Published in final edited form as:
J Commun Disord. 2010 ; 43(6): 498–510. doi:10.1016/j.jcomdis.2010.05.002.

Language Sample Measures and Language Ability in Spanish English Bilingual Kindergarteners

Lisa M. Bedore[a], **Elizabeth D. Peña**[a], **Ronald B. Gillam**[b], and **Tsung-Han Ho**[a]

Lisa M. Bedore: lbedore@mail.utexas.edu; Elizabeth D. Peña: lizp@mail.utexas.edu; Ronald B. Gillam: ron.gillam@usu.edu; Tsung-Han Ho: tsunghan@mail.utexas.edu

[a] The University of Texas at Austin

[b] Utah State University

Abstract

Measures of productivity and sentence organization are useful metrics for quantifying language development and language impairments in monolingual and bilingual children. It is not yet known what measures within and across languages are most informative when evaluating the language skills of bilingual children. The purpose of this study was to evaluate how measures of language productivity and organization in two languages converge with children's measured language abilities on the Bilingual English Spanish Assessment (BESA), a standardized measure of language ability. 170 kindergarten age children who produced narrative language samples in Spanish and in English based on a wordless picture book were included in the analysis. Samples were analyzed for number of utterances, number of different words, mean length of utterance, and percentage of grammatical utterances. The best predictors of language ability as measured by the BESA scores were English MLU, English grammaticality, and Spanish grammaticality. Results are discussed in relationship to the nature of the measures in each of the languages and in regard to their potential utility for identifying low language ability in bilingual children.

1. Introduction

Narrative samples are often recommended for use as part of a language evaluation battery for bilingual children (Damico & Oller, 1983; Gutiérrez-Clellen, 2002; Jax, 1988). Narratives are ecologically valid, reflecting cultural norms as well as the demands of academic language use (Westby, Van Dongen, & Maggart, 1989). One practical advantage of using narratives to describe language development is that they are a source of information about discourse level organization as well as productivity and sentence level organization (Fiestas & Peña, 2004). The creation of a good narrative requires children to coordinate the organization of events with the production of utterances that convey specific meaning. Thus, narrative story-telling challenges children to produce longer and more complex utterances than they might routinely use in conversational speech. Language formulation difficulties are especially likely to be evident in narrative (Hadley, 1998; Leadholm & Miller, 1995). These difficulties are likely to be manifested in narrative sample measures such as number of words used, mean length of utterance (MLU) and proportion of grammatical utterances.

Corresponding author: Lisa M. Bedore, Permanent address: Department of Communication Sciences and Disorders, 1 University Station A1100, The University of Texas at Austin, Austin, Texas 78712, Phone-512-232-5101.

When evaluating the language abilities of bilingual children, it is important to account for their distributed language knowledge. One solution is to administer standardize tests that contain culturally and linguistically appropriate items that have been normed on a bilingual population. Some examples of tests that follow these principles include the *Bilingual English Spanish Assessment* (BESA) (Peña, Gutierrez-Clellen, Iglesias, Goldstein, & Bedore, in development), the *Preschool Language Scale - Spanish 4* (Zimmerman, Steiner, & Pond, 2002) and the *Clinical Evaluation of Language Fundamentals-4 Spanish* (Wiig, Secord, & Semel, 2006). Within the semantics domain, one solution to this challenge is to use conceptual scoring to credit responses in either language (Bedore, Peña, García, & Cortez, 2005; Brownell, 2001; Pearson, Fernández, & Oller, 1993). Further study of bilingual language production is needed to evaluate solutions that focus on the morphosyntactic domain.

Our goal is to explore the potential of using language sample measures from two languages to inform assessments of language ability. We begin by providing background on the use of language sample measures on children with language impairment (LI) and then discuss what is known about them for bilingual children with and without LI. We focus on productivity and sentence organization. In the work presented here children with language impairment demonstrate delayed language development relative to their typically developing age-matched peers. In the majority of the studies of monolingual children, participants meet the additional diagnostic criteria for specific language impairment of cognitive development within the average range. In contrast, in the studies of bilingual children participants have often not been given cognitive testing due to the lack of valid instruments for this population (Peña & Bedore, 2008).

1.1 Measures of productivity

Productivity, or the amount of language used by a child, is one indicator of language knowledge. Language productivity can be measured as the amount of language (words or utterances) that a person generates in a set period of time (Leadholm & Miller, 1995). Miller and Iglesias (2008) have standardized a narrative sample analysis for bilingual children using Systematic Analysis of Language Transcripts software. An important advantage of a standardized narrative analysis is that it makes it possible to determine if various measures of language productivity are within the average range for a particular task (Miller & Iglesias, 2008). Number of total words (NTW) and number of different words (NDW) both tend to increase with age in typically developing children, but NDW tends to be a better indicator of language growth since it reflects diversity of vocabulary. Across studies of children with language impairment (LI), NDW is consistently low relative to their typically developing peers suggesting that it is a potentially sensitive measure (Hewitt, Hammer, Yont, & Tomblin, 2005; Klee, Stokes, Wong, Fletcher, & Gavin, 2004; Owen & Leonard, 2002; Redmond, 2004; Watkins, Kelly, Harbers, & Hollis, 1995).

Typically developing bilingual children, like their monolingual peers, appear to systematically increase the number and variety of words they produce in their stories over the early school years. Table 1 summarizes findings from studies of Spanish-English bilingual children that have reported data on a single story and shows how measures of lexical diversity change over time. Older children typically produce stories with more words and a greater variety of words than do younger children. For example, Muñoz, Gillam, Peña, and Gulley-Faehnle (2003) elicited the *Frog Where are You* (Mayer, 1969) story in English from Spanish-English bilingual preschoolers. NTW and NDW both increased but did not differ significantly by age perhaps due to the limited age range under consideration. Using the same story, Miller and colleagues followed children from kindergarten to third grade to evaluate the relationship between oral language and literacy outcomes (Miller, et al., 2006). There was an overall significant linear trend in the number of different words children produced in English and Spanish by grade

level, and NDW was significantly associated with literacy outcomes. In a study by Uccelli and Paez'(2007), children increased their lexical productivity in both languages, but only the change in English was statistically significant.

Data on lexical development in bilingual children with LI show that they produce fewer words and make more naming errors than their typically developing bilingual peers (Peña & Bedore, 2008). For example, Simonsen (2002) compared the performance of Swedish Finnish bilingual 6-year-olds with and without LI to their monolingual Swedish peers on a standardized picture-naming task. The bilingual children with LI named fewer pictures relative to their Swedish monolingual peers with LI and relative to their typically developing bilingual peers. The bilingual and monolingual children with LI had different response patterns. The monolingual children with LI produced more semantically related and classification errors. The bilingual children with LI also produced semantically related errors but were more likely present non-responses. Along the same lines, a comparison of NDW in the narrative samples of kindergarten age bilinguals showed that children with LI used less diverse vocabularies in each of their languages relative to their typically developing peers (73.8 vs. 80.6 in English and 68.1 vs. 79.6 in Spanish) (Simón-Cereijido & Gutiérrez-Clellen, 2009). These findings indicate that NDW may be sensitive to language development and language ability in bilingual children as has been observed in other languages (e.g., English and Cantonese) (Hewitt, et al., 2005; Klee, et al., 2004).

1.2 Measures of sentence organization

One of the most well-documented difficulties of children with LI is difficulty in the production of grammar. These difficulties are reflected in measures of sentence organization such as mean length of utterance (MLU) and grammaticality. Mean length of utterance is the most widely used clinically and in research; MLU increases systematically in early language development (Brown, 1973; Miller, 1981). In sufficiently challenging contexts (such as narrative) sentence length continues to increase for English speaking children into the early school years (Loban, 1976). Data on Spanish learning children also demonstrate change over time in MLU (Gutiérrez-Clellen, Restrepo, Bedore, Peña, & Anderson, 2000). In Spanish the growth trajectory in MLU is flatter relative to English increases related to increased sentence length due to added clauses in Spanish can be offset by the speakers ability to drop subjects or other information available contextually (Bedore, 2001; Gutiérrez-Clellen, et al., 2000).

Children with LI have consistently lower MLU than do their typically developing peers in English (Hewitt, et al., 2005; Klee, 1992; Leonard, 1998). In Spanish, children with LI also demonstrate MLUs that are significantly lower than their TD peers MLU (Bedore & Leonard, 2001; Jacobson & Schwartz, 2002). However, in these studies the difference between MLU in children with typically developing language skills and LI is smaller and overlaps more in Spanish than in English speakers. This may be due to the same factors that contribute to flatter growth trajectories in Spanish MLU.

Bilingual children demonstrate increases in MLU in Spanish and in English over the early school years as summarized in Table 2. Both Muñoz et al. (2003) and Miller et al. (2006) documented statistically significant growth in MLU over the early school years. MLU in both languages of bilingual children with LI is lower than that of their typically developing peers (5.5 vs. 6.23 in English and 5.2 vs. 5.8 in Spanish) (Simón-Cereijido & Gutiérrez-Clellen, 2009). These data, in combination with data showing that MLU increases in this time period for bilingual children, suggest that MLU may contribute to our ability to differentiate children with low language ability in both Spanish and English.

Difficulties in grammatical production in English and Spanish have been extensively documented at the level of individual grammatical constructions and at the level of global

measures of grammaticality (e.g., percentages of grammatical utterances, composite scores). At level of grammatical constructions English-speaking children have more consistent difficulties with tense marking (e.g., third person present tense singular forms, regular past tense forms) while forms such as articles, plurals, and possessives are less consistently observed to be difficult (e.g., Leonard, Eyer, Bedore, & Grela, 1997; Rice, Wexler, & Hershberger, 1998). In contrast, Spanish-speaking children have less difficulty with tense marking but often have difficulties with articles and clitics. They also have difficulties with number agreement in noun and verb phrases, person agreement in verb phrases, and gender agreement in noun phrases (e.g., Anderson & Souto, 2005; Bedore & Leonard, 2001, 2005).

Grammaticality is positively associated with judgments of language proficiency and ability (e.g., Bedore & Leonard, 1998; Gillam & Johnston, 1992; Restrepo, 1998; Rice & Wexler, 1996). Researchers have quantified grammatical development (or difficulties) in two ways. Some work focuses on global judgments of grammaticality. For example, Gillam and Johnston (1992) found that children with LI produced more ungrammatical utterances in their stories than children without LI. In a study of bilingual Spanish-English speaking children, Muñoz et al. (2003) found that the percentage of grammatical utterances produced by bilingual children in English rose from 58.75 percent at age 4;8 years to 80.08 by 5;3 years.

Other work has focused on building composite scores based on the coding of individual errors. This approach is potentially more specific since it can account for the possibility of multiple errors in an utterance and can be focused on error types of interest. For example, Rice and Wexler (1996) and Bedore and Leonard (1998) created composites based on errors that are characteristic of the language of children with specific language impairment (e.g., verb tense errors, noun phrase errors) to use as predictors of language abilities in English, and these contributed to the correct classification of children with and without LI. Similarly, Restrepo (1998) found that number of grammatical errors per utterance in Spanish contributed significantly to the identification of children with LI.

A number of researchers have documented variability in grammatical production in bilingual children. Spanish speaking children learning English for example acquire some but not all of the grammatical structures expected relative to their MLU (Bland-Stewart & Fitzgerald, 2001). There are cross-linguistic influences that facilitate the acquisition of some structures but inhibit the acquisition of others. For example, French learning children acquiring Swedish use a low frequency Swedish structure to express past tense in their narratives that is more similar to the French construction (Schlyter, 1996). Bilingual children also produce some errors at higher rates than do their monolingual peers (Jacobson & Schwartz, 2005). For example typically developing bilingual children are more likely to produce overgeneralizations of tense markers *(ranned* for *ran)* than their similar aged monolingual peers.

In spite of some specific differences in the language of typically developing monolingual and bilingual children, it appears that the same kinds of grammatical constructions that are challenging for monolingual children with SLI are challenging for bilingual children with SLI. For example, Paradis, Crago, Genesee and Rice (2003) found that the errors produced by groups of bilingual French-English learning children were similar in type and rate to those of their monolingual peers with SLI in French and English. Bilingual Spanish-English speaking children with LI can be differentiated from their peers on the basis of the same types of errors produced by monolingual English and Spanish children with LI (Gutiérrez-Clellen, Restrepo, & Simón-Cereijido, 2006; Gutiérrez-Clellen & Simón-Cereijido, 2007; Gutiérrez-Clellen, Simón-Cereijido, & Wagner, 2008; Simón-Cereijido & Gutiérrez-Clellen, 2009). These findings suggest that measures of grammaticality should contribute to our ability to distinguish between of differing language learning abilities.

In sum, data on bilingual children acquiring English and Spanish provide evidence that children make progress on basic measures of linguistic productivity and sentence organization in the early school years. Data on children with LI indicate that these children have difficulties in productivity and sentence organization that would likely be captured in MLU and grammaticality measures. Work to date has focused on these measures in one language or another but not evaluated how such information could be combined. Thus the goal of the current study was to evaluate the extent to which data from language sample measures in Spanish and English converge with performance on a standardized measure of language ability. Specifically we addressed the following questions:

1. Do measures of productivity and sentence organization correlate with ability ratings based on the Bilingual English Spanish Assessment (BESA)?
2. Do measures of productivity and sentence organization correlate within and across languages?
3. Which measure or combination of measures based on narrative sample analysis converges with a measure of bilingual language ability?

2. Material and methods

2.1 Participants

One hundred and eighty-six kindergarten children who were enrolled in an on-going, longitudinal study of identification of LI in bilingual Spanish-English speakers were included in this study. Sixteen cases were eliminated from the analysis because the children did not have narrative samples available in both languages due to absence from school at the time of testing or failure of recording equipment. All children were enrolled in kindergarten and had a mean age of 68.13 months (SD =4.39). The children were from one of three school districts (two in central Texas and one in Utah) that enroll large numbers of Hispanic children.

2.2 Methods

As part of the test battery for the longitudinal study, children's parents and teachers completed the questionnaires about the child's patterns of language input and output (Gutiérrez-Clellen & Kreiter, 2003; Restrepo, 1998), children completed the experimental version of the Bilingual English Spanish Assessment (BESA) (Peña, et al., in preparation) and produced narrative samples in Spanish and English based on wordless picture books. All children completed the narrative and the BESA in Spanish before they completed the corresponding English tasks due to the constraints of the larger project.

To establish the children's level of exposure to Spanish and English a parent interview was administered by telephone. Children's parents rated their current levels of language input and output on an hour-by-hour basis (Gutiérrez-Clellen & Kreiter, 2003; Restrepo, 1998). Parents also provided information about their children's history of exposure to both languages at home and school from birth. Teachers provided information on the children's classroom language use. To be eligible to participate in the study children had to have at least 20% input and output in each language when they were in pre-kindergarten. At the time children were tested (in kindergarten) their use of English and Spanish spanned the full range from predominant Spanish use to predominant English use. Based on the children's patterns of language output 38.2% of the children were Spanish dominant (using Spanish over 60% of time), 32.5% of the children were balanced bilinguals (using Spanish and English 40–60% of the time), and the remaining 29.2% of the children were English dominant (using English over 60% of the time). The average number of years of exposure to English was 2.21 years of age. Thirty eight percent of the children learned English and Spanish from age 2 or earlier and 61.8 % of the children started learning English at age three years or later.

To establish the children's level of language ability, all of the children completed the experimental version of the BESA (Peña, et al., in preparation). The *BESA* is a standardized test of language ability for Spanish-English bilingual children. In the semantics subtests children listen to a story related to a culturally relevant theme (e.g., a picnic in Spanish or a birthday party in English). As the child hears the story and is shown illustrations of the objects or activities, she is asked to respond to 48 questions that are embedded in the story. Questions tap semantic knowledge such as category generation (e.g., Tell me all the zoo animals you can think of), similarities and differences (e.g., Show me which two invitations go together) and functions (e.g., Show me what you do with a handkerchief). Children's correct responses are credited in either English or Spanish on both subtests (see Bedore, et al., 2005; Peña, Bedore, & Rappazzo, 2003; Peña, Bedore, & Zlatic-Guinta, 2002). There are equivalent but not translated items on the Spanish and English versions. The Spanish and English item sets were selected to have equivalent item difficulty. Most children complete this task in approximately 30 minutes per language. Preliminary analysis of the Spanish semantics test data with a group of 284, 4–6 year-old children indicates co-efficient alpha between.78 and.84.For English, co-efficient alpha ranged from.81–.916 with a group of 244 children.

The morphosyntax subtests tap grammatical forms that are challenging for children with LI in English and in Spanish. On the cloze task, examiners read a complete sentence while pointing to a picture. They then read a second sentence corresponding to another, similar picture that the child completes using the targeted morphosyntactic markers (e.g., possessives in English, clitics in Spanish). Sentence repetition comprises the second part of the morphosyntactic test and is used to test more complex forms that cannot be elicited using cloze tasks. In Spanish there are 23 cloze item and 50 sentence repetition items and in English there are 32 cloze and 30 sentence repetition items. Typically children complete this task in 20 minutes per language. For the population and age range under consideration the BESA morphosyntax subtests have good to excellent discriminant classification (Gutiérrez-Clellen, et al., 2006; Gutiérrez-Clellen & Simón-Cereijido, 2007).

A language ability composite *z*-score was generated for each child based on the higher language performance on the six *BESA* subtests. The rationale for selecting the higher language score for each subtest was to obtain the best estimate of the child's language ability (rather than focusing on proficiency by selecting the score based only on one language). To generate the *z*-scores, means and standard deviations were calculated for each subtest in each language for the experimental sample. The subtest means were transformed into *z*-scores to standardize the scores. We compared each subtest *z*-score (semantics, morphosyntax cloze, and sentence repetition) by language (Spanish and English) and selected the higher *z*-score for each subtest. For the morphosyntax subtests we weighted the cloze and sentence repetition score equally because the task types and forms assessed both differentiate children with and without language impairment (Gutiérrez-Clellen, et al., 2006; Gutiérrez-Clellen & Simón-Cereijido, 2007) and averaged these with the semantics *z*-scores. Children's composite scores spanned the full range ability from high (1.77) to low (−2.46). This yielded a language ability composite used as the response variable in the multiple regression.

Table 3 shows that all possible combinations of *z*-scores contributed to the children's composite scores. This range of performance patterns illustrates the need for a cross-language composite to consider bilingual children's language performance in both languages simultaneously (Bedore & Peña, 2008). Using the higher score in each domain is consistent with findings that bilingual children often demonstrate mixed dominance as they shift from home to school demands (Gutiérrez-Clellen, 2002;Kohnert, Bates, & Hernandez, 1999;Kohnert & Danahy, 2007;Peña & Stubbe Kester, 2004).

2.2 Materials

Children told a story based on one of two wordless picture books for each language. For each language, the examiner first provided the child a story model that was based on the script provided by Miller and Iglesias (2008). The children retold the same story to the examiners while they looked at the pictures. Next, the children were given a new book. Examiners directed them to look through the pictures and then tell a story. Based on procedures developed by Berman and Slobin (1994), examiners provided backchannel responses (e.g., "Aha," "Sí," "Tell me more") as the children told the stories. If children stopped telling the story, the examiners restated the children's last utterance and encouraged them to continue to the end of the story. The second story, told without a model, was the focus of the analysis reported in this article. Approximately half of the children told a story based on Frog Goes to Dinner (Mayer, 1974) in Spanish and Frog, Where are You? (Mayer, 1969) in English. For the remaining children the order of the narratives by language was reversed. All samples were recorded using a digital audio recorder (Sony MS-515 or ICD-P320) with an external microphone (ECM 115) and then transcribed using Sony digital voice editor version 2.4.04. To ensure transcription reliability all transcripts were transcribed by a trained research assistant and checked by a second research assistant (usually the individual who had collected the language sample data). An independent transcriber resolved discrepancies. Transcript checks included checking spelling to ensure accurate word counts, checking segmentation to ensure accurate MLU counts, and checking grammaticality coding to ensure that grammaticality coding rules were adhered to.

2.3 Coding and Analysis Procedures

All stories were coded using *Systematic Analysis of Language Transcripts* (SALT) (Miller & Iglesias, 2008). In Spanish, verb forms were linked to their word roots following SALT conventions (e.g., *corre* 'runs' and *corrió* 'ran' were both linked to *correr* 'to run'). This ensures that the measures of word use were not inflated by the presence of multiple forms of single words. The only other basic grammatical structure that was marked in Spanish was noun plural. In English, samples were coded for basic grammatical structures such as noun plural, possessive, and verb tense markers in English. All irregular verb forms were linked to their root forms. This was consistent with the treatment of verb forms in Spanish in SALT. It also ensured that children were not over credited for lexical diversity when they produced creative verbs forms such as *sawed* and *seened* in addition to *saw* and *sees* in their stories.

In addition, several project specific codes for grammaticality were used. All complete and intelligible utterances were classified as influenced, grammatical or ungrammatical. Utterances were marked as influenced if children codeswitched or if the children produced an utterance with a dislocated subject. Consistent with our focus on the relationship between production in each language and language ability, influenced utterances were eliminated from the analysis. This resulted in the elimination of 11% of utterances from the English samples and 12.5% of utterances from Spanish samples. Determination of utterances as ungrammatical was based on a preset list of errors in English and Spanish. Appendices A and B contain lists of the error types that resulted in an utterances being coded as ungrammatical in English or Spanish. Only utterances containing these errors were classified as ungrammatical. For example, some utterances contained imprecise word choices but these were not classified as ungrammatical as long as they did not contain the grammatical errors from these lists. To verify reliability of the grammaticality 10% of the samples were randomly selected for recoding. Intercoder reliability was 96.66% for Spanish and 98.15% for English.

All complete and intelligible utterances that were not marked as influenced were included in the analysis. Using the SALT for Research rectangular database function, the following measures were obtained for the analyzed data set (complete and intelligible utterances with influenced utterances excluded) number of utterances, number of different words, and mean

length of utterance in words. Counts were generated of the number of grammatical and ungrammatical utterances. Based on these measures, we calculated the percentage of grammatical utterances (grammatical utterances/grammatical and ungrammatical utterances). The means and standard deviations for each of the predictor variables are summarized by language in Table 4.

3. Results

3.1. The associations among predictor variables

To address the question of relationships between measures Pearson correlation coefficients for the predictor variables are reported in Table 5. The correlation coefficients among all variables ranged from −0.23 to 0.88 (see Table 5). Correlations between the BESA ability score and MLU, NDW, and Grammaticality in English were low but significant. Similarly the correlations between BESA ability *z*-score and Spanish MLU and Grammaticality were low but statistically significant. However, the correlation between the BESA ability *z*-score and number of utterances in English and Spanish were not significant. Within the narrative measures, the correlation between NDW and Number of Utterances were high and the correlation with MLU was moderate and significant in English as well as Spanish.

3.2 Regression analysis

Three multiple regression models were explored to address the question of which combination of language sample measures converge with language ability as measured by the BESA. The goal of the multiple-regression analysis was to determine a parsimonious model that explained the most variance in the dependent variable containing the fewest number of independent variables. Three selection methods, *forward selection, backward selection,* and the *stepwise selection* were used to determine the target multiple-regression model in this study. Table 6 lists the model comparison results based on the selection methods.

The stepwise multiple-regression method was adopted because it resulted in the most parsimonious model, and it combined forward and backward selection methods. Specifically, the stepwise regression selected predictors by adding variables when they made a significant contribution to the model and removed variables already in the model when they were not significant. Furthermore, the R^2 of the model with 3 predictors was nearly identical to the R^2 of the full model. There was no need to add additional 5 predictors to the model.

The stepwise multiple-regression results are listed in Table 7. The model with three independent variables in predicting the BESA *z*-scores was statistically significant ($F_{3,166}$ = 19.70, p <.0001). The variables English MLU, English Grammaticality, and Spanish grammaticality accounted for 24.92% of the variance in the BESA *z*-score. It should be noted that English MLU contributed more to the BESA *z*- score variance than the other two predictors in terms of the relatively higher beta coefficient.

4. Discussion

This work explores the extent to which language sample measures from both of a bilingual's languages converge with language ability as indexed by a standardized test (the *BESA*). One unique feature of our study is the way in which we used both Spanish and English across the domains to develop a composite that incorporated both languages. Measures of language of productivity and sentence organization correlated with the *BESA z*-score although the total number of utterances produced in the narrative samples did not. The combination of language measures that best converged on language ability was MLU and grammaticality in English and grammaticality in Spanish.

The measures of productivity UTT and NDW were not significant predictors of ability in this analysis in either language, although NDW in English was correlated with ability as indexed by the BESA. This is consistent with the idea that children need to have some critical amount of vocabulary to be able to use the language to complete the testing tasks (Bohman, Bedore, Peña, Mendez-Perez, & Gillam, in press; Castilla, Restrepo, & Perez-Leroux, 2009; Simón-Cereijido & Gutiérrez-Clellen, 2009). The finding that NDW was not a significant predictor is not inconsistent with past literature. One problem with using NDW for narrative discourse is that the there may be too much overlap between the children with low and high ability.

MLU in Spanish and English were significantly correlated with the children's language ability as measured by the *BESA.* However, when Spanish and English MLU were put into the same model only MLU in English accounted for sufficient variance to enter the model as a predictor of language ability as measured by the BESA Z score. The finding that English MLU was a useful predictor of ability is consistent with findings from studies of monolingual and bilingual children with language impairment. The finding that Spanish was correlated with ability but did not enter the model as a predictor might seem inconsistent with past findings that it was a useful predictor of language impairment (Simón-Cereijido & Gutiérrez-Clellen, 2007). However, it is important to keep in mind that in this study both Spanish and English predictors were evaluated. The children in this study had been acquiring Spanish since birth and flatter patterns of development are observed in children with and with out language impairment (relative to English). But as these children are starting to learn English it is possible that the ability to form utterances and communicate emerges before the children refine their grammatical knowledge of the language. Thus the ability to start to produce utterances may be an earlier indicator of ability than grammaticality at this point in second language learning.

English and Spanish grammaticality were both significant predictors in the regression model, but English grammaticality contributed slightly more to the model than did Spanish grammaticality. Following the logic discussed above, grammaticality could be expected to develop after the ability to produce simple utterances. English learning children demonstrate higher rates of grammatical errors compared to Spanish learners. The combination of challenges related to producing narratives and creating grammatically-correct sentences may have resulted in more grammatical errors in English. This finding parallels the findings of Fiestas and Peña (2004) in study of 4–6 year bilinguals had greater exposure to Spanish than to English. These children were able to produce stories of equal complexity as measured by story grammar elements but produced more Spanish-influenced utterances in English, their relatively less proficient language. Another reason that grammaticality may contribute significantly to the regression model is the range of variation observed in the measure. On average children's grammaticality in Spanish was higher than their grammaticality in English but the standard deviations were similar for both languages. Finally, grammaticality was the only Spanish factor of those explored in this analysis that contributed to ability. When grammaticality is low in both languages, clinicians should be concerned about low language ability.

Finally, it is important to consider differences in the way that language ability was defined in the current study in comparison to other studies using these predictors. In this study we measured language ability in children with a full range of language abilities based on the children's semantics and morphosyntactic skills in both languages. Thus, the finding that sentence organization measures were better predictors than NDW was not due to a match between morphosyntax measures (e.g., Simón-Cereijido & Gutiérrez-Clellen, 2007).

In sum, this work represents an initial step in developing a procedure for evaluating language sample measures in two languages for bilingual kindergarteners. Predictors from both languages converged with a general language ability rating indexed by the *BESA.* Some

predictors, such as NDW and MLU in Spanish, that might be expected to be useful based on literature about monolingual development did not account for significant amount of variance in an overall measure of language ability in our sample of bilingual children. This may have occurred because bilingual children's performance on these measures was too variable. Future work should consider the extent to which degree of bilingual development or language dominance needs to be considered. It is also important to consider the extent to which the same predictors would be informative over a broader age span. Refinements of the language sample measures that were coded in this study might also help make these measures more sensitive to overall indices of language ability.

Acknowledgments

This work was supported by the grant 1 R01 DC007439-01 from the National Institute on Deafness and Other Communication Disorders (NIDCD). We are grateful to the families that participated in the study. We would also like to thank all of the interviewers for their assistance with collecting the data for this project and the school districts for allowing us access to collect the data. This report does not necessarily reflect the views or policy of NIDCD.

References

Anderson R, Souto S. The use of articles by monolingual Puerto RicanSpanish-speaking children wth specific language impairment. Applied Psycholinguistics 2005;26:621–647.

Bedore LM. Assessing Morphosyntax in Spanish-speaking children. Seminars in Speech and Language 2001;22(1):65–77. [PubMed: 11302457]

Bedore LM, Leonard LB. Specific language impairment and grammatical morphology: A discriminant function analysis. Journal of Speech, Language, & Hearing Research 1998;41(5):1185–1192.

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

Bedore LM, Leonard LB. Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. Applied Psycholinguistics 2005;26(2): 195–225.

Bedore LM, Peña ED. Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. International Journal of Bilingual Education and Bilingualism 2008;11(1):1–29.

Bedore LM, Peña ED, García M, Cortez C. Conceptual versus monolingual scoring: When does it make a difference? Speech, Language, Hearing Services in Schools 2005;36:188–200.

Berman, RA.; Slobin, DI. Relating events in narrative: A crosslinguistic developmental study. Hillsdale, NJ: Lawrence Erlbaum Associates; 1994.

Bland-Stewart LM, Fitzgerald SM. Use of Browns 14 grammatical morphemes by bilingual Hispanic preschoolers: a pilot study. Communication Disorders Quarterly 2001;22(4):171–186.

Bohman T, Bedore LM, Peña ED, Mendez-Perez A, Gillam RB. What you hear and what you say: Language performance in Spanish English bilinguals. International Journal of Bilingual Education and Bilingualism. in press. 10.1080/13670050903342019

Brown, R. A First Language: The Early Stages. Cambridge: Harvard University Press; 1973.

Brownell, R. Expressive One-Word Picture Vocabulary Test Spanish-Bilingual Edition. Novato, CA: Academic Therapy Publications; 2001.

Castilla A, Restrepo MA, Perez-Leroux AT. Individual differences and language interdependence: A study of sequential bilingual development in Spanish-English preschool children. International Journal of Bilingual Education and Bilingualism 2009;99999(1)

Damico JS, Oller JW Jr. The diagnosis of language disorders in bilingual children: Surface-oriented and pragmatic criteria. Journal of Speech and Hearing Disorders 1983;48:385–394. [PubMed: 6645433]

Fiestas CE, Peña ED. Narrative discourse in bilingual children: Language and task effects. Language, Speech, & Hearing Services in Schools 2004;35(2):155–168.

Gillam RB, Johnston JR. Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. Journal of Speech & Hearing Research 1992;35(6):1303–1315. [PubMed: 1283612]

Gutiérrez-Clellen VF. Narratives in two languages: Assessing performance of bilingual children. Linguistics and Education 2002;13(2):175–197.

Gutiérrez-Clellen VF, Kreiter J. Understanding child bilingual acquisition using parent and teacher reports. Applied Psycholinguistics 2003;24(2):267–288.

Gutiérrez-Clellen VF, Restrepo MA, Bedore LM, Peña ED, Anderson RT. Language sample analysis in Spanish-speaking children: Methodological considerations. Language, Speech & Hearing Services in the Schools 2000;31(1):88–98.

Gutiérrez-Clellen VF, Restrepo MA, Simón-Cereijido G. Evaluating the discriminant accuracy of a grammatical measure with Spanish-speaking children. Journal of Speech, Language, and Hearing Research 2006;49(6):1209–1223.

Gutiérrez-Clellen VF, Simón-Cereijido G. The discriminant accuracy of a grammatical measure with Latino English-speaking children. Journal of Speech, Language, and Hearing Research 2007;50(4): 968–981.

Gutiérrez-Clellen VF, Simón-Cereijido G, Wagner C. Bilingual children with language impairment: A comparison with monolinguals and second language learners. Applied Psycholinguistics 2008;29:3–19.

Hadley PA. Language-sampling protocols for eliciting text-level discourse. Language, Speech & Hearing Services in the Schools 1998;29(3):132–147.

Hewitt LE, Hammer CS, Yont KM, Tomblin JB. Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. Journal of Communication Disorders 2005;38(3):197. [PubMed: 15748724]

Jacobson PF, Schwartz RG. Morphology in incipient bilingual Spanish-speaking preschool children with specific language impairment. Applied Psycholinguistics 2002;23(1):23.

Jacobson PF, Schwartz RG. English past tense use in bilingual children with language impairment. American Journal of Speech Language Pathology 2005;14(4):313–323. [PubMed: 16396614]

Jax, VA. Unpublished doctoral dissertation. University of California; Los Angeles: 1988. Narrative construction by children learning English as a second language: A precursor to reading comprehension.

Klee T. Developmental and diagnostic characteristics of quantitative measures of children's language production. Topics in Language Disorders 1992;12(2):28–41.

Klee T, Stokes SF, Wong A, Fletcher P, Gavin WJ. Utterance length and linguistic diversity in Cantonese-speaking children with and without language impairment. Journal of Speech, Language, and Hearing Research 2004;47:1396–1410.

Kohnert K, Bates E, Hernandez A. Balancing bilinguals: Lexical-semantic production and cognitive processing in children learing Spanish and English. Journal of Speech, Language, and Hearing Research 1999;42(6):1400–1413.

Kohnert K, Danahy K. Young L2 learners' performance on a novel morpheme task. Clinical Linguistics & Phonetics 2007;21(7):557–569. [PubMed: 17564857]

Leadholm, B.; Miller, J. Language Sample Analysis: The Wisconsin Guide. Madison, WI: Bureau for Exceptional Children, Wisconsin Department of Public Instruction; 1995.

Leonard, LB. Children with Specific Language Impairment. Cambridge, MA: Bradford Books; 1998.

Leonard LB, Eyer JA, Bedore LM, Grela BG. Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. Journal of Speech, Language, and Hearing Research 1997;40(4):741–753.

Loban, W. Language development: kindergarten through grade twelve. Urbana, IL: National Council of Teachers of English; 1976.

Mayer, M. Frog, where are you?. New York, NY: Penguin Books; 1969.

Mayer, M. Frog goes to dinner. New York, NY: Penguin books; 1974.

Miller, J. Assessing Language Production in Children. Boston: Allyn and Bacon; 1981.

Miller J, Heilmann J, Nockerts A, Iglesias A, Fabiano L, Francis DJ. Oral language and reading in bilingual children. Learning Disabiities Research and Practice 2006;21:30–43.

Miller, J.; Iglesias, A. Systematic Analysis of Language Transcripts (SALT) Research Version 2008. Madison, WI: SALT Software LLC; 2008.

Muñoz ML, Gillam RB, Peña ED, Gulley-Faehnle A. Measures of language development in fictional narratives of Latino children. Language, Speech & Hearing Services in the Schools 2003;34:332–342.

Owen AJ, Leonard LB. Lexical diversity in the spontaneous speech of children with specific language impairment: Application *of D*. Journal of Speech, Language, and Hearing Research 2002;45(5):927–937.

Paradis J, Crago M, Genesee F, Rice ML. French-English bilingual children with SLI: How do they compare with their monolingual peers? Journal of Speech, Language, & Hearing Research 2003;46 (1):113.

Pearson BZ, Fernández S, Oller DK. Lexical development in bilingual infants and toddlers:comparison to monolingual norms. Language and Learning 1993;43(1):93–120.

Peña, ED.; Bedore, LM. Bilingualism and Language Impairment in Children. In: Schwartz, R., editor. Handbook of Child Language Disorders. New York: Psychology Press; 2008. p. 281-307.

Peña ED, Bedore LM, Rappazzo C. Comparison of Spanish, English,and bilingual children's performance across semantic tasks. Language, Speech & Hearing Services in the Schools 2003;34(1):5–16.

Peña ED, Bedore LM, Zlatic-Guinta R. Category-generation performance of bilingual children: The influence of condition, category, and language. Journal of Speech, Language and Hearing Research 2002;45:938–947.

Peña, ED.; Gutierrez-Clellen, V.; Iglesias, A.; Goldstein, B.; Bedore, LM. Bilingual English Spanish Assessment (BESA). indevelopment

Peña, ED.; Stubbe Kester, E. Semantic Development in Spanish-English Bilinguals: Theory, Assessment, and Intervention. In: Goldstein, B., editor. Bilingual Language Development and Disorders in Spanish-English Speakers. Baltimore: Brookes; 2004.

Redmond SM. Conversational profiles of children with ADHD, SLI and typical development. Clinical Linguistics and Phonetics 2004;18(2):107–125. [PubMed: 15086133]

Restrepo MA. Identifiers of predominantly Spanish-speaking children with language impairment. Journal of Speech, Language, & Hearing Research 1998;41(6):1398–1411.

Rice ML, Wexler K. Toward tense as a clinical marker of specific language impairment in English-speaking children. Journal of Speech & Hearing Research 1996;39(6):1239–1257. [PubMed: 8959609]

Rice ML, Wexler K, Hershberger S. Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. Journal of Speech, Language, and Hearing Research 1998;41(6):1412–1431.

Schlyter S. Bilingual children's stories: French passé composé/imparfait and their correspondences in Swedish. Linguistics 1996;34(5):1059–1085.

Simón-Cereijido G, Gutiérrez-Clellen VF. Spontaneous language markers of Spanish language impairment. Applied Psycholinguistics 2007;28(2):317–339.

Simón-Cereijido G, Gutiérrez-Clellen VF. A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. Applied Psycholinguistics 2009;30(2): 315–337. [PubMed: 19444336]

Simonsen, A. Working Papers. Vol. 50. Lund University: Department of Linguistics; 2002. Naming amongst 6 year old children; p. 79-84.

Uccelli P, Páez MM. Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. Language, Speech, and Hearing Services in Schools 2007;38(3):225–236.

Watkins RV, Kelly DJ, Harbers HM, Hollis W. Measuring children's lexical diversity: Differentiating typical and atypical language learners. Journal of Speech and Hearing Research 1995;34:1134–1141.

Westby C, Van Dongen R, Maggart Z. Assessing narrative competence. Seminars in Speech and Language 1989;10:63–75.

Wiig, EH.; Secord, WA.; Semel, E. Clinical Evaluation of Language Fundamentals-4 Spanish. San Antonio: Pearson Asessment; 2006.

Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool Language Scale-4 Spanish Edition. San Antonio: The Psychological Corporation; 2002.

CEU Questions

1. Narratives are often recommended as an assessment task for bilingual children because
    - **a.** Narratives are representative of children's conversational speech abilities.
    - **b.** Narratives are a good source of language productivity data.
    - **c.** Narratives are a culturally and linguistically appropriate language-sampling task.
    - **d.** Narratives provide information about children's ability to produce grammatical utterances.
2. To evaluate the language ability of bilingual children clinician's should
    - **a.** Attend only to the child's first language.
    - **b.** Use information only about the second language.
    - **c.** Interview parents about the child's abilities because testing is not appropriate,
    - **d.** Look for ways to systematically inform decisions based on data from both languages.
3. The BESA *z score* was based on children's better performance in semantics and morphosyntax in any combination of English or Spanish scores.
    - **a.** Because individual language scores are not meaningful.
    - **b.** Because children may show mixed dominance while they are acquiring a second language.
    - **c.** Because it is important to consider both language simultaneously.
    - **d.** A and B.
    - **e.** B and C
4. Number of different words correlated with language ability but was not a significant predictor of language ability. It is likely that is true because:
    - **a.** Because the measure may not be sensitive to small changes.
    - **b.** Because children with LI use as many words as their typically developing peers.
    - **c.** Because growth is NDW is likely to have a flat growth in bilingual children.
    - **d.** None of the above.
5. Grammaticality ratings were based on different errors sets in English and Spanish because:
    - **a.** Children have higher levels of grammaticality in Spanish.

b. Because children have difficulty with different grammatical structures in English and Spanish.

c. Children have a lower MLU in Spanish than in English.

d. All of the above.

Response Key

1. C

2. D

3. E

4. A

5. B

Table 1

Developmental change in lexical productivity measures based on a single narrative

Study	Age	English		Spanish	
		TNW M (SD)	NDW M (SD)	TNW M (SD)	NDW M (SD)
Muñoz et al. (2003	4;4	117.58 (57.73)	48.58 (16.12)	--	--
	5;3	209.83 (160.39)	69.75 (35.25)	--	--
Miller et al. (2006)	6;0	--	67.3 (25.0)	--	69.4 (19.9)
	7;0	--	75.8 (26.0)	--	79.2 (21.9)
Uccelli &Páez(2007)	5;9	38.29 (24.23)	21.38 (13.07)	35.79 (36.30)	20.79 (14.92)
	6;8	45.67 (21.13)	25.54 (36.30)	42.50 (23.67)	24.92 (11.14)

Note. TNW = Total number of words; NDW = Number of different words

Table 2

Change in MLU in bilingual children

Study	Age	English	Spanish
		MLU M (SD)	MLU M (SD)
Muñoz et al. (2003)	4;4	4.85 (.97)	--
	5;3	5.85 (.66)	--
Castilla, Restrepo, and Pérez-Leroux (2009)	4;8	--	4.51 (1.02)
Simon-Cereijido & Gutierrez-Clellen (2009)	5;2	5.96 (1.09)	5.63 (1.13)
Miller et al. (2006)	6;0	6.1 (1.1)	5.3 (.91)
	7;0	6.3 (1.0)	5.7 (.87)

Note. MLU = Mean length of utterance

Table 3

Percentage of children in each configuration of higher language z-scores contributing to BESA Language Ability Score

Semantics	Morphosyntax		Percent
	Cloze	Sentence repetition	
Spanish	Spanish	Spanish	21.0
English	Spanish	Spanish	33.3
Spanish	English	Spanish	8.2
Spanish	Spanish	English	1.7
English	English	English	18.2
Spanish	English	English	7.1
English	Spanish	English	3.0
English	English	Spanish	7.1

Table 4

Mean values (and standard deviations) for language sample measures in English and Spanish

Measure	English	Spanish
Utterances	26.57 (13.29)	19.17 (15.36)
MLU - words	5.04 (1.24)	4.41 (1.41)
NDW	51.91 (21.88)	40.80 (23.29)
Grammaticality	46.09 (25.83)	67.86 (24.96)

Note. MLU = Mean length of utterance; NDW = Number of different words

Table 5

Correlations between language sample measures and BESA language sample ability

Measure	BESA ability (Z)	English UTT	English MLU	English NDW	English Gram	Spanish UTT	Spanish MLU	Spanish NDW	Spanish Gram
BESA ability (Z)	1.00								
English UTT	0.12	1.00							
English MLU	0.46***	0.23**	1.00						
English NDW	0 29***	0.79***	0.57***	1.00					
English Gram	0.36***	0.09	0.41***	0.33***	1.00				
Spanish UTT	−0.05	0.12	0.00	0.05	−0.23**	1.00			
Spanish MLU	0.18*	0.06	0.26***	0.12	0.03	0.40***	1.00		
Spanish NDW	0.05	0.10	0.13	0.12	−0.12	0.88***	0.65***	1.00	
Spanish Gram	0 29***	−0.09	0.20**	0.03	0.03	0.12	0.43***	0 27***	1.00

Note.

* $p < .05$,

** $p < .01$,

*** $p < .001$;

BESA Ability (z) = Bilingual English Spanish Assessment Language Ability Score; UTT = Total number of utterances, MLU = Mean Length of Utterances words; NDW= Number of different words; Gram = Grammaticality

Table 6

Multiple regression model comparisons based on four selection methods

Selection Method	Number of predictors	*F*	*p*-value	R^2
Full	8	7.48	< .0001	0.2709
Forward	4	15.04	< .0001	0.2672
Backward	3	19.7	< .0001	0.2625
Stepwise	3	19.7	< .0001	0.2625

Table 7

Predictors of English MLU, English Grammaticality, and Spanish Grammaticality to the average BESA z-score: Multiple Regression

	Beta Weight	Beta	*p*-value	Adj. R^2	*F*	*p*-value
Intercept	0	−1.575	< .001	0.2492	19.7	<.0001
English MLU	0.350	0.222	< .001			
English Grammaticality	0.156	0.005	0.031			
Spanish Grammaticality	0.195	0.006	0.005			

Note. MLU = Mean length of utterance

J Commun Disord. Author manuscript; available in PMC 2011 November 1.

Appendix A

Utterances containing these English errors were coded ungrammatical

Error type	Sample errors
Omission	
Article omission	**The* dog went over there.
Preposition omission	Dog (and girl no) and boy fell **on* the floor.
Possessive pronoun omission	The boy lost **his* frog.
Verb omission	(And she) she **is* not hurt.
Number agreement	
Article number substitution	*A* frogs go over there.
Demonstrative pronoun number substitution	*Those* frog jumped high.
Possessive pronoun gender substitution	The boy put the frog in *her* shoe.
Verb tense	
Verb tense omission	Then he *see* a deer.
Verb over-regularization	He *runned* fast.
Verb person substitution	He *saws* a deer
Other	
Preposition substitution	Boy put the frog *to* the bucket.
Pronoun case substitution	*Him* went over there.

Appendix B

Utterances containing these Spanish errors were classified as ungrammatical

Error type	Sample error	Gloss
Omissions		
Article omission	La rana se fue con **el* niño	The frog went away with **the* boy.
Preposition omission	El niño está jugando *con la rana	The boy is play **with* the frog
Possessive pronoun omission	La rana mamá dijo que la rana chiquita era **su* bebé.	The mother frog said the little frog was **her* baby
Verb omission	El niño *verbo el instrumento	The boy **verb* the instrument.
Number agreement errors		
Article number substitution	*La* ranas se fue con el niño.	*The-sing* frogs left with the boy
Adjective number substitution	El sapo y el niño están *dormido.*	The toad and the boy are *asleep-singular*
Clitic pronoun number substitution	Vió el sapo y *los* agarró.	(he) Saw the frog and grabbed *it-plural.*
Possessive pronoun number substitution	El niño vió *su* amigos.	The boy say *his-sing* friends
Verb number substitution error	Las ranas *está* en el agua	The frogs *is-sing* in the water
Gender agreement		
Article gender substitution	*El* rana se fue con el niño.	*The-masc frog-fem* left with the boy.
Adjective gender substitution	El sapo y el niño están *dormidas.*	The toad-masc and the boy-masc are *asleep-fem.*
Clitic pronoun gender substitution	Vió el sapo y *la* agarró.	(He) saw the toad-masc and grabbed *it-fem.*
Demonstrative pronoun gender substitution	El niño se tropezó con *estos* avispas.	The boy ran into *those- masc* wasps-fem.
Possessive pronoun gender substitution	Esta cama es *mío.*	This bed-fem is *mine*-masc.
Verb tense		
Tense substitution error	Anoche la rana *sale.*	Last night the frog *gets-pres* out.
Other		
Preposition substitution	Y el venado llevó al niño *a* los cuernos.	And the deer carried the boy *to* the horns.
Pronoun case substitution	Se los dió a *lo.*	(He) gave them to *it-dative*