.0) | .34 | 0.44 |
| Parents' rating of use of Spanish | 3.6 (0.52) | 3.3 (0.95) | .39 | 0.39 |
| Parents' rating of proficiency in English | 3.9 (0.32) | 3.6 (0.70) | .24 | 0.55 |
| Parents' rating of proficiency in Spanish | 3.7 (0.48) | 3.1 (1.10) | .14 | 0.71 |
| Teachers' rating of use of English | 2.83 (1.53) | 3.12 (0.83) | .65 | 0.27 |
| Teachers' rating of use of Spanish | 2.33 (1.50) | 1.92 (1.36) | .60 | 0.30 |
| Teachers' rating of proficiency in English | 3.83 (0.41) | 3.0 (0.92) | .05 | 1.11 |
| Teachers' rating of proficiency in Spanish | 3.0 (0.82) | 2.71 (1.38) | .72 | 0.24 |
| Proportion of input at school in English | 0.44 (0.35) | 0.53 (0.37) | .67 | 0.25 |
| Proportion of input at school in Spanish | 0.56 (0.35) | 0.47 (0.37) | .67 | 0.22 |

## Table 4

Means and standard deviations of amount of language input as well as ratings of use and proficiency for the Northeastern participants ($N = 35$).

| Characteristic | Language ability | | *p* | *d* |
|---|---|---|---|---|
| | Typical ($N = 14$) | Impaired ($N = 14$) | | |
| Age (years;months) | 6;1 | 6;1 | .98 | 0.01 |
| Proportion of English input at home | 0.61 (0.35) | 0.60 (0.34) | .96 | 0.03 |
| Proportion of Spanish input at home | 0.39 (0.35) | 0.40 (0.34) | .96 | 0.03 |
| Parents' rating of use of English | 3.52 (0.62) | 3.25 (0.70) | .32 | 0.43 |
| Parents' rating of use of Spanish | 1.82 (1.28) | 2.12 (1.12) | .58 | 0.24 |
| Parents' rating of proficiency in English | 3.70 (0.46) | 3.0 (0.75) | .009 | 1.23 |
| Parents' rating of proficiency in Spanish | 1.76 (1.30) | 2.12 (1.12) | .51 | 0.29 |
| Teachers' rating of use of English | 3.91 (0.33) | 3.73 (0.49) | .27 | 0.48 |
| Teachers' rating of use of Spanish | 0.56 (0.62) | 0.76 (0.94) | .51 | 0.28 |
| Teachers' rating of proficiency in English | 3.26 (1.48) | 3.45 (0.82) | .70 | 0.15 |
| Teachers' rating of proficiency in Spanish | 0.58 (1.17) | 1.22 (1.56) | .25 | 0.48 |
| Proportion of input at school in English | 0.83 (0.25) | 0.85 (0.24) | .84 | 0.08 |
| Proportion of input at school in Spanish | 0.10 (0.12) | 0.14 (0.24) | .57 | 0.22 |

**Table 5**

Means and standard deviations of English Morphosyntax Test (E-MST) scores and percentages of children correctly classified as having typical language development (TLD) or language impairment (LI) across Southwestern English-only proficient (SW EP), Southwestern English-dominant bilingual (SW EDB), and Northeastern (NE) samples.

| | TLD | | LI | | | |
|---|---|---|---|---|---|---|
| **Classification** | ***M*** | ***SD*** | ***M*** | ***SD*** | **Sensitivity** | **Specificity** |
| SW EP | 0.77 | 0.14 | 0.38 | 0.17 | 82.1% (23/28) | 89.3% (25/28) |
| SW EDB | 0.79 | 0.13 | 0.39 | 0.22 | 80% (8/10) | 90% (9/10) |
| NE | 0.64 | 0.20 | 0.38 | 0.20 | 85.7% (12/14) | 61.9 % (13/21) |

*J Speech Lang Hear Res*. Author manuscript; available in PMC 2012 June 05.

**Table 6**

Means and standard deviations of E-MST scores by Southwestern and Northeastern children with typical language development.

| | Southwestern (*N* = 38) | | Northeastern (*N* = 21) | | | |
|---|---|---|---|---|---|---|
| **Characteristic** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***p*** | ***d*** |
| Possessives | 0.86 | 0.24 | 0.67 | 0.38 | .043 | 0.64 |
| Third-person singular | 0.75 | 0.25 | 0.55 | 0.38 | .041 | 0.66 |
| Past tense | 0.74 | 0.26 | 0.70 | 0.27 | .657 | 0.15 |
| Plurals | 0.78 | 0.30 | 0.59 | 0.41 | .056 | 0.55 |
| Copula | 0.69 | 0.27 | 0.60 | 0.30 | .253 | 0.32 |
| Negatives | 0.86 | 0.27 | 0.67 | 0.35 | .022 | 0.63 |
| Passives | 0.70 | 0.33 | 0.37 | 0.34 | .001 | 0.99 |

*Note.* E-MST scores represent the proportion of correct items produced on the cloze task. Results were compared using an unequal variance *t* test.

*J Speech Lang Hear Res*. Author manuscript; available in PMC 2012 June 05.