AUTOMATED GRAMMATICAL TAGGING OF LANGUAGE SAMPLES FROM SPANISH-SPEAKING CHILDREN LEARNING ENGLISH

by

Nikki Redd

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Audiology and Speech-Language Pathology

Brigham Young University

August 2006

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Nikki Redd

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Ron W. Channell, Chair
Date	Martin Fujiki
Date	Shawn Nissen

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Nikki Redd in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Ron W. Channell
Chair, Graduate Committee

Accepted for the Department

Ron W. Channell
Graduate Coordinator

Accepted for the College

K. Richard Young
Dean, David O. McKay School of Education

ABSTRACT

AUTOMATED GRAMMATICAL ANALYSIS OF LANGUAGE SAMPLES FROM SPANISH-SPEAKING CHILDREN LEARNING ENGLISH

Nikki Redd

Department of Audiology and Speech-Language Pathology

Master of Science

Research has demonstrated that automated grammatical tagging is fast and accurate for both English and Spanish child language, but there has been no research done regarding its accuracy with bilingual children. The present study examined this topic using English and Spanish language samples taken from 254 children living in the United States. The subjects included school-aged children enrolled in public schools in the United States in grades 2, 3, or 5. The present study found high automated grammatical tagging accuracy scores for both English (M = 96.4%) and Spanish (M = 96.8%). The study suggests that automated grammatical analysis has potential to be a valuable tool for clinicians in the analysis of the language of bilingual children.

ACKNOWLEDGMENTS

I could not have completed this thesis on my own. I would like to thank my committee chair, Dr. Ron Channell, for his constant support and guidance. I would like to thank my committee members, Dr. Martin Fujiki and Dr. Shawn Nissen for their input and advice. I express my appreciation for my family and friends for their support during my thesis and throughout my education.

TABLE OF CONTENTS

Page

LIST OF TABLES

Introduction

Researchers and clinicians would find software for automated (part-of-speech) grammatical tagging of child language useful if the software proved to be fast and accurate. The present study evaluates English and Spanish software for bilingual (English-Spanish) child language. Previous research has shown accuracy levels of up to 95% for automated grammatical tagging in English and just slightly lower levels (averaging 92%) for Spanish automated grammatical tagging. The present study evaluates the effects that unique differences in bilingual language development may have on the accuracy of both English and Spanish software for automated grammatical tagging.

Existing software for automated grammatical tagging of English is both fast and accurate. Channell and Johnson (1999) evaluated the accuracy of automated grammatical tagging for language samples of 30 normally developing English-speaking children. They found accuracy levels ranging from 92.9% to 97.4% on a word-by-word basis. Long (2001) compared five manual and automated analyses across 30 participants including students and clinicians with experience in the analyses they performed. Analyses included basic utterance and morpheme counts, comprehensive structural analyses, and interactional analyses. In this study, Long found that manual grammatical analyses generally took more time and were no more accurate than analyses done with the aid of a computer.

The study of automated grammatical tagging for Spanish is still in its infancy. Wilson (2005) evaluated the accuracy of a Spanish version of automated grammatical tagging for child language. Twenty-four language samples taken from children in Mexico

that spoke only Spanish were evaluated. Results of this study showed word-level accuracy levels ranging from 87.7% to 95.6%. This study showed that Spanish software for automated grammatical tagging is a relatively fast and accurate way to analyze Spanish-speaking children's language.

Bilingual language development is an area of significant interest in public schools because of the increasing number of bilingual children, particularly children that speak Spanish at home and are learning English in school. Bilingual children present unique challenges to speech language pathologists because of their unique language development. Combined research of the accuracy of both English and Spanish automated analysis for bilingual child language would be relevant and beneficial for clinicians practicing in the United States that serve bilingual populations. The present study extends current research on English and Spanish automated grammatical tagging to bilingual (English-Spanish) child language. It will use manual and automated grammatical tagging for both Spanish and English language samples from bilingual children. The manual and computer-generated tags will be compared and the accuracy of automated grammatical tagging will be determined for both, allowing conclusions to be drawn regarding the efficiency and accuracy of current software.

Review of Literature

As the United States has become more culturally diverse, the identification of language disorders in bilingual children has become a significant concern for speech-language pathologists. Clinical language assessment of bilingual Spanish-English-speaking children is problematic because of the possibility of overrepresentation or underrepresentation of bilingual students in therapy. The difficulty in identifying language disorders in bilingual children lies in the unique characteristics of bilingual language development that differ so greatly from the characteristics of monolingual child language. The purpose of this review is to discuss current research in formal and informal assessment techniques for bilingual Spanish-English-speaking child language.

Formal Assessment

The lack of materials, such as standardized tests, for bilingual language assessment presents unique challenges for speech language pathologists when assessing bilingual children. Appropriate and non-biased tests are required in order to perform valid evaluations of bilingual child language. A current practice for assessing the language skills of bilingual children involves translating standardized English tests into the child's native language. Although these tests are presented in the child's native language, they are not norm-referenced on bilingual children and therefore cannot reflect the unique aspects of bilingual language development (Kester & Peña, 2002). The current methods the majority of speech-language pathologists use for bilingual child language analysis are insufficient when intervention decisions are based solely on bilingual children's performances on standardized tests. If formal tests are to be used in bilingual language assessment, analysis of both languages is necessary to determine a child's linguistic

comprehension and performance. Speech-language pathologists must also remember that results of standardized tests must be interpreted with caution, as most standardized tests are not norm-referenced on bilingual populations.

Validity of the Spanish Preschool Language Scale-3. There are several standardized tests that have been norm-referenced on Spanish-speaking populations, but there is little research whether or not these tests are valid for bilingual populations. Restrepo and Silverman (2001) evaluated the validity of the *Spanish Preschool Language Scale-3* (SPLS-3; Zimmerman, Steiner, & Pond, 1993) for use with bilingual children. Thirty-seven bilingual children between the ages of 4;4 and 6;6 participated in the study. Each child participated in several English and Spanish language measures, including a parent interview, language sampling, and criterion-referenced measures. Eighty-one percent of children in the study scored one *SD* or more below the mean on the SPLS-3 (*M* = 1.52 *SD* below the mean). It was discovered that the test norms were not representative of the study sample. The researchers emphasized the need for tests to focus on Spanish and bilingual language development as well as for clinicians to consider children's different language backgrounds and uses. The researchers found that there was no evidence of concurrent or predictive validity and no evidence of test-retest reliability when administering the SPLS-3 to bilingual children. The use of the SPLS-3 in identifying bilingual children for speech therapy could result in over-identification of children with language impairment.

In addition to the participants' scores on the SPLS-3, researchers analyzed the appropriateness of vocabulary on the SPLS-3 and found that several words within the test were developmentally inappropriate (Restrepo and Silverman, 2001). Furthermore, the

test did not include several areas of language development that would be important in identifying language disorders. Based on the results of this study, researchers concluded that standardized measures, specifically the SPLS-3, are not appropriate for bilingual language analysis. They recommended that until reliable and valid bilingual standardized testing is available, clinicians should use language sample analysis and parent interview to assist with the analysis of language of the bilingual children in their caseloads.

Assessing vocabulary and language exposure. A speech-language pathologist must not only consider bilingualism as a complicating factor in language assessment, but should realize that each bilingual child comes from a different background with varying levels of language exposure in each language. Umbel, Pearson, Fernández, and Oller (1992) completed a study that evaluated the vocabulary of bilingual children in both Spanish and English and the effect of amount of language exposure on vocabulary acquisition. The participants in the study were 105 middle-class English-Spanish bilingual children from first grade in four public schools. Participants were divided into two groups, including children that spoke only Spanish at home (OSH) and children that spoke both English and Spanish at home (ESH). Children from both groups were given two standardized tests including the *Peabody Picture Vocabulary Test—Revised* (PPVT-R; Dunn & Dunn, 1981) for English language analysis and the *Test de Vocabulario en Imágenes Peabody—Adaptación Hispanoamericana* [Peabody Picture Vocabulary Test—Latin American adaptation] (TVIP-H; Dunn, Padilla, Lugo, & Dunn, 1986) for Spanish language analysis. Both the OSH and the ESH groups received scores on the PPVT-R that were significantly lower than the norms. On the TVIP-H, the scores of both the OSH group and the ESH group were similar both to the norms and to each other.

While evaluating the results of the PPVT-R and the TVIP-H, the researchers concluded that most participants had singlet vocabulary, in other words, their vocabulary was distributed between English and Spanish. This study stressed the importance of evaluating bilingual children in both their languages in order to get a more accurate description of the children's full vocabulary. The study also found that experience with language in the home was important for bilingual vocabulary development. This study suggested that clinicians should not only assess bilingual children in both languages, but should also consider the amount of exposure to each language during assessment.

Alternatives to standardized testing. Saenz and Huer (2003) outlined the various alternatives to standardized testing in the assessment of bilingual child language. The alternatives they outlined included renorming tests for different populations, using dynamic assessment rather than formal tests, using other nonstandardized methods of assessment, and modifying standardized tests. Their study evaluated the impact of modifying a particular standardized test, the *Clinical Effectiveness of Language Functioning—Third Edition* (CELF-3; Semel, Wiig, & Secord, 1995). The test was modified for bilingual children by rewording directions, allowing multiple attempts, allowing longer response times, giving more detailed explanations of trial items, and testing above the ceiling.

The participants in the Saenz and Huer (2003) study included 28 children between the ages of 8 and 9. Participants received two versions of the CELF-3, including a modified and an unmodified version. Overall, participants received higher scores on the CELF-3 when test modifications were used. If testing modification is used as an alternative method of assessment, normalization data cannot be used to determine

language ability. However, the authors proposed that testing modification saves time and allows students to perform well when they otherwise would not have. While this modification method cannot give a standardized score, it would give clinicians an idea of areas in which the child had difficulty, even when provided with support.

Standardized versus nonstandardized assessment. Anderson (1996) looked at the cultural bias in standardized testing for bilingual children and attempted to provide an alternative means of analysis. In Anderson's study, 20 four-year-old bilingual children completed two forms of language assessment including the *Structured Photographic Expressive Language Test—Preschool* (SPELT-P, Werner & Kresheck, 1989) and a testing instrument developed by the author, which was a measure performed while the child and a clinician were playing. The testing instrument developed by the author included activities designed to elicit various grammatical targets, all of which were also targeted on the SPELT-P. All 20 participants scored higher on the assessment developed by the author. Because both assessments targeted the same grammatical forms, the results of this study showed that the SPELT-P was unable to elicit grammatical forms that children were able to produce in play. Thus, a natural, more interactive setting proved to be a more appropriate method for obtaining information about preschool bilingual children's language form and function than standardized testing.

Informal Assessment

The assessment of bilingual children also presents unique challenges to speech language pathologists because of bilingual children's unique language development. When assessing bilingual child language, a clinician must understand that a bilingual child is not two monolingual children in one and that the child cannot be separated into

two languages (Goldstein, 2004). A bilingual child's language knowledge may not be complete in one language or the other, but may be complete when both languages are taken into account. A bilingual child may be proficient in their first language (L1) in one context and proficient in their second language (L2) in a different context due to different patterns of use (Gutierrez-Clellan, Restrepo, Bedore, Peña, & Anderson, 2000). For example, if a child speaks Spanish at home and English at school, it is likely that the child will be more proficient with Spanish with every day vocabulary and conversation, but more proficient with English in academic vocabulary and conversation. Because of this knowledge, more comprehensive assessments of bilingual child language, such as language sampling, are helpful in understanding bilingual children's unique linguistic characteristics.

Conceptual versus monolingual scoring. In a recent study, Bedore, Peña, García, and Cortez (2005) evaluated monolingual and conceptual scoring in the classification of typically developing bilingual child language. Monolingual scoring involves assessing one language without incorporating the other whereas conceptual scoring involves considering both languages and scoring language on meaning rather than language choice (Pearson, Fernández, & Oller, 1993).

The Bedore et al. study involved two studies of bilingual children. In Study 1, the participants included 55 typically developing (TD) Hispanic children between the ages of 4;0 and 7;11 living in the United States. The participants were divided among four groups based on percentage of English and Spanish use, including a primarily English-speaking group (PE; spoke English 80% or more of the time), a bilingual English group (BE; spoke English 50-80% of the time), a bilingual Spanish group (BS; spoke Spanish 50-80% of

the time), and a primarily Spanish-speaking group (PS; spoke Spanish 80% or more of the time). The children participated in description tasks that were given three different scores: (a) monolingual scores, which scored nonduplicated words in the target language, (b) total response scores, which included the number of correct items in both languages, and (c) conceptual scores, which scored the number of unique concepts in both languages. Participants in the BS and BE groups produced more errors overall, which showed they lacked necessary vocabulary in their nondominant language. Based on the results of this study, the authors concluded that total and conceptual scoring is more valuable and more appropriate than monolingual scoring in analyzing TD bilingual child language.

In Study 2, 40 TD Hispanic children between the ages of 5;0 and 6;1 were divided into the same language groups as in Study 1. Participants completed receptive and expressive tasks found in three stories and were scored based on the same monolingual and conceptual scoring as used in Study 1. In Study 2, children received higher scores when the conceptual score was used, rather than the monolingual score, which the researchers concluded was the more appropriate score. The results of the Bedore et al. study showed that monolingual scoring and analysis did not offer a complete analysis of bilingual child language. It can therefore be concluded that speech-language pathologists need to consider both languages when assessing bilingual children.

Narrative Discourse. A study by Fiestas and Peña (2004) evaluated the Spanish and English story grammars and story elements of 12 bilingual children between the ages of 4;0 and 6;11 in Texas. Their study also included a brief evaluation of grammaticality of the bilingual children's language. Each child produced four narratives, two in English

and two in Spanish, elicited by two methods of stimuli. The first was a picture description of a typical Mexican-American birthday party and the second was a wordless picture book, *Frog, Where Are You?* (Mayer, 1969). The children's utterances were divided into C-units and assessed for grammaticality. The participants in the study produced similar amounts of grammatical utterances in English and Spanish, with no significant differences between the two languages. The study showed that narrative tasks are effective for eliciting representative language samples for bilingual children between the ages of 4 and 7. The narrative task using the wordless picture book proved to be more effective in assessing overall language proficiency and ability in English and Spanish. The picture description task was not as effective in challenging the children in English, their second language. This study suggests that clinicians should be selective when choosing materials to elicit language samples from their bilingual clients.

Grammatical characteristics of a Spanish-English bilingual with SLI. There is very limited research in the area of bilingual children with SLI. In one study of Spanish-speaking children with SLI, their grammatical morphologies were compared with same-age peers and with younger TD children with similar MLUs (Bedore & Leonard, 2001). Forty-five children were included in the study, all of which were Mexican Spanish-speaking, and 15 of them were identified as having SLI. The children participated in several grammatical probes, including picture naming, sentence completion tasks, and description of events in order to evaluate their production of 14 grammatical morphemes. The morphemes assessed included nine types of verb inflections and five types of noun phrases. Analyses of variance (ANOVA) were computed for each grammatical morpheme, using group (SLI, MLU, age) as the between-subjects variable. There were

significant differences in three of the nine verb inflections, including present third person plural, past third person singular, and the past third person plural. There were also significant differences for all five of the grammatical morphemes associated with noun phrases, including definite articles, indefinite articles, direct object clitic pronouns, noun plural inflections, and adjective agreement inflections. These results showed that the children with SLI not only performed more poorly than their same-age peers, but also performed more poorly than younger MLU-matched peers in several categories. However, the children with SLI and younger MLU-matched children made similar verb inflection errors and noun phrase-related errors.

There is some evidence that difficulties with grammar for monolingual children may be even more severe in bilingual children (Restrepo & Kruth, 2000). Restrepo and Kruth (2000) sought to describe the grammatical characteristics of a bilingual Spanish-English speaking child with SLI and assessed whether these characteristics were different from monolingual children with SLI. The study compared two bilingual Spanish-English speaking children, one with and one without SLI. The authors' attempt was to compare children with similar linguistic backgrounds in order to distinguish among SLI, second language learning, and language loss. They also sought to determine if the grammatical differences of bilingual children differed from those of monolingual children with SLI or second language learners, as reported in the literature. Each child participated in Spanish and English spontaneous language samples. Language sampling was chosen as the method of assessment because Restrepo (1998) found that a combination of parent report and language sampling correctly distinguished between Spanish-speaking children with and without SLI. The language samples were then transcribed and analyzed using MLU,

sentence complexity and type, Brown's morphemes (Brown, 1973), and error type. In the English language sample analyses, the child without SLI used more grammatical forms and had mastered morphological skills better than the child with SLI. The child with SLI had limited use of verb forms, pronouns, prepositions, limited syntactic complexity, and many errors. There were also significant differences in Spanish for both children. The child without SLI had very few grammatical errors and used many different grammatical styles and sentence types. The child with SLI had problems using articles, pronouns, and prepositions. She had a limited number of sentence types and grammatical forms in her utterances. In addition to being different from each other, the data showed that bilingual children with SLI have grammatical difficulties that differ from monolingual children with SLI and ESL learners.

Parent Report Measures. Parent report measures have become an accepted method of assessment for bilingual language development. However, while research has documented the validity of these measures in monolingual English speaking, there is little research regarding their validity for bilingual Spanish- and English-speaking children. Patterson (2000) evaluated the validity of the Spanish-English Vocabulary Checklist (SEVC; Patterson, 1998), a parent report measure developed by the author. Participants in this study included 12 bilingual toddlers between the ages of 21 and 27 months. A parent of each participant completed the SEVC and answered questions regarding their child's language development and exposure to each language. The children then participated in 30-minute language samples with their reporting parent. The correlations of reported vocabulary on the SEVC and observed vocabulary in language assessments was comparable to monolingual English correlations on other parent report measures.

Thus, the SEVC may be considered valid in the assessment of bilingual toddler's language.

In a similar study by Marchman and Sussmann-Martínez (2002), the validity of two parent report measures was analyzed. Participants in the study were 26 TD bilingual toddlers learning Spanish and English. Participants' caregivers completed two report measures including the *MacArthur Communicative Development Inventories* (CDI; Fenson, Dale, Reznick, Thal, Bates, Hartung, et al., 1993) and the *Inventario del Desarrollo de Habilidades Comunicativas: Palabras y Enunciados* (IDHC; Jackson-Maldonado, Bates, & Thal, 1992). Lab sessions for language assessment were then conducted in English and Spanish by bilingual researchers. There were significant correlations between the CDI and IDHC reported measures and observation in language assessment lab sessions. The authors concluded that parent report measures, specifically the CDI and the IDHC, could be good predictors of skill in bilingual child language.

Syntactic complexity in narratives. Evaluating syntactic complexity in children may give important information about how children will linguistically function in school. In studies of the narratives of English-speaking children, it has been found that students with language difficulties often have difficulty producing complex sentences and paragraphs (Gutiérrez-Clellen and Hofstetter, 1994). Gutiérrez-Clellen and Hofstetter (1994) looked at syntactic complexity in the narratives of 77 bilingual preschoolers, first graders, and third graders. The participants watched a short movie and were asked to explain the movie to an interviewer. The children's narratives were transcribed and segmented into T-units. Each T-unit was analyzed for subordination and phrase elaboration grammatical structures such as relative clauses, nominal clauses, infinitive

clauses, adverbial clauses, adverbial phrases, and prepositional phrases. The third graders produced more words per T-unit, more subordination, and more relative and prepositional phrases than the first grade group. They produced more prepositional phrases than the first grade group. There were no significant differences found among number of T-units, nominal clauses, infinitive clauses, adverbial clauses, or adverbial phrases among the three age groups. This study showed that, with age, children begin to embed more complex information in their sentences to allow for more descriptive explanations.

In a similar study, the syntactic complexity of the narratives of Spanish-speaking children learning English with low and average school achievement were compared (Gutiérrez-Clellen, 1998). Twenty-eight children with average achievement (*M* age = 8.1) and 29 children with low achievement (*M* age = 7.7) completed two language samples. In one, they described a brief, wordless movie and in the other they described the wordless picture book *Frog Goes to Dinner* (Mayer, 1974). The same grammatical structures as in the previously discussed study (Gutiérrez-Clellen & Hofstetter, 1994) were analyzed. The average achieving group produced more infinitive clauses, nominal clauses, and relative clauses. They also demonstrated more ability with subordination. There were no significant differences found among number of T-units, adverbial or prepositional phrases, or adverbial clauses. The limited achievement group produced less-complex sentence structures overall. This study also showed that the movie description task was more useful in assessing children's complex language than the book description task. Almost all of the syntactic structures being measured were produced with more frequency in the movie description task in comparison to the book description task.

Impact of first language loss on grammar. There is some concern that children may lose their first language as they begin to acquire and use their second language more often. Anderson (1999) sought to provide data on primary language loss and its effect on grammar in a Spanish-English bilingual child. The author completed a two-year longitudinal study of a primarily Spanish-speaking girl from Puerto Rico that was learning English in the United States. She was 4;7 at the beginning of the research and 6;5 at the last taping session. Interactions between the participant and her mother in Spanish were taped. Before or after each interaction, she interacted with another familiar adult in English. Each interaction was 30 minutes long and a total of 12 recordings were made over a period of 22 months. The participant's utterances were transcribed and were analyzed using mean length of utterance (MLU), mean length of response (MLR), and incidence of embedding across samples MLU. The participant's grammatical errors were divided into the following categories: (a) noun phrase agreement errors, (b) verb phrase errors, (c) syntactic errors, and (d) incorrect use of functor words and clitic pronouns.

The participant's Spanish MLU decreased over the course of the taping sessions from 6.1 in the first taping session to 3.2 in the last two sessions. Similarly, MLR and use of complex syntax decreased. Most of the participant's productive errors in the language samples were grammatical errors. Many of the grammatical errors were similar to those made by monolingual speakers' of Romance languages that have been diagnosed with SLI. Even though the participant's expressive language in Spanish decreased, there was no evidence of receptive loss. The participant was able to understand most of everything that was said to her and was able to follow all directions in Spanish. She expressed preference for speaking English rather than Spanish. Over the course of 22 months, the

participant demonstrated expressive L1 loss as a result of learning and use of L2. This study showed that a child's decreased productive use and input of L1 may result in loss of linguistic skill in that language, particularly expressive skill. Because the participants' errors were similar to those of monolingual speakers of Romance languages diagnosed with SLI, this study showed that it is important to consider the impact of L1 loss when evaluating a child learning a second language.

Spontaneous language sampling. In a dissertation by Lennon (1983), 30 Spanish/English bilingual children from the St. Louis area between the ages of 4;0 and 6;11 participated in a spontaneous language samples in the children's preferred language. The language samples were scored using Developmental Sentence Scoring (DSS; Lee, 1974) or Toronto's Developmental Assessment of Spanish Grammar (DASG; Toronto, 1976). For DSS, eight syntactic and morphologic structures were evaluated including: indefinite pronouns, personal pronouns, main verbs, secondary verbs, negatives, conjunctions, interrogative reversal and wh-questions. Because of a small sample size, statistical analyses were not used in the study. However, it was observed that mean sentence point, total score, and DSS mean score all increased with age. The sample requiring DASG analysis was even smaller than the sample requiring DSS, so statistical analyses were not possible. The limited data showed that, on average, 4;0-4;11 year olds and 6;0-6;11 year olds scored higher than 5;0-5;11 year olds. The author reported that the methods used for language sampling did not give satisfactory assessment of the children's language competence, but rather offered a limited, compartmentalized view of the children's true linguistic capabilities.

Comprehensive Language Analysis

Spanish-speaking children with SLI. Restrepo (1998) attempted to identify the best measures of discriminating between primarily Spanish-speaking children with normal language and primarily Spanish-speaking children with SLI. The participants in the study were 62 primarily Spanish-speaking preschoolers, kindergarteners, and first graders in Tuscon, Arizona. Half of the children were diagnosed as having moderate to severe language impairment and half had normal language. Each participant completed several verbal measures, including: (a) spontaneous language samples, which were subsequently analyzed using T-units, Developmental Assessment of Spanish Grammar (DASG; Toronto, 1976), mean length of T-unit (MLTU), and total number of errors per T-unit (NETU); (b) the Spanish Structured Photographic Expressive Language Test—II (SSPELT-II; Werner & Kresheck, 1989), and (c) language learning tasks. The parents of each participant also completed an interview. Analysis of the data showed that the best identifiers of Spanish-speaking children having language impairment were parental report, family history of speech and language problems, MLTU, and NETU.

Bilingual children compared to monolingual children with SLI. There is a large amount of scientific evidence that the developing English of a bilingual child learning English as a second language will often have many errors, but there has been little research regarding the types of errors that second language learners make. Paradis (2005) evaluated whether the English of bilingual children learning English as a second language compared to the English of same-age monolingual children with SLI. Paradis chose to compare these two groups because both groups have intact sensory and nervous systems and appear to be TD besides language, but both have limited language abilities.

In the study, Paradis evaluated the grammatical morphology of 24 bilingual children between the ages of 4;4 and 7;10. Each child participated in an ongoing study and was visited every six months. During visits, an interview with parents was completed and children participated in the *Columbia Mental Maturity Scale* (CMMS; Burgemeister, Hollander Blum, & Lorge, 1972), the *Test of Early Grammatical Impairment* (TEGI; Rice & Wexler, 2001), and a 45-minute language sample. The bilingual children's language was evaluated in the following areas: percentage correct, omission error, and commission error scores. Data analysis showed that the majority of the scores for the bilingual children's language were within the range of scores for children with SLI, even though there was no reason to suspect that any of the participants had language disorders. It can be seen in this study that bilingual child language is similar to the language of monolingual children with SLI. Because of this, misidentification of bilingual children having SLI is possible.

Conclusion

Research has evaluated the accuracy of automated grammatical tagging in both English-speaking and Spanish-speaking monolingual children. The present study furthers research in this area by extending the project to bilingual Spanish-speaking children learning English by evaluating the accuracy of automated grammatical tagging in both languages.

Method

The present study is a continuation of research being done at Brigham Young University in the development and evaluation of automated grammatical tagging software. The study extends this research in examining the accuracy of tagging English and Spanish language samples taken from bilingual children living in the United States.

Participants

All child language samples were taken from the Child Language Data Exchange System (CHILDES) database (MacWhinney, 2000). The present study used bilingual language samples in Spanish and English taken from two sources: (a) the Snow-Velasco samples, taken from a study of Puerto Rican bilinguals enrolled in a bilingual school program (Davidson, Kline, & Snow, 1986; Velasco; 1989) and (b) "Frog Stories," taken from a set of narrative samples of bilingual and monolingual Spanish- and English-speaking children in Miami, Florida (Pearson, 2002).

Snow-Velasco samples. The Snow-Velasco samples (Davidson, Kline, & Snow, 1986; Velasco; 1989) included an English and a Spanish language sample from each of 80 bilingual children from Puerto Rico. The participants were all enrolled in bilingual classes in New Haven, Connecticut public schools. Students received their education based on a "pairing model" for bilingual education in which students received education for half of each day in Spanish and half in English. The Spanish curriculum included reading and content area teaching. The English curriculum included repetition in English of some content students had already learned in Spanish, English reading, and ESL instruction. Half of the participants were third graders and the other consisted of fifth graders.

Participants were selected based on teacher ratings. In teacher ratings, half the children in each grade were described as poor readers and half were described as good readers. Each participant was either born in Puerto Rico or was born in the United States to Puerto Rican parents. All participants spoke primarily Spanish in the home. Participants selected for the study had been involved in the school system's bilingual program for two years or more. All participants qualified for a free lunch program, which indicated poverty status, and all came from homes where parents were either unemployed or were unskilled laborers. Participants completed the California Test of Basic Skills in Spanish. Results showed reading scores between the 60th and 89th percentile for the good readers and the 11th to 40th percentile for the poor readers. Each participant had English decoding skills at a third grade level or higher as determined by the Word Recognition Achievement Test.

For language assessment in both English and Spanish, participants completed two tasks including a definition task and a picture description task. In the definitions task, participants were given the following instructions: "What does ___ mean?"/ "¿Qué quiere decir ___ ?". Participants were asked to define: bicycle, bird, clock, diamond, donkey, flower, foot, hat, knife, nail, stool, thief, and umbrella. In the picture description task, participants were shown a picture of children playing or participating in common household activities. They were given the following specific instructions: "Please describe this picture so that another child that will be coming after you can draw a picture exactly like this one but without looking at it, just by listening to you" / "Por favor describe lo que está pasando en este dibujo, para que el niño que venga después pueda hacer un dibujo igual a este, pero sin verlo, solo escuchandote a ti."

Frog stories. The second set of samples was collected in Miami, Florida in a study of bilingualism (Pearson, 2002). The children from the Miami study that were included in the present study were the 173 bilingual children with both English and Spanish language samples. The children were enrolled in one of four different educational programs in Miami including: (a) English immersion for Hispanic students, (b) "two-way" bilingual programs for Hispanic students in which students received half Spanish instruction and half English instruction each day, (c) regular English instruction for non-Hispanic students, or (d) monolingual English-speaking students in schools with a majority of Hispanic students.

Language samples in English and Spanish were gathered using *Frog, Where Are You?* (Mayer, 1969), a wordless picture book. Participants were instructed to look through the book, and then tell the story while they looked at the pictures. Adults who collected language samples were careful to limit verbal feedback and give only neutral comments as to not influence the children's' narrations (Berman & Slobin, 1994). Other precautions were taken to ensure that children used their own natural language.

Manual Analysis

All English and Spanish language samples were manually tagged, or coded for part-of-speech. Examples of grammatical tagging for both English and Spanish are included in Table 1. The English samples were tagged using the same tags as in the Channell and Johnson study (1999). The Spanish samples were tagged using a Spanish grammatical tagging scheme inspired by the Language Assessment, Remediation, and Screening Procedure (LARSP; Crystal, Garman, & Fletcher, 1989). Spanish tags were similar to English tags, with some modifications due to differences in verbs and

auxiliaries, more complex Spanish morphology, different prepositional combinations with determiners and pronouns, and differences in syntax (Wilson, 2005).

Table 1
Examples of English and Spanish Grammatical Tagging

Language	Utterance	Example
English	Child Utterance	there was a boy that had a frog in a jar.
	Tagged Utterance	#g there <EX was <BCDZ a <DA boy <NN that <PRL had <VBD a <DA frog <NN in <IN a <DA jar <NN . <.
Spanish	Child Utterance	había una vez un niño que tenía un sapo.
	Tagged Utterance	#g había <V.d una <D vez <N un <D niño <N que <SB tenía <V.d un <D sapo <N . <.

A second clinician independently tagged approximately 1,000 utterances in English and approximately 1,000 utterances in Spanish to determine interrater reliability. Interrater reliability was found to be 96.5% on a tag-by-tag basis.

Software for Automated Grammatical Tagging

The software for both the Spanish and English for analysis of language samples included two types of probability information. The first was relative tag probability. This probability took a given word and determined the likelihood of it having a specific grammatical tag based on likelihood information in an electronic dictionary (Channell & Johnson, 1999). There was normally more than one tag option for each word, so probability determined which tag most likely would fit best the given word. The other probability information was tag transition probability, which determined the likelihood of one tag option coming after another tag. For example, the word pequeño, when preceded or followed by a noun, would most likely be tagged as an adjective (el <D hombre <N pequeño <AJ or el <D pequeño <AJ hombre <N / the small man). However, when

preceded by a determiner and not preceded or followed by a noun, pequeño would be a pronoun (quiero <V el <D pequeño <PO / I want the small [one]). Tag transition probability took these variations of word order into account and determined the likelihood of one tag coming after another.

The Spanish and English dictionaries included words in all categories, tag options, tag option frequencies, and tag sequence frequencies (Wilson, 2005). The dictionaries both included a corpus of words taken from adult and child language samples.

Procedure

All of the English and Spanish language samples were formatted into plain text documents and unnecessary information was removed. All child utterances were formatted in lowercase text, except for proper nouns. Parentheses were placed around the children's minor utterances, false starts, and unintelligible words so the computer would skip these words. An asterisk was placed before the interviewers' utterances so the program would not tag these utterances, but allowed the utterances to remain in order to give background and context for the language samples.

Manual and automated analyses were compared using utility programs in each language. Each tag applied by the software was compared with each tag applied manually. English analyses were carried out using tagging software developed by Channell and Johnson (1999) for English automated tagging. Spanish samples were tagged using software developed by Channell and Wilson (2003). The samples were compared on a word-by-word basis to determine accuracy of the software.

Results

Accuracy

The accuracy with which the software's tags matched with manual tags at the word level in English ranged from 95.6% to 97.6% with a mean of 96.4% (*SD* = 1.8). The word-level accuracy for the Spanish language samples ranged from 96.2% to 97.3% with an average of 96.8% (*SD* = 1.8). Table 2 lists the mean number of tags and percent accuracy scores with standard deviations for each group of participants in the English and Spanish child language samples.

Table 2
Mean Number of Tags and Percent Accuracy Scores with Standard Deviations for English and Spanish Child Language Samples

	English		Spanish	
Population	Tags	Accuracy (*SD*)	Tags	Accuracy (*SD*)
Miami Grade 2	263.0	96.4% (1.4)	232.4	96.8% (1.8)
Miami Grade 5	296.8	95.6% (1.7)	250.0	97.3% (1.5)
Snow-Velasco Grade 3	179.4	97.1% (1.8)	179.6	96.2% (2.1)
Snow-Velasco Grade 5	212.5	97.6% (1.6)	217.1	96.4% (1.7)
Combined	252.9	96.4% (1.8)	227.9	96.8% (1.8)

It can be seen that the Spanish accuracy scores were slightly higher than English accuracy scores in the Miami samples and slightly lower in the Snow-Velasco samples. However, these differences were minimal. Grade level did not have a significant effect on accuracy level. In the Miami samples, accuracy was higher for grade 2 in English, but higher for grade 5 in Spanish. In the English Snow-Velasco samples, grade 5 had higher accuracy than grade 3. In Spanish, grade 5 had slightly higher accuracy than grade 3. Thus, although there were differences among languages and samples, these differences

and interactions among variables were found to be not statistically significant using a two-way mixed analysis of variance at an alpha level of .05.

One measure of linguistic complexity of the samples was found by dividing the number of utterances by the number of tags in each sample. Using Pearson's correlations, there was a slight tendency for the accuracy of English tagging to decrease with increased sentence complexity, $r = -0.26$, $p < .01$. This was not true for Spanish, nor for other correlations of number of tags or utterances with percentage of accuracy, $p > .01$.

Tag-by-Tag Accuracy

Tag descriptions, frequencies of occurrence, tag-by-tag percent accuracy levels, and tag confusions for English and Spanish are shown in Table 3 and Table 4 respectively. These tables give detailed information about which tags were confused with the correct tags in automated analysis.

In English, when considering only the grammatical categories that occurred 30 times or more, automated grammatical tagging demonstrated high levels of accuracy (90% or better) for all variations of copula (96-100%), existential words (97%), infinitive markers (99%), determiners (97-100%), prepositions (91%), adjectives (95-100%), and nouns (99-100%). Automated grammatical tagging demonstrated low levels of accuracy (70% or below) for conjunctions including phrasal conjunctions (52%) and subordinating conjunctions (22%).

In Spanish, when considering only the grammatical categories that occurred 30 times or more, automated grammatical tagging demonstrated high levels of accuracy (90% or better) for the copula (99-100%), connectives (100%), determiners (100%), intensifiers (90%), question wh-words (92%), negation (96%), all forms of nouns (95-

100%), all forms of pronouns (95-100%) and several types of verbs including simple verbs (97%), simple verbs with an attached pronoun (90%), past tense verbs (97%), present participle (98%), and subjunctive verbs (94%). Automated grammatical tagging demonstrated low levels of accuracy (70% or below) for singular and plural adjectives (53-64%), and several categories of verbs including present participle verbs with an attached pronoun (32%), imperative verbs (39%), imperative verbs with an attached pronoun (31%).

Table 3
Description, Frequency, Accuracy, and Tag Confusion for English Tags

Tag	Description	Frequency	Accuracy (%)	Confusions (%)
Copula				
BC	be	83	100	
BCD	were	83	96	XBD (4)
BCDZ	was	669	97	XBDZ (3)
BCM	am	1	100	
BCR	are	111	98	XBR (2)
BCZ	is	908	99	XBZ (1)
Connectives/Other				
CC	coordinating conjunctions	3606	86	CP (14)
CP	phrasal conjunctions	743	52	CC (48)
CS	subordinating conjunctions	87	22	DD (1) PD (67) PRL (9) RQL (1)
EX	existential	337	97	PS3 (1) RB (2)
TO	infinitive marker	680	99	IN (1)
Determiners				
D$	possessives	1029	100	
DA	articles	10334	100	
DCN	cardinal numbers	249	99	PN (1)
DD	demonstrative singular	98	97	PD (3)
DDS	demonstrative plural	8	100	
DN	indefinites	377	99	RQL (1)
DON	ordinal numbers	6	67	NN (33)
DPA	predeterminer/initiator	50	100	
DWN	wh-nominal clause	1	100	
DWQ	wh-question	1	100	
DWX	wh-exclamative/qualifier	1	0	RBQ (100)
Prepositions				
IN	prepositions	5404	91	RP (8)
Adjectives				
JJ	adjectives	1652	95	NN (2) RB (1) VBN (1)
JJR	comparative	10	100	
JJT	superlative	2	50	NN (50)
Nouns				
NN	singular nouns	12005	99	
NNS	plural nouns	1526	100	
NP	proper nouns	151	100	
Pronouns				
P$	possessives	4	50	D$ (50)
PD	demonstratives singular	180	95	CS (1) DD (1) PRL (3)
PDS	demonstratives plural	1	100	
PI	indefinites	388	100	
PL	reflexive/intensive singular	19	95	NN (5)
PLS	reflexive/intensive plural	1	100	

Tag	Description	Frequency	Accuracy (%)	Confusions (%)
PN	non-precise quantifiers and ordinals	315	91	DCN (3) DN (1) NNS (1) RB (2) VB (1)
PO	object forms	1060	100	
PRL	relative pronouns	401	88	CS (3) PD (7)
PS	subject forms	1567	100	
PS3	third person singular	2125	100	
PWN	wh-nominal	30	83	DN (13) DWN (3)
PWQ	wh-interrogative	20	90	PWN (10)
Adverbs				
RB	adverbs	2534	97	NN (1) RP (1)
RBN	introducing noun clause	27	89	RBQ (4) RBS (7)
RBQ	question wh-adverb	31	77	DWX (3) NN (3) RBN (6) RBS (10)
RBR	comparative	10	90	JJR (10)
RBS	introducing adverbial clause	813	99	
RBT	superlative	1	100	
RP	particles of verbs	740	91	IN (5) RB (3)
RQL	intensifier/qualifier	143	77	DD(1) JJ(1) NN(2) PN(9) RB(8) RQLP(2)
RQLP	postqualifiers	1	100	
Verbs				
VB	verb	2284	96	NN (4)
VBD	past	4150	98	
VBG	participle/gerund	1788	95	NN (5)
VBN	past participle	177	73	JJ (3) NN (2) VB (2) VBD (19)
VBTO	verb+to	62	100	
VBZ	third person present singular	912	98	NNS (2)
VPO		1	100	
Auxiliaries				
XB	be	3	100	
XBD	were	119	92	BCD (8)
XBDZ	was	642	99	BCDZ (1)
XBG	being	3	100	
XBM	am	2	100	
XBN	been	3	100	
XBR	are	102	93	BCR (7)
XBZ	is	418	96	$ (1) BCZ (3)
XD	do	86	100	
XDD	did	100	97	VBD (3)
XDZ	does	14	100	
XG	get	13	15	VB (46) VBD (15) XGD (23)
XGD	got, getted	53	85	VBD (15)
XGZ	gets	3	33	VBZ (67)
XH	have	4	75	VB (25)
XHD	had, haved	22	23	VBD (77)
XHZ	has	5	20	VBZ (80)
XM	modal	223	100	
XM*	modal+negation	30	100	

Table 4
Description, Frequency, Accuracy, and Tag Confusion for Spanish Tags

Tag	Description	Frequency	Accuracy (%)	Confusions (%)
Adjectives				
AJ	adjective singular	684	64	N (33) V (1) V.n (1)
AJ.s	adjective plural	144	53	N.s (44) V.i+P (1) V.n (1)
Adverbs				
AV	adverb	2517	89	PR (9)
Copula				
B	copula	857	100	
B.d	past	506	99	X.d (1)
B.i	imperative	1	100	
B.s	subjunctive	4	75	X.i (25)
Connectives/Other				
CC	conjunction	4233	100	
SB	subordinator	2487	100	
D	determiner	11227	100	
IF	intensifier/qualifier	133	90	AJ (3) AV (1) D (5) IN (1) PO (1)
IN	initiator	103	78	D (21) PO (1)
Q	question wh-word	36	92	IF (3) SB (6)
TO	infinitive marker	438	84	PR (15)
NG	negation	542	96	D (1) IF (1) N (2)
Nouns				
N	noun	10742	98	V (1)
N.s	plural noun	1498	95	AJ.s (1) V (3)
NP	proper noun	88	100	
Pronouns				
PO	other pronoun	950	95	D (1) V (3)
PP	personal pronoun	4230	97	D (3)
PR	preposition	4682	99	
PR+D	preposition+determiner	748	100	
Verbs				
V	verb	3272	97	N (2) N.s (1)
V+P	verb+pronoun	176	90	N (6) N.s (1) V (2) V.i+P (1) V.s (1)
V+P+P	verb+pronoun+pronoun	8	75	N (25)
V.c	conditional	1	100	
V.d	past (preterit/imperfect)	4581	97	N (3)
V.g	present participle	1321	98	N (2)
V.g+P	present participle+pronoun	85	32	N (68)
V.i	imperative	18	39	V (22) V.s (39)
V.i+P	imperative+pronoun	13	31	N (62) V+P (8)
V.n	past participle	143	77	AJ (3) N (15) N.s (1) SB (1) V (1) V.i+P (3)
V.n+P	past participle+pronoun	1	0	V.i+P (100)
V.s	subjunctive	158	94	N (1) V.d (3) V.i (1) X.i (1)

Tag	Description	Frequency	Accuracy (%)	Confusions (%)
Auxiliaries				
X	auxiliary	561	74	B (15) V (11)
X.d	past	739	93	B.d (4) V.d (3)
X.i	imperative	1	100	
X.s	subjunctive	5	80	V.s (20)

Discussion

Automated grammatical tagging produces high levels of tag accuracy for both the English and Spanish of bilingual children's language.

Tagging accuracy in English was similar to tagging accuracy found in previous research in English automated grammatical tagging (Channell & Johnson, 1999). The participants in the present study included 254 bilingual children in either second, third, or fifth grade. The participants in the Channell and Johnson study included 30 monolingual children between the ages of 2;6 and 7;11. However, the samples in the present study were much shorter, averaging 25 utterances, than the Channell and Johnson study, which averaged 203 utterances in length. Despite differences in age level and mean number of utterances, English tagging accuracy was similar between the present study and Channell and Johnson's study.

Unlike the findings in English, the Spanish tagging accuracy was not similar to previous research in Spanish automated grammatical tagging. Wilson's (2005) study found lower accuracy levels (87.7% to 95.6%) in Spanish child language samples than those of the present study. The difference in accuracy could be due to differing levels of sample length, linguistically simpler samples, or improved program performance. The average number of utterances per sample in the Spanish language samples used in the present study was 24 utterances, whereas the average number of utterances per sample in the Wilson study was 232 utterances. The present study included a greater number of language samples, but the samples were much shorter in length than those evaluated in Wilson's study.

There were several limitations of the present study that may have influenced the results. First, as previously noted, the child language samples included in the present study were quite short, containing on average 25 utterances per sample for English and 24 utterances per sample for Spanish. Next, the presence of code switching, or the mixing of a child's two languages, in both the English and Spanish child language samples may have had an effect on the results of the study. Because the children in the samples were bilingual, they often used a word or phrase of one language while speaking in the other language. The software programs for English and Spanish are specific to their respective languages and cannot accommodate words from the other language, so these words were not included in the tagging. Finally, the Spanish skills of the researchers were not at the same level of proficiency as native Spanish speakers. More advanced Spanish skills of the researchers may have resulted in greater differences in automated and manual tagging. Future research in the area of automated grammatical tagging for bilingual child language is warranted. It would be of particular interest for researchers to be able to account for code switching in bilingual automated grammatical tagging.

The study of automated grammatical tagging in Spanish remains early in its development and is therefore not yet directly applicable to clinical practice. In English, the output of the tagging program is used by Computerized Profiling (CP; Long, Fey, & Channell, 2003), which offers clinically useful child language analyses. The CP software gives an inventory of children's syntactic structures using analyses such as LARSP (Crystal et al., 1989) and provides quantitative measures through analyses such as DSS (Lee, 1974) and MLU. As more research is conducted in this area, the output of Spanish software may be similarly used for use in more clinically applicable programs such as

CP. The present study provides important preliminary data in the area of automated grammatical tagging for bilingual child language. Further research is required to make the Spanish software clinically beneficial; however, automated grammatical tagging has potential to be a useful tool for the clinician serving bilingual populations.

References

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

Anderson, R. T. (1999). Impact of first language loss on grammar in a bilingual child. *Communication Disorders Quarterly, 21,* 4-16.

Bedore, L. M., Peña, E. D., García, M., & Cortez, C. (2005). Conceptual versus monolingual scoring: When does it make a difference? *Language, Speech, and Hearing Services in Schools, 36*, 188-200.

Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44,* 905-924.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Burgemeister, B., Hollander Blum, L., & Lorge, I. (1972). *Columbia Mental Maturity Scale.* New York: The Psychological Corporation.

Channell, R. W., & Johnson, B. W. (1999). Automated grammatical tagging of child language samples. *Journal of Speech, Language, and Hearing Research, 42,* 727-734.

Channell, R. W., & Wilson, K. (2003, November). *Automated grammatical analysis of children's Spanish language samples.* Poster session presented at the annual meeting of the American Speech-Language and Hearing Association, Chicago.

Crystal, D., Garman, M., & Fletcher, P. (1989). *The grammatical analysis of language disability: A procedure for assessment and remediation* (2nd ed.). London: Cole and Whurr.

Davidson, R., Kline, S., & Snow, C. E. (1986). Definitions and definite noun phrases: indicators of children's decontextualized language skills. *Journal of Research in Childhood Education, 1*, 37-40.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test—Revised.* Circle Pines, MN: American Guidance Service.

Dunn, L., Padilla, E., Lugo, D., & Dunn, L. (1986). *Test de Vocabulario en Imágenes Peabody—Adaptación Hispanoamericana* [Peabody Picture Vocabulary Test—Latin American adaptation]. Circle Pines, MN: American Guidance Service.

Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., et al. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual.* San Diego, CA: Singular Publishing Group.

Fiestas, C. E., & Peña, E. D. (2004). Narrative discourse in bilingual children: Language and task effects. *Language, Speech, and Hearing Services in Schools, 35*, 155-168.

Goldstein, B. A. (2004). *Bilingual language development & disorders in Spanish-English Speakers*. Baltimore: Paul H. Brookes.

Gutiérrez-Clellen, V. F. (1998). Syntactic skills of Spanish-speaking children with low school achievement. *Language, Speech, and Hearing Services in Schools, 29,* 207-215.

Gutiérrez-Clellen, V. F., & Hofstetter, R. (1994). Syntactic complexity in Spanish narratives: A developmental study. *Journal of Speech & Hearing Research, 37,* 645-654.

Gutiérrez-Clellen, V. F., Restrepo, M. A., Bedore, L., Peña, E., & Anderson, R. (2000). Language sample analysis in Spanish-speaking children: methodological considerations. *Language, Speech, and Hearing Services in Schools, 31*, 88-98.

Jackson-Maldonado, D., Bates, E., & Thal, D. (1992). *Fundación MacArthur: Inventario del desarrollo de habilidades comunicativas.* San Diego, CA: San Diego State University.

Kester, E. S., & Peña, E. D. (2002). Language ability assessment of Spanish-English bilinguals: future directions. *Practical Assessment, Research & Evaluation, 8*.

Lee, L. (1974). *Developmental sentence analysis.* Northwestern University Press.

Lennon, G. H. (1983). Evaluating St. Louis Hispanic Children's Language Dominance Through Formal and Spontaneous Measures. (Doctoral dissertation, Towson University, 1983). Dissertation Abstracts International, 44 (07), 2073.

Long, S. H. (2001). About time: A comparison of computerized and manual procedures for grammatical and phonological analysis. *Clinical Linguistics & Phonetics, 15*, 399-426.

Long, S. H., Fey, M. E., & Channell, R. W. (2003). *Computerized Profiling (Version 9.6)* [Computer software]. Retrieved from http://www.computerizedprofiling.org/

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed; Vol. 2). Mahwah, NJ: Lawrence Erlbaum Associates.

Marchman, V. A., & Sussmann-Martínez, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research, 45,* 983-997.

Mayer, M. (1969). *Frog, where are you*? New York: Dial Books for Young Readers.

Mayer, M. (1974). *Frog goes to dinner.* New York: The Dial Press.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools, 36,* 172-187.

Patterson, J. L. (1998). Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *American Journal of Speech-Language Pathology, 7, 46-56.*

Patterson, J. L. (2000). Observed and reported expressive vocabulary and word combinations in bilingual toddlers. *Journal of Speech, Language, and Hearing Research, 43, 121-128.*

Pearson, B. Z. (2002). Narrative competence among monolingual and bilingual school children in Miami. In D.K. Oller and R.E. Eilers (Eds.), Language and literacy in bilingual children (pp. 135-174). Clevedon, UK: Multilingual Matters.

Pearson, B. Z., Fernández, S., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language and Learning, 43*, 93-120.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1398-1411.

Restrepo, M. A., & Kruth, K. (2000). Grammatical characteristics of a Spanish-English bilingual child with specific language impairment. *Communication Disorders Quarterly, 21,* 66-76.

Restrepo, M. A., & Silverman, S. W. (2001). Validity of the Spanish Preschool Language Scale-3 for use with bilingual children. *American Journal of Speech-Language Pathology, 10*, 382-393.

Rice, M., & Wexler, K. (2001). *Test of Early Grammatical Impairment.* New York: The Psychological Corporation.

Saenz, T. I., & Huer, M. B. (2003). Testing strategies involving least biased language assessment of bilingual children. *Communication Disorders Quarterly, 24,* 184-193.

Semel, E., Wiig, E., & Secord, W. (1995). *Clinical Effectiveness of Language Functioning—Third Edition.* San Antonio: The Psychological Corporation.

Toronto, A. S. (1976). Developmental assessment of Spanish grammar. *Journal of Speech and Hearing Disorders, 41,* 150-171.

Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development, 63*, 1012-1020.

Velasco, P. (1989). *The relationship between decontextualized oral language skills and reading comprehension in bilingual children.* Unpublished doctoral dissertation. Harvard Graduate School of Education.

Werner, E. O., & Kresheck, J. D. (1989). *Spanish Structure Photographic Expressive Language Test—Preschool.* Sandwich, IL: Janelle Publications.

Werner, E. O., & Kresheck, J. D. (1989). *Spanish Structure Photographic Expressive Language Test—II.* DeKalb, IL: Janelle.

Wilson, K. M. (2005). Automated grammatical tagging of children's language samples in Spanish. Unpublished master's thesis, Brigham Young University, Provo.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1993). *Preschool Language Scale-3: Spanish Edition.* San Antonio, TX: Psychological Corporation.