Published in final edited form as:

Am J Speech Lang Pathol. 2014 November 1; 23(4): 574–586. doi:10.1044/2014_AJSLP-13-0026.

Conceptual scoring of receptive and expressive vocabulary measures in simultaneous and sequential bilingual children

Megan Gross, Milijana Buac, and Margarita Kaushanskaya

University of Wisconsin-Madison

Abstract

Purpose—This study examined the effects of conceptual scoring on the performance of simultaneous and sequential bilinguals on standardized receptive and expressive vocabulary measures in English and Spanish.

Method—Participants included 40 English-speaking monolingual children, 39 simultaneous Spanish-English bilingual children, and 19 sequential bilinguals, ages 5–7. The children completed standardized receptive and expressive vocabulary measures in English and also in Spanish for bilinguals. After the standardized administration, bilinguals were given the opportunity to respond to missed items in their other language to obtain a conceptual score.

Results—Controlling for group differences in socioeconomic status (SES), both simultaneous and sequential bilinguals scored significantly below monolinguals on single-language measures of English receptive and expressive vocabulary. Conceptual scoring removed the significant difference between monolinguals and simultaneous bilinguals in the receptive modality, but not in the expressive modality; differences remained between monolinguals and sequential bilinguals in both modalities. However, in both bilingual groups conceptual scoring increased the proportion of children with vocabulary scores within the average range.

Conclusions—Conceptual scoring does not fully ameliorate the bias inherent in single-language standardized vocabulary measures for bilinguals, but the procedures employed here may assist in ruling out vocabulary deficits, particularly in typically-developing simultaneous bilingual children.

“The bilingual is not two monolinguals in one person” (Grosjean, 1989). Grosjean issues this cautionary note to neurolinguists, but the message is equally important for speech-language pathologists and other educational professionals working with bilingual children. Expecting a bilingual child to show the same vocabulary knowledge as a monolingual child in each language is not realistic because bilingual children generally show distributed vocabulary knowledge; there are some concepts for which they know the corresponding word only in language A and other concepts they know only in language B (Pearson, 1998). For example, a child may know the word *frijoles* in Spanish, but not its translation equivalent *beans* in English; similarly, a child may know *subtraction* in English, but not its translation equivalent *resta* in Spanish. If, during a language assessment, a speech-language pathologist considers a bilingual child’s single-language skills only, there is a risk of over-

identification of language impairment (e.g., Bedore, Peña, Garcia, & Cortez, 2005; Kohnert, 2010; Pearson, 1998; Teoh, Brebner, & McCormack, 2012). Conceptual scoring has been suggested as a less-biased alternative to single-language measures for bilinguals (Pearson, Fernandez, & Oller, 1993).

Conceptual scoring considers the total number of concepts for which a child has a word in at least one language. For example, on a conceptual vocabulary measure, a child would receive three points for knowing *frijoles*, *subtraction*, *mano*, and *hand* because these words reflect knowledge of three concepts (beans, subtraction, hand), even though one word is known only in Spanish and one is known only in English. Although conceptual scoring has received general support from empirical work as a method to obtain a more complete picture of a bilingual child's lexical knowledge (Atkins-Burnett & Aikens, 2008; Bedore et al., 2005; Hemsley, Holm, & Dodd, 2010; Junker & Stockman, 2002; Kan & Kohnert, 2005; Mancilla-Martinez, Pan, & Vagh, 2011; Mancilla-Martinez & Vagh, 2013; Marchman & Martinez-Sussmann, 2002; Paez, 2008; Pearson et al., 1993; Sheng, Lu, & Kan, 2011; Sheng, Peña, Bedore, & Fiestas, 2012; Wang, Castilleja, Sepulveda, & Daniel, 2011), there remain some uncertainties (e.g., Thordardottir, Rothenberg, Rivard, & Naves, 2006) as to the utility of conceptual scoring across different bilingual populations, across different testing modalities (receptive vs. expressive), and across different languages. In the current study, we examine the use of conceptual scoring with standardized measures of receptive and expressive vocabulary in Spanish and English in both simultaneous and sequential typically-developing bilingual children. The study addresses two main goals. First, we investigate whether conceptual scoring can close the gap in vocabulary scores between bilingual children and their monolingual peers. Second, we consider test-specific factors (i.e., language and modality) and child-specific factors (i.e., acquisition history and SES) that may affect the performance of bilingual children when conceptual scoring procedures are used. An understanding of these factors and how they interact with conceptual scoring can assist speech-language pathologists in reducing the risk of over-identifying deficits when measuring vocabulary skills in typically-developing bilingual children.

The Monolingual-Bilingual Gap on Single-Language Vocabulary Measures as a Reflection of Distributed Vocabulary Knowledge in Bilinguals

Bilingual children have repeatedly been found to score below their monolingual peers (e.g., Ben-Zeev, 1977; Bialystok, Luk, Peets, & Yang, 2010; Hemsley, Holm, & Dodd, 2006; Leseman, 2000) or below monolingual norms (e.g., Fernandez, Pearson, Umbel, Oller, & Molinetmolina, 1992; Mancilla-Martinez et al., 2011; Teoh et al., 2012; Uccelli & Paez, 2007; Uchikoshi, 2006; Umbel, Pearson, Fernandez, & Oller, 1992) on single-language vocabulary measures in one or both of their languages. Oller and colleagues (2007) highlight the distributed nature of lexical knowledge in bilinguals as a key contributor to their lower single-language vocabulary scores when compared to monolinguals. Due to context-specific learning, there may be certain areas of vocabulary to which children are only exposed in one of their languages, resulting in "singlet" vocabulary knowledge. For example, a child who is exposed to cooking vocabulary in a home where Spanish is spoken may know the word *sartén* in Spanish, but not *frying pan* in English. Conversely, a child who speaks English at

school may know the word *desk* in English, but not *escritorio* in Spanish. In this case, *sartén* and *desk* would be singlets in the child's vocabulary. Other areas of vocabulary may be experienced in both languages, and as a result children may possess "doublet" vocabulary knowledge (e.g., they know both *mano* and *hand*). Bilingual children often exhibit a greater proportion of singlets than doublets, especially in their productive vocabularies (Conboy & Thal, 2006; Deuchar & Quay, 2000; Hemsley et al., 2010; Junker & Stockman, 2002; Kan & Kohnert, 2005; Mancilla-Martinez et al., 2011; Pearson, Fernandez, & Oller, 1995; Peña, Bedore, & Zlatić-Giunta, 2002; Sheng et al., 2011). Because of this distributed nature of lexical-semantic knowledge, measuring bilinguals' vocabulary skills in only one of their languages is likely to underestimate their knowledge, even if the language being tested appears to be their dominant language (Kan & Kohnert, 2005; Pearson et al., 1993). Therefore, measures of bilingual vocabulary knowledge need to take into account the child's vocabulary knowledge across both languages.

Conceptual Scoring as a Measure of Lexical-Semantic Knowledge across Languages

Conceptual scoring, which considers the number of concepts for which a child has a word in any language instead of counting language-specific lexical items, has been suggested as a tool for a more accurate assessment of bilinguals' vocabulary performance. This approach was first described by Pearson and colleagues (1993) who demonstrated that, while single-language scores for expressive vocabulary were lower for bilingual than for monolingual toddlers, the conceptual scores did not differ, suggesting that these scores may be a better way to assess vocabulary knowledge in bilinguals. Conceptual scoring has been implemented for a variety of language assessment procedures, including parent report measures in toddlers (Junker & Stockman, 2002; Mancilla-Martinez et al., 2011; Mancilla-Martinez & Vagh, 2013; Marchman & Martinez-Sussmann, 2002; Patterson, 1998; Pearson et al., 1993; Rescorla & Achenbach, 2002; Thordardottir et al., 2006), criterion-referenced measures of vocabulary (Hemsley et al., 2010; Kan & Kohnert, 2005; Sheng et al., 2011) and semantic skills (Bedore et al., 2005; Sheng et al., 2012), and standardized assessments (Atkins-Burnett & Aikens, 2008; de Abreu, Baldassi, Puglisi, & Befi-Lopes, 2013; Paez, 2008; Wang et al., 2011). The standardized assessments that incorporate conceptual scoring include the Receptive and Expressive One-Word Picture Vocabulary Tests – Spanish Bilingual Edition (Brownell, 2000), the Bilingual Verbal Ability Tests (Muñoz-Sandoval, Cummins, Alvarado, & Ruef, 1998), and the Preschool Language Scales 5th Edition Spanish (Zimmerman, Steiner, & Pond, 2012).

In previous studies, conceptual scoring has been shown to boost the vocabulary scores of bilingual children to the level of their monolingual peers (Hemsley et al., 2010; Junker & Stockman, 2002; Pearson et al., 1993; Thordardottir et al., 2006; but for an exception see de Abreu et al., 2013), to improve the classification accuracy of lexical-semantic measures used with bilingual children (Bedore et al., 2005; Mancilla-Martinez et al., 2011; Sheng et al., 2012), and to result in significant score improvements relative to single-language scoring (Atkins-Burnett & Aikens, 2008; Bedore et al., 2005; de Abreu et al., 2013; Hemsley et al., 2010; Kan & Kohnert, 2005; Paez, 2008; Sheng et al., 2011; Wang et al., 2011). However,

these positive findings must be qualified depending on whether the measure under consideration was designed for use with the child's first (L1) or second language (L2) (Atkins-Burnett & Aikens, 2008; Bedore et al., 2005; Hemsley et al., 2010; Mancilla-Martinez & Vagh, 2013; Sheng et al., 2011; Thordardottir et al., 2006), and depending on the modality of the measure (i.e., receptive vs. expressive; Hemsley et al., 2010; Sheng et al., 2011). In addition, characteristics of the bilingual samples in these prior studies, including acquisition history (sequential vs. simultaneous) and socioeconomic background, have varied across studies. Thus, it is important to consider test-specific and child-specific factors that may influence outcomes of conceptual scoring.

Factors Influencing the Vocabulary Performance of Bilingual Children

Test language

The benefits of conceptual scoring over single-language scoring have been shown to vary depending on the language under consideration. Bedore and colleagues (2005), who used the Bilingual English Spanish Assessment (BESA), and Atkins-Burnett and Aikens (2008), who used the Expressive One-Word Picture Vocabulary Test, found greater gains when the base language of the measure was Spanish and conceptual scoring allowed the child to switch into English. Similarly, Sheng and colleagues (2011) found that the conceptual scores of Mandarin-English bilinguals in their sample represented a significant improvement over their single-language scores in Mandarin, but not over their single-language scores in English. In contrast, Kan and Kohnert (2005) found significant gains from conceptual scoring over single-language scores in both Hmong and English. These patterns could reflect the dominance profiles of the children, social prestige issues surrounding the perceived acceptability of responding in the minority language, or a combination of the two. In the present study, we administered both English and Spanish measures of vocabulary to English-Spanish bilingual children. This enabled us to contrast the impact of allowing for Spanish responses during testing in English and allowing for English responses during testing in Spanish.

Modality

The majority of available studies on the utility of conceptual scoring have focused on the expressive modality (e.g., Atkins-Burnett & Aikens, 2008; Bedore et al., 2005; Junker & Stockman, 2002; Patterson, 1998; Rescorla & Achenbach, 2002). Those studies that have examined both receptive *and* expressive skills have varied in whether gains from conceptual scoring were observed in both modalities (Kan & Kohnert, 2005) or only in receptive skills (Hemsley et al., 2010). In the study of Samoan-English bilinguals by Hemsley and colleagues, the children's conceptual expressive vocabulary scores did not differ significantly from their single-language expressive abilities in English. However, for receptive vocabulary, the children did know words in Samoan that they did not know in English; as a result, conceptual scoring resulted in significant gains over single-language scoring and removed the statistically significant difference between bilingual and monolingual receptive vocabulary scores. Because conceptual scoring can have differential effects depending on the modality, in the present study we implemented conceptual scoring for both receptive and expressive vocabulary measures in both English and Spanish.

Acquisition history

Studies of language development in bilingual children often distinguish *simultaneous* bilinguals, who begin acquiring both languages in infancy, from *sequential* bilinguals, who receive exposure to their second language after infancy, often at school entry (Kohnert, 2010). Differences have been identified in the language acquisition patterns of these two groups. For example, Hammer and colleagues (Hammer, Lawrence, & Miccio, 2008) documented differences between simultaneous and sequential bilinguals in their English and Spanish receptive vocabulary scores at the beginning and end of a two-year period in Head Start. During the school-age years, a phenomenon of shifting dominance has been documented in sequential bilinguals such that a gap emerges between their receptive and expressive skills in the home language, with continued receptive development and plateaued expressive skills (Gibson, Oller, Jarmulowicz, & Ethington, 2012; Hemsley et al., 2010; Kan & Kohnert, 2005; Oller et al., 2007; Sheng et al., 2011). However, previous studies of conceptual scoring have generally investigated *either* simultaneous *or* sequential bilinguals rather than drawing direct comparisons between the two groups. Specifically, studies of conceptual scoring using parent checklists have generally been conducted with young simultaneous bilinguals (Junker & Stockman, 2002; Mancilla-Martinez et al., 2011; Mancilla-Martinez & Vagh, 2013; Marchman & Martinez-Sussmann, 2002; Pearson et al., 1993; Thordardottir et al., 2006), while studies of conceptual scoring using criterion-referenced and standardized measures in preschool and school-age children have generally been conducted with sequential bilinguals (Atkins-Burnett & Aikens, 2008; de Abreu et al., 2013; Hemsley et al., 2010; Kan & Kohnert, 2005; Sheng et al., 2011). These differences in age and measures make it difficult to compare the outcomes for simultaneous and sequential bilinguals across studies. In the present study, therefore, we examined the impact of conceptual scoring on vocabulary performance in two groups of bilingual children: those who had acquired both languages from infancy and those who had acquired English only after age 3.

Socioeconomic status

Although the gap observed between the single-language scores of monolingual and bilingual children can be at least partially attributed to the distributed nature of bilinguals' lexical knowledge, the role of socioeconomic status (SES) also needs to be considered. The effect of SES on vocabulary development has been long established in monolingual children (e.g., Engel, Santos, & Gathercole, 2008; Hart & Risely, 1995; Rescorla & Achenbach, 2002) and has also been documented in bilingual children (e.g., Hakuta, Butler, & Witt, 2000; Oller et al., 2007). In the United States, it is a demographic reality that bilingual children from immigrant families often have lower SES than their monolingual peers and than the children included in the standardization samples for test norms (e.g., Mancilla-Martinez et al., 2011; Uccelli & Paez, 2007; Umbel & Oller, 1994). However, some studies that have controlled for SES have still found vocabulary discrepancies between bilingual and monolingual children (e.g., Fernandez et al., 1992; Hemsley et al., 2006; Leseman, 2000; Umbel & Oller, 1994; Umbel et al., 1992). When conceptual scoring was used with parent checklists to examine the vocabulary growth of Spanish-English bilingual toddlers from low-SES homes, their conceptual scores were below both English and Spanish monolingual norms (Mancilla-

Martinez & Vagh, 2013). Thus, it is important to consider how SES and the use of conceptual scoring may interact with each other to influence vocabulary outcomes in bilingual children. Previous conceptual scoring studies have varied in whether they focused on bilingual children from low-SES families (Hemsley et al., 2010; Kan & Kohnert, 2005; Mancilla-Martinez et al., 2011; Mancilla-Martinez & Vagh, 2013), targeted bilingual children from mid-SES families (Junker & Stockman, 2002; Pearson et al., 1993; Sheng et al., 2011), or did not specifically consider SES. In the present study, we included bilingual children from diverse SES backgrounds and statistically controlled for these differences in SES when examining the effects of conceptual scoring on the performance of the bilingual children in our sample.

The Current Study

Previous studies of conceptual scoring have yielded mixed results depending on the language and the modality of the measure used. Characteristics of the bilingual samples, including acquisition history and SES, have also varied across studies. Therefore, the goal of the present study was to examine the contributions of these factors to the effects of conceptual scoring. We were especially interested in examining the utility of conceptual scoring in the context of standardized vocabulary assessment tools.

While many previous studies have applied conceptual scoring to different types of assessments, only a few studies examined conceptual scoring in the context of standardized assessments frequently used with school-age children. The practice policy statement from the American Speech-Language Hearing Association (ASHA) with regard to the assessment of culturally and linguistically diverse children highlights the need for “appropriate use of published test materials in language assessment including standardized norm-referenced tests and criterion-referenced tests, including analyzing normative sampling limitations, general psychometric issues... and inherent cultural and linguistic biases in these test materials” (American Speech-Language Hearing Association, 2004). However, while clinicians may be aware that the monolingual norms on standardized assessments are not appropriate for use with bilingual children, they are often less certain of what assessment procedures they should use instead, especially when measures are not widely available in both languages (Thordardottir et al., 2006), as is the case for some of the languages most commonly spoken in the U.S., including Chinese, Tagalog, Vietnamese, and Korean (Ryan, 2013).

If a particular measure is not available in the child’s home language, one option might be to implement conceptual scoring during the administration of a test originally published in the majority language. However, using a test designed for a single language, even in conjunction with conceptual scoring, raises some concerns. First, the items may not be culturally appropriate for use in two languages. Second, translating the items into the other language for use during conceptual scoring may not preserve the original difficulty hierarchy on which the basal and ceiling rules of the test are based. Third, test language effects may interact with the other factors, including test modality, child’s acquisition history, and SES. In order to understand the possible benefits of conceptual scoring, it is necessary to systematically consider how these different factors may affect bilingual

vocabulary performance within the context of a single study. The current study was designed to fill this need and to address the following two research questions:

1. When group differences in SES are taken into account, does conceptual scoring yield receptive and expressive vocabulary scores among typically developing bilingual children that are statistically equivalent to those obtained by their monolingual peers?
2. How do test language, modality, SES and bilingual acquisition history modulate the impact of conceptual scoring on the vocabulary performance of typically-developing school-age bilingual children?

Method

Participants

The 98 participants in the current study were drawn from a larger project on novel word learning and vocabulary knowledge in monolingual and bilingual children between the ages of 5 and 7. The study took place in Madison, WI and Chicago, IL. The children were recruited through the Madison public schools by distributing informational fliers printed in Spanish and English and by disseminating information through community centers, including Latino cultural centers and churches. Three groups of children were included in the current study: 1) 40 monolingual¹ English-speaking children with no more than 5% exposure to another language during a typical week; 2) 39 simultaneous Spanish-English bilingual children who began producing 2-word phrases in both languages before age 3, according to parent report; and 3) 19 sequential Spanish-English bilingual children who began producing 2-word phrases in English after age 3, according to parent report. We chose the production of word combinations to index age of acquisition because this production milestone would be a salient marker for parents to facilitate consistency in reporting. The three groups did not differ in age or nonverbal intelligence. They did, however, differ in socioeconomic status (SES). Specifically, the sequential bilingual children came from families with significantly lower SES than both the monolingual and simultaneous bilingual children. SES was indexed by the years of education completed by the primary caregiver, in accordance with common practice in studies of language development (e.g., de Abreu et al., 2013; Junker & Stockman, 2002; Marchman & Martinez-Sussmann, 2002; Sheng et al., 2011) and in line with general trends for measuring SES in recent child development and education research (Ensminger & Fothergill, 2003; Sirin, 2005). In addition to these differences in SES, the three groups came from different racial/ethnic backgrounds, with the majority of monolinguals being Caucasian, the majority of simultaneous bilinguals being Latino or multi-racial, and the majority of sequential bilinguals being Latino. Table 1 presents the demographic characteristics of the three groups.

The simultaneous and sequential bilingual groups differed in their language background characteristics. On average, compared to the simultaneous bilingual group, the children in

¹In the larger project, 68 children met the criteria for inclusion in the monolingual English-speaking group. However, to accommodate the smaller size of the two bilingual groups ($n = 39$ and $n = 19$), 40 of the 68 monolingual children were randomly selected for the current study.

the sequential bilingual group were exposed to Spanish more during a typical week and were more likely to use mostly Spanish in the home. A clear majority of the simultaneous bilingual children preferred English, according to parent report, while the sequential bilingual children were more divided between preferring English or Spanish. Table 2 presents the language background characteristics for the two bilingual groups.

Exclusion criteria included diagnosed language impairment, learning disabilities, psychological/behavioral disorders, neurological impairment, and other developmental disabilities. The majority of children passed a pure tone hearing screening at 25 dB at 1000 Hz, 2000 Hz, and 4000 Hz bilaterally. For 8 children (4 simultaneous bilinguals, 2 sequential bilinguals, and 2 monolinguals), hearing screening equipment was not available at the time of testing. However, their primary caregivers reported no hearing-related concerns.

Measures

Vocabulary—To measure English receptive vocabulary, all children completed the *Peabody Picture Vocabulary Test-III* (Form B) (PPVT-III, Dunn & Dunn, 1997). To measure Spanish receptive vocabulary, the bilingual children completed the *Test de Vocabulario en Imágenes Peabody* (TVIP, Dunn, Padilla, Lugo, & Dunn, 1986). Following the recommendations in the technical manual, the *normas hispanas compuestas* (composite Hispanic norms) were used to derive standard scores. The PPVT-III was selected over the more recent PPVT-IV because it is closer in form to the TVIP, which was adapted from the PPVT-R, but still represents improvements over the PPVT-R in reducing cultural bias. This combination of the PPVT-III and the TVIP has been used in recent studies of receptive vocabulary in Spanish-English bilingual children (e.g., Gibson et al., 2012; Hammer et al., 2008). To measure English expressive vocabulary, all children completed the *Picture Vocabulary* subtest of the *Woodcock-Johnson III Tests of Achievement* (Form A) (Woodcock, McGrew, & Mather, 2001). The norms used to derive standard scores came from the normative update (Woodcock, McGrew, Schrank, & Mather, 2007). To measure Spanish expressive vocabulary, the bilingual children completed the *Vocabulario sobre dibujos* subtest of the *Batería III Woodcock-Muñoz Pruebas de aprovechamiento* (Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005), which is an adaptation, rather than a translation, of the *Picture Vocabulary* subtest in English from the *Woodcock-Johnson III Tests of Achievement*. The English receptive and expressive measures were normed on separate demographically representative samples based on the U.S. Census, but the bilingual status of the children in the samples is not stated. The Spanish receptive and expressive measures were normed on separate groups of monolingual Spanish speakers or, for a subset of the *Batería III* sample, individuals who were clearly Spanish-dominant at the time of testing.

For the bilingual children, both *language-specific* and *conceptual* scores were obtained for each vocabulary measure. The *language-specific scores* were obtained by administering each assessment in its original language (i.e., English for English measures and Spanish for Spanish measures) according to the standardized procedures. The *conceptual scores* were obtained by returning to any items the child had answered incorrectly and re-administering them in the opposite language. For example, after the standard administration of the PPVT

in English, the examiner told the child, “Now we are going to do some in Spanish.” For receptive measures, the child heard a translation of the missed item and was asked to point to the matching picture. For the expressive measures, the child was asked to name the missed item in the opposite language (e.g., “¿Qué es esto?” instead of “What’s this?”), and the response was checked against acceptable translations for the test item. Translations were obtained for all test items using a Spanish-English dictionary (Rollin, Carvajal, & Horwood, 2005) and *SpanishDict*, an online dictionary, and were verified by consulting two native speakers of Spanish. For each measure, items answered correctly during conceptual scoring were added to the language-specific raw score to obtain a conceptual raw score that reflected word knowledge across languages. These conceptual raw scores were in turn converted to standard scores using the test’s norms. Although conceptual scoring was not part of the procedure used to develop these norms, standard scores were derived as a point of comparison (Pearson et al., 1993) to determine where the child’s conceptual scores would fall relative to the monolingual normative sample.

Nonverbal intelligence—The children completed the *Visual Matrices* subtest of the *Kaufman Brief Intelligence Test* (KBIT-2, Kaufman & Kaufman, 2004) to assess their nonverbal intelligence. In accordance with the KBIT-2 administration procedure, instructions to the children were presented verbally in the child’s preferred language (English or Spanish).

Hearing screening—The children’s hearing was screened bilaterally at 25 dB at 1000, 2000, and 4000 Hz using a GSI-17 audiometer in the testing room.

Procedure

The majority of the testing sessions took place in a laboratory at the Waisman Center; however, 6 children were tested in a quiet room in a community space and 4 children were tested in a quiet room in their home. The measures of vocabulary and nonverbal intelligence were administered to the children by trained research assistants. The research assistants were instructed to vary and record the order in which they administered the tasks so as to avoid introducing consistent order effects. All Spanish measures were administered by bilingual research assistants. For the bilingual children, the conceptual scoring procedure was employed after the standardized administration of each vocabulary measure. Parents filled out the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld, & Kaushanskaya, 2007) about themselves, and they were also interviewed face to face in their preferred language about their child’s developmental history. They provided information about their own educational history, as well as information about their child’s education, language use and exposure, relevant medical history, and family background.

Analyses

In the first set of analyses, we focused on the comparisons among monolinguals, simultaneous bilinguals, and sequential bilinguals on their vocabulary scores in English. These analyses examined whether the use of conceptual scoring procedures with standardized vocabulary assessments could reduce differences between monolinguals and bilinguals. To take into account the potential impact of significant group differences in SES,

we included SES (as measured by total years of education completed by the primary caregiver) as a covariate in analyses of covariance (ANCOVA). First, we compared the groups on their language-specific scores on the English receptive and expressive vocabulary measures. Second, we compared the scores obtained by the monolingual group to the conceptual scores obtained by the simultaneous and sequential bilingual groups. For all analyses, significant main effects of group were further examined in pairwise comparisons using the Bonferroni adjustment for multiple comparisons.

To examine the factors that modulate the effects of conceptual scoring, we conducted a repeated measures ANCOVA with scoring method (language-specific vs. conceptual), modality (receptive vs. expressive), and original test language (English vs. Spanish) as within-subject factors and group (simultaneous vs. sequential bilinguals) as the between-subjects factor. Given the significant difference in SES between the two bilingual groups, SES served as a covariate.

Finally, to consider the clinical utility of conceptual scoring for providing a more accurate picture of the performance of typically-developing bilingual children, we implemented an analysis modeled on that of Bedore and colleagues (2005). Bedore et al. (2005) identified the proportion of children correctly classified as typically developing with single-language vs. conceptual scoring. The authors considered classification below 80% as poor, 80–90% as fair, and at least 90% as good. We computed the proportion of children in each bilingual group who obtained language-specific scores of at least 85 (i.e., within or above average) on the English and Spanish measures of receptive and expressive vocabulary. Then we computed the proportions when conceptual scoring procedures were employed. Finally, we considered the proportion of children scoring within the average range on receptive measures in at least one language and on expressive measures in at least one language using language-specific and conceptual scores.

Results

Monolingual/Bilingual Comparisons

Given the significant difference in SES among groups, all group comparisons of language-specific and conceptual vocabulary scores from the English measures were conducted as analyses of covariance (ANCOVA) with SES as the covariate.²

English receptive vocabulary language-specific scores—The ANCOVA revealed a main effect of group, $F(2, 94) = 19.9, p < .001, \eta_p^2 = .30$. Follow-up comparisons revealed that the monolinguals outperformed the simultaneous bilinguals, $p = .002$, and the sequential bilinguals, $p < .001$, and the simultaneous bilinguals outperformed the sequential bilinguals, $p = .001$.

²When we conducted the group comparisons without controlling for SES, the patterns of results was the same as those obtained from the ANCOVA for the comparison between the monolingual and bilingual groups. The only result that differed between the two analyses was the comparison between simultaneous and sequential bilinguals on conceptual scores from the English expressive vocabulary measure; the group difference was significant without SES as a covariate, but it was not significant once SES differences between the two bilingual groups were controlled in the ANCOVA. Given that the comparisons of interest in this section were between the monolingual group and each bilingual group, we report the results from the ANCOVA.

English expressive vocabulary language-specific scores—There was a significant main effect of group, $F(2, 94) = 15.8, p < .001, \eta_p^2 = .25$. All follow-up comparisons between groups were significant. Monolinguals outperformed simultaneous bilinguals, $p = .002$, and sequential bilinguals, $p < .001$, and simultaneous bilinguals outperformed sequential bilinguals, $p = .012$.

English receptive vocabulary with conceptual scoring—A significant main effect of group was obtained using conceptual scores, $F(2, 93) = 8.6, p < .001, \eta_p^2 = .16$. Follow-up comparisons revealed that the simultaneous bilinguals' conceptual scores did not differ from the monolinguals' scores, $p = .375$. Sequential bilinguals' conceptual scores were significantly lower than monolinguals' scores, $p < .001$, and simultaneous bilinguals' conceptual scores, $p = .008$.

English expressive vocabulary with conceptual scoring—A significant main effect of group was obtained using conceptual scores, $F(2, 93) = 13.7, p < .001, \eta_p^2 = .23$. Monolinguals outperformed both the simultaneous bilinguals, $p < .001$, and the sequential bilinguals, $p < .001$, but the difference between the two bilingual groups was not significant, $p = .529$. Figure 1 shows the language-specific and conceptual scores for all three groups.

Factors Modulating the Effect of Conceptual Scoring for Bilingual Children

A 2 (scoring method) x 2 (modality) x 2 (original test language) x 2 (group) repeated measures ANCOVA with SES as a covariate examined factors that may modulate the impact of conceptual scoring for the two bilingual groups. There was a significant main effect of scoring method, $F(1, 47) = 28.2, p < .001, \eta_p^2 = .31$. Overall, children received significantly higher scores with conceptual scoring ($M_{\text{adjusted marginal}} = 102.4, SE = 1.5$) than with language-specific scoring ($M_{\text{adjusted marginal}} = 90.5, SE = 1.6$). There was a main effect of language, $F(1, 47) = 9.38, p = .004, \eta_p^2 = .17$, with children receiving higher scores in English ($M_{\text{adjusted marginal}} = 97.5, SE = 1.7$) than in Spanish ($M_{\text{adjusted marginal}} = 95.3, SE = 2.1$). A significant interaction between language and modality, $F(1, 47) = 8.58, p = .005, \eta_p^2 = .15$, reflected a greater effect of language in the expressive modality (English $M_{\text{adjusted marginal}} = 93.4, SE = 1.6$ vs. Spanish $M_{\text{adjusted marginal}} = 88.5, SE = 2.3$) than in the receptive modality (English $M_{\text{adjusted marginal}} = 101.6, SE = 2.2$ vs. Spanish $M_{\text{adjusted marginal}} = 102.2, SE = 2.3$). However, there was no main effect of modality, $F(1, 47) = 1.71, p = .197, \eta_p^2 = .035$, and there was no interaction between modality and group, $F(1, 47) = 1.55, p = .219, \eta_p^2 = .032$.

The interactions between scoring method and the other variables under consideration (modality, language, group) were not significant (all $ps > .1$), indicating that these variables did not significantly modulate the effects of conceptual scoring, once SES was taken into account. Table 3 provides a summary of the language-specific and conceptual scores for each bilingual group on all four measures.

The clinical significance of the effect of conceptual scoring is considered in Table 4. The proportion of children in the simultaneous and the sequential bilingual groups who scored within or above the average range on each measure with and without the use of conceptual

scoring is shown. For each measure, conceptual scoring increased the proportion of typically-developing bilingual children scoring within or above the average range. For the simultaneous bilinguals, conceptual scoring resulted in “good” classification for all measures, with at least 90% of children scoring in the average range. This represents an improvement over single-language scoring, which yielded 90% classification only for the English receptive measure.

For the sequential bilinguals, single-language scoring yielded scores within the average range for fewer than 75% of the children on each measure. Conceptual scoring improved the percentage of children scoring within the average range to 84% for the Spanish expressive measure and 90% for the Spanish receptive measure. For the English measures, these proportions still fell below 75%. However, when considering conceptual scores derived from both English and Spanish measures, all sequential bilinguals scored within the average range on at least one receptive measure and 89% scored within the average range on at least one expressive measure.

Discussion

The current study examined conceptual scoring in the context of standardized assessments of receptive and expressive vocabulary in English and Spanish with two goals: (a) to determine whether conceptual scoring would reduce the discrepancy in vocabulary performance between bilingual and monolingual children in English, particularly when controlling for group differences in SES; (b) to examine the effects of base language of the assessment, modality, SES, and acquisition history on the gains afforded by conceptual scoring. Therefore, the current study addressed both theoretical and clinical questions. From a theoretical standpoint, given the phenomenon of distributed vocabulary knowledge, we were interested in whether using conceptual scoring, while statistically controlling for discrepancies in SES, could close the gap in vocabulary scores obtained by bilingual children and their monolingual peers. Clinically, we were interested in whether conceptual scoring could be used to reduce the risk of over-identifying vocabulary deficits in typically-developing bilingual children.

Monolingual/Bilingual Comparisons

Even controlling for differences in socioeconomic status, both simultaneous and sequential school-aged bilingual children were found to lag behind their monolingual peers on standardized measures of English receptive and expressive vocabulary, with sequential bilinguals also lagging behind simultaneous bilinguals. However, the effects of conceptual scoring diverged for the two bilingual groups. Specifically, conceptual scoring removed the significant difference between simultaneous bilinguals and monolinguals for receptive, but not expressive, vocabulary. This finding is similar to that of Hemsley and colleagues (2010) and may reflect the commonly documented gap between receptive and expressive vocabulary skills that emerges when bilingual children enter school (Gibson et al., 2012; Hemsley et al., 2010; Kan & Kohnert, 2005; Oller et al., 2007; Sheng et al., 2011; Uccelli & Paez, 2007). If expressive vocabulary skills in Spanish are limited, then allowing children to respond in Spanish during conceptual scoring of an English expressive vocabulary measure

may not yield much benefit. Yan and Nicoladis (2009) alternatively suggest that low expressive vocabulary scores in bilinguals may not reflect small expressive vocabularies, but rather difficulty accessing and retrieving the appropriate word in the moment of the assessment.

Sequential bilinguals lagged behind both monolinguals and simultaneous bilinguals on receptive and expressive vocabulary, even with the use of conceptual scoring. This finding is similar to that of Bedore and colleagues (2005) who observed that a significant difference remained between Spanish-dominant bilinguals and primarily English-speaking Latino children on the English subtest of the BESA even when conceptual scoring was employed. Moreover, in the current study, significant differences between the conceptual scores of sequential bilinguals and the scores of monolinguals were observed even though differences in SES were controlled statistically. This finding is in line with other studies (Fernandez et al., 1992; Hemsley et al., 2006, 2010; Leseman, 2000; Umbel & Oller, 1994; Umbel et al., 1992) in which group differences in vocabulary scores persisted when differences in SES were accounted for statistically or by careful group selection. However, these studies left open the possibility that the persisting vocabulary gap after controlling for SES was due to the phenomenon of distributed vocabulary. The current study demonstrated that even when conceptual scoring is implemented to account for distributed vocabulary knowledge *and* differences in SES are statistically controlled, sequential bilinguals still performed significantly worse than their monolingual English-speaking peers on both receptive and expressive vocabulary. Previous studies that have both implemented conceptual scoring and equated monolingual and bilingual groups on SES have obtained similar findings of persisting monolingual-bilingual gaps in expressive vocabulary in other language groups, including French-English bilinguals (Thordardottir et al., 2006) and Portuguese-Luxembourgish bilinguals (de Abreu et al., 2013). This persisting gap could reflect genuine differences between bilingual and monolingual children, in that bilingual children from lower SES backgrounds may know fewer words overall, even across both languages. Alternatively, it is possible that conceptual scoring does not fully account for the phenomenon of distributed vocabulary knowledge and/or that other factors (beyond SES and distributed vocabulary knowledge) contribute to the discrepancies in vocabulary observed between monolinguals and bilinguals.

Factors Modulating the Effects of Conceptual Scoring for Bilingual Children

Previous studies have identified differential effects of conceptual scoring depending on the language (i.e., L1 vs. L2; Atkins-Burnett & Aikens, 2008; Bedore et al., 2005; Hemsley et al., 2010; Mancilla-Martinez & Vagh, 2013; Sheng et al., 2011; Thordardottir et al., 2006) and modality (i.e., receptive vs. expressive; Hemsley et al., 2010; Sheng et al., 2011) of the measure under consideration. However, these studies also varied in whether they included sequential or simultaneous bilinguals, in the type of assessment they used, in whether they considered the impact of SES, and in how they measured the effects of conceptual scoring. Thus, there was a need to examine all of these factors more comprehensively within the context of a single study.

The analyses conducted on the two bilingual groups revealed that the children received higher scores in English than in Spanish, and that this was particularly true for expressive vocabulary. This finding highlights the phenomenon of reduced expressive vocabulary skills in the minority language in school-age bilinguals (Gibson et al., 2012; Hemsley et al., 2010; Kan & Kohnert, 2005; Oller et al., 2007; Sheng et al., 2011; Uccelli & Paez, 2007). Further, these analyses revealed that conceptual scoring resulted in a significant improvement in vocabulary scores for the bilingual children in this study³ and that this effect of conceptual scoring was not significantly modulated by test language (English vs. Spanish), test modality (receptive vs. expressive), or group (simultaneous vs. sequential). This finding is consistent with Kan and Kohnert (2005), who found a significant improvement from conceptual scoring in both Hmong and English and in both receptive and expressive vocabulary. However, this finding stands in contrast to studies that identified differential effects of conceptual scoring depending on language and/or modality of testing (e.g., Atkins-Burnett & Aikens, 2008; Bedore et al., 2005; Hemsley et al., 2010; Mancilla-Martinez & Vagh, 2013; Sheng et al., 2011; Thordardottir et al., 2006). The differences between prior studies (which have observed language and modality effects on conceptual scoring) and the current study (which did not) may lie in the fact that, while the present study considered the contributions of SES, prior studies had not. It is important to point out, however, that future work where SES is considered as an independent contributor to conceptual-scoring effects rather than as a covariate (as in the present study) is necessary in order to fully delineate the contribution of SES to conceptual-scoring effects across different languages and testing modalities.

While the current study documented significant score improvements as a result of conceptual scoring, it is important to ask whether these score improvements are of clinical significance in terms of reducing the risk of over-identification of vocabulary deficits in typically-developing bilingual children. Table 4 shows the proportion of children in the two bilingual groups whose scores fell within or above the average range on each measure with single-language vs. conceptual scoring. For the simultaneous group, when single-language scoring was employed, only the English receptive measure exhibited “good” classification (Bedore et al., 2005), with 90% of typically-developing children scoring within the average range. However, all measures exhibited at least 90% correct classification when conceptual scoring was employed. This finding suggests that the use of conceptual scoring in conjunction with frequently used standardized measures may help in ruling out the presence of a vocabulary deficit in typically-developing simultaneous bilingual children.

For the sequential bilingual group, single-language scoring resulted in poor classification (less than 80%) on all measures. With conceptual scoring, classification rose to 90% for the Spanish receptive measure and 84% for the Spanish expressive measure, but remained poor for the English measures. Thus, although test language did not modulate the effect of

³For the simultaneous bilingual group, only one child (3%) showed no improvement on the English receptive measure, but 27 children (71%) showed no improvement on the English expressive measure. All children showed improvement on the Spanish receptive measure. Four children (11%) showed no improvement on the Spanish expressive measure. For the sequential bilingual group, one child (5%) showed no improvement on the English receptive measure, 5 children (26%) showed no improvement on the English expressive measure, all children improved on the Spanish receptive measure, and 4 children (11%) showed no improvement on the Spanish expressive measure. The majority of the children, therefore, benefited from the opportunity to use both languages.

conceptual scoring in the Analysis of Covariance, it did affect the extent to which conceptual scoring boosted the performance of sequential bilinguals into the average range. This finding underscores the non-equivalence of tests designed in different languages to test the same construct (e.g., receptive vocabulary), as there may be differences in the type of items included and in the characteristics of the normative samples. This finding further demonstrates that implementing conceptual scoring together with standardized English vocabulary measures would not be effective in helping to rule out vocabulary deficits in typically-developing sequential bilinguals; even using conceptual scoring with standardized Spanish vocabulary measures has limited utility for sequential bilinguals in the area of expressive vocabulary. Thordardottir and colleagues (2006) similarly caution that conceptual scoring is not necessarily “an accurate yardstick of bilingual children’s language development when compared to monolingual norms” (p. 19) and highlight the need to conduct assessments in both languages.

It is important to note that, in the current study, when single-language measures were administered in *both* English and Spanish, more than 20% of the sequential bilinguals still scored below the average range in both languages. When the tests were administered in both languages using conceptual scoring, however, all sequential bilinguals obtained a score within the average range in at least one language for receptive vocabulary and 89% obtained a score within the average range in at least one language for expressive vocabulary. Thus, administering separate measures in each language may still over-identify children as having vocabulary deficits, but allowing children to respond in both languages during the administration of both English and Spanish measures may improve classification.

Limitations

The methodology of the present study necessarily limits the ability to generalize findings to other populations of children and to instantiations of conceptual scoring other than the very specific conceptual scoring approach implemented here. First, the small sample size (especially for the sequential bilingual group) may have limited our ability to detect significant effects in comparisons involving the sequential bilinguals. Second, similarly to Pearson and colleagues (1993), we used monolingual norms “for purposes of illustration and comparison” (p. 104). Although the use of monolingual norms is known to yield biased outcomes with bilinguals (e.g., Pearson, 1998) and is not in line with best practices (American Speech-Language Hearing Association, 2004), it is also the case that obtaining meaningful bilingual norms would be unrealistic given the great variability among bilinguals. Thus, we examined whether conceptual scoring could be used in conjunction with monolingual norms to counteract this exact bias.

Third, because we implemented conceptual scoring within the context of standardized measures, we translated the items on the tests into the language of conceptual scoring. This method is inherently problematic because translation equivalents often differ in difficulty level (driven by frequency of use; age of acquisition; phonotactic properties; etc.) and in cultural appropriateness. Therefore, the age-based hierarchy of items in the original language does not carry over to the translated items, and the rules for establishing basals and ceilings among the translated items cease to be meaningful. In the present study, the basal

and the ceiling were established in the language of administration, and conceptual scoring was derived based on the re-administration of items within that range. However, it is likely that the effects of conceptual scoring could have been even greater if we had established a new ceiling in the other language, especially when the original language was the child's weaker language. Procedures for implementing conceptual scoring in other standardized tests, such as the PLS-5 Spanish (Zimmerman et al., 2012) and the BVAT (Muñoz-Sandoval et al., 1998), have taken this approach in establishing a new ceiling. However, this alternative is also problematic given the nonequivalent difficulty of items across languages.

A fourth issue is that of possible testing effects from repeating items, particularly during receptive measures. If children realize that they are being asked to repeat missed items, this may allow them to eliminate their initial response from the response set, thereby increasing the probability of selecting the correct item by chance. We attempted to reduce the likelihood of this testing effect by waiting until the end of the standardized administration to repeat missed items so that children would be less likely to realize they had answered the item incorrectly and to remember their initial response.

Conclusions and Clinical Implications

The best practices in the assessment of language skills in a bilingual child entail a comprehensive set of procedures that involve gathering a detailed language history from the parents, observing the child in the classroom and other settings, and conducting direct assessments of the child (e.g., Kohnert, 2008). With regard to direct assessment, it is important to test both languages (e.g., Kohnert, 2010; Pearson, 1998). However, assessments are not always available in the child's home language. In addition, single-language assessments in both languages may still underestimate bilinguals' skills because of the distributed nature of lexical-semantic knowledge. Conceptual scoring, which combines children's lexical knowledge across their two languages, may serve to equalize the demands of vocabulary tasks for monolingual and bilingual children and reduce the risk of over-identification of deficits.

In the present study, we implemented a simple conceptual scoring method within the context of frequently administered standardized receptive and expressive vocabulary measures with both simultaneous and sequential bilingual children. Conceptual scoring removed the statistically significant difference between simultaneous bilinguals and monolinguals for receptive vocabulary. Conceptual scoring did not ameliorate the difference between sequential bilinguals and monolinguals, even when controlling for SES. However, conceptual scoring did result in a significant increase in the scores obtained by both bilingual groups. For simultaneous bilinguals, this score increase translated into improved classification. Allowing children to respond in Spanish during the administration of the English assessments yielded scores within the average range for over 90% of the simultaneous bilinguals. These results suggest that if tests in both languages are not available, speech-language pathologists may be able to use a standardized English measure with the assistance of an interpreter to re-administer missed items in the child's native language. While this approach would likely still over-identify some bilingual children if used to *rule in* language impairment, it may improve the accuracy with which clinicians can

rule out vocabulary deficits in simultaneous bilinguals. However, this procedure may not be appropriate for sequential bilinguals, as allowing them to respond in Spanish during the administration of English assessments still yielded below-average scores for over 25% of the children. Although time intensive, the use of measures in both languages, in combination with conceptual scoring, may be necessary for assessing vocabulary in sequential bilinguals if standardized vocabulary tests with monolingual norms are to be used during an assessment.

Acknowledgments

This research was supported by NIH grants R03 DC010465-01 and R01 DC011750 awarded to Margarita Kaushanskaya, Training Grant T32 HD049899-08 awarded to Megan Gross, and the NIH Diversity Supplement R03 DC0104565-01 and Training Grant T32 DC005359-10 awarded to Milijana Buac. The authors would like to thank Barbara Pearson for her insightful suggestions about the uses and limitations of conceptual scoring. We would like to acknowledge Christina Ausick for experiment development; Michelle Batko, Nicole Compt, Katie Engh, Regina Estrada, Kiran Gosal, Allison Holt, Shu-ting Hsieh, Liz Jaramillo, Hailey Kuettner, Eva Lopez, Jessica Martalock, Breana Mudrock, Nivi Nair, Sarah Naumann, Emily Silverberg, and Kris Wright for assistance with recruitment, data collection, and coding; and members of the Language Acquisition and Bilingualism Lab for helpful comments during the preparation of this paper. In addition, we are grateful to Sara Lopez, Eduardo Montoya (Iglesia Presbiteriana Ebenezer), and Gerson and Lillian Amaya (Iglesia Evangelica Hazieli) for providing community testing space. Finally, we deeply appreciate all of the children and parents who participated in the study.

References

- American Speech-Language Hearing Association. Knowledge and skills needed by speech-language pathologists and audiologists to provide culturally and linguistically appropriate services [Knowledge and Skills]. 2004. Available from www.asha.org/policy
- Atkins-Burnett, S.; Aikens, N. Use of Conceptual Scoring to Increase Reliability and Validity of Direct Child Assessments with Linguistically Diverse Preschoolers. Paper presented at the Head Start's Ninth National Research Conference; Washington, DC. 2008.
- Bedore LM, Peña ED, Garcia M, Cortez C. Conceptual versus monolingual scoring: When does it make a difference? *Language Speech and Hearing Services in Schools*. 2005; 36(3):188–200.10.1044/0161-1461(2005/020)
- Ben-Zeev S. The influence of bilingualism on cognitive strategy and cognitive development. *Child Development*. 1977; 48(3):1009–1018.10.1111/j.1467-8624.1977.tb01260.x
- Bialystok E, Luk G, Peets KF, Yang S. Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*. 2010; 13(04):525–531.10.1017/s1366728909990423
- Brownell, R. Receptive and Expressive One-Word Picture Vocabulary Tests - Spanish-Bilingual Edition. San Antonio, TX: The Psychological Corporation; 2000.
- Conboy BT, Thal DJ. Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development*. 2006; 77(3):712–735.10.1111/j.1467-8624.2006.00899.x [PubMed: 16686797]
- de Abreu PMJE, Baldassi M, Puglisi ML, Befi-Lopes DM. Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech Language and Hearing Research*. 2013; 56:630–642.10.1044/1092-4388(2012/12-0079
- Deuchar, M.; Quay, S. *Bilingual Acquisition: Theoretical Implications of a Case Study*. Oxford: Oxford University Press; 2000.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Test*. 3. Circle Pines, MN: AGS; 1997.
- Dunn, LM.; Padilla, ER.; Lugo, DE.; Dunn, LM. *Test de Vocabulario en Imágenes Peabody (Adaptación Hispanoamericana) [Peabody Picture Vocabulary Test Hispanic American version]*. Circle Pines, MN: AGS; 1986.

- Engel PMJ, Santos FH, Gathercole SE. Are working memory measures free of socioeconomic influence? *Journal of Speech, Language, and Hearing Research*. 2008; 51:1580–1587.
- Ensminger, ME.; Fothergill, K. A decade of measuring SES: What it tells us and where to go from here. In: Bornstein, MH.; Bradley, RH., editors. *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum; 2003. p. 13-28.
- Fernandez MC, Pearson BZ, Umbel VM, Oller DK, Molinetmolina M. Bilingual receptive vocabulary in Hispanic preschool children. *Hispanic Journal of Behavioral Sciences*. 1992; 14(2):268–276.10.1177/07399863920142006
- Gibson TA, Oller DK, Jarmulowicz L, Ethington CA. The receptive–expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism: Language and Cognition*. 2012; 15(Special Issue 01):102–116.10.1017/S1366728910000490
- Grosjean F. Neurolinguists, beware: The bilingual is not two monolinguals in one person. *Brain and Language*. 1989; 36(1):3–15. [PubMed: 2465057]
- Hakuta, K.; Butler, YG.; Witt, D. Policy Reports. Santa Barbara, CA: University of California Linguistic Minority Research Institute; 2000. How long does it take English language learners to attain proficiency?.
- Hammer CS, Lawrence FR, Miccio AW. Exposure to English before and after entry into Head Start: Bilingual children's receptive language growth in Spanish and English. *Bilingual Education and Bilingualism*. 2008; 11(1):30–56.10.2167/beb376.0
- Hart, B.; Risely, TR. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Brookes; 1995.
- Hemsley G, Holm A, Dodd B. Diverse but not different: The lexical skills of two primary age bilingual groups in comparison to monolingual peers. *International Journal of Bilingualism*. 2006; 10(4): 453–476.10.1177/13670069060100040401
- Hemsley G, Holm A, Dodd B. Patterns in diversity: Lexical learning in Samoan-English bilingual children. *International Journal of Speech-Language Pathology*. 2010; 12(4):362–374.10.3109/17549501003721064 [PubMed: 20441411]
- Junker DA, Stockman IJ. Expressive vocabulary of German-English bilingual toddlers. *American Journal of Speech-Language Pathology*. 2002; 11(3):381–394.10.1044/1058-0360(2002/042)
- Kan PF, Kohnert K. Preschoolers learning Hmong and English: Lexical-semantic skills in L1 and L2. *Journal of Speech Language and Hearing Research*. 2005; 48(2):372–383.10.1044/1092-4388(2005/026)
- Kaufman, AS.; Kaufman, NL. *Kaufman Brief Intelligence Test*. 2. Circle Pines, MN: American Guidance Service; 2004.
- Kohnert, K. *Language Disorders in Bilingual Children and Adults*. San Diego, CA: Plural Publishing; 2008.
- Kohnert K. Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*. 2010; 43(6):456–473.10.1016/j.jcomdis.2010.02.002 [PubMed: 20371080]
- Leseman PPM. Bilingual vocabulary development of Turkish preschoolers in the Netherlands. *Journal of Multilingual and Multicultural Development*. 2000; 21(2):93–112.10.1080/01434630008666396
- Mancilla-Martinez J, Pan BA, Vagh SB. Assessing the productive vocabulary of Spanish-English bilingual toddlers from low-income families. *Applied Psycholinguistics*. 2011; 32(2):333–357.10.1017/s0142716410000433
- Mancilla-Martinez J, Vagh SB. Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. *Early Childhood Research Quarterly*. 2013; 28(3):555–567. <http://dx.doi.org/10.1016/j.ecresq.2013.03.004>.
- Marchman VA, Martinez-Sussmann C. 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*. 2002; 45:983–997.
- Marian V, Blumenfeld H, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and monolinguals. *Journal of Speech Language and Hearing Research*. 2007; 50(4):940–967.

- Muñoz-Sandoval, AF.; Cummins, J.; Alvarado, CG.; Ruef, ML. Bilingual Verbal Ability Tests. Itasca, IL: Riverside; 1998.
- Muñoz-Sandoval, AF.; Woodcock, RW.; McGrew, KS.; Mather, N. Batería III Woodcock-Muñoz Pruebas de aprovechamiento. Rolling Meadows, IL: Riverside; 2005.
- Oller DK, Pearson BZ, Cobo-Lewis AB. Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*. 2007; 28:191–230. [PubMed: 22639477]
- Paez MA. English language proficiency and bilingual verbal ability among Chinese, Dominican, and Haitian immigrant students. *Equity & Excellence in Education*. 2008; 41(3):311–324.10.1080/10665680802177380
- Patterson JL. Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *Am J Speech Lang Pathol*. 1998; 7(4):46–56.
- Pearson BZ. Assessing lexical development in bilingual babies and toddlers. *The International Journal of Bilingualism*. 1998; 2(3):347–372.
- Pearson BZ, Fernandez S, Oller DK. Cross-language synonyms in the lexicons of bilingual infants - one language or two? *Journal of Child Language*. 1995; 22(2):345–368. [PubMed: 8550727]
- Pearson BZ, Fernandez SC, Oller DK. Lexical development in bilingual infants and toddlers - Comparison to monolingual norms. *Language Learning*. 1993; 43(1):93–120.
- Peña ED, Bedore LM, Zlatic-Giunta R. Category-generation performance of bilingual children: The influence of condition, category, and language. *Journal of Speech Language and Hearing Research*. 2002; 45(5):938–947.10.1044/1092-4388(2002/076)
- Rescorla L, Achenbach TM. Use of the Language Development Survey (LDS) in a national probability sample of children 18 to 35 months old. *Journal of Speech, Language, and Hearing Research*. 2002; 45(4):733–743.10.1044/1092-4388(2002/059)
- Rollin, N.; Carvajal, CS.; Horwood, J. *Diccionario Oxford Compact*. Oxford: Oxford University Press; 2005.
- Ryan, C. *American Community Survey Reports*. Washington, DC: US Census Bureau; 2013. Language use in the United States: 2011; p. ACS-22
- Sheng L, Lu Y, Kan PF. Lexical development in Mandarin-English bilingual children. *Bilingualism-Language and Cognition*. 2011; 14(4):579–587.10.1017/s1366728910000647
- Sheng L, Peña ED, Bedore LM, Fiestas CE. Semantic deficits in Spanish-English bilingual children with language impairment. *Journal of Speech Language and Hearing Research*. 2012; 55(1):1–15.10.1044/1092-4388(2011/10-0254)
- Sirin SR. Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*. 2005; 75(3):417–453.
- SpanishDict. 2008. from <http://www.spanishdict.com/>
- Teoh WQ, Brebner C, McCormack P. Assessing the language skills of children from culturally and linguistically diverse backgrounds: The expressive vocabulary performance of Singaporean English-Mandarin bilingual pre-schoolers. *International Journal of Speech-Language Pathology*. 2012; 14(3):281–291.10.3109/17549507.2011.652171 [PubMed: 22335605]
- Thordardottir E, Rothenberg A, Rivard ME, Naves R. Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*. 2006; 4(1):1–21.10.1080/14769670500215647
- Uccelli P, Paez 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):225–236.10.1044/0161-1461(2007/024)
- Uchikoshi Y. English vocabulary development in bilingual kindergarteners: What are the best predictors? *Bilingualism*. 2006; 9(01):33.10.1017/s1366728905002361
- Umbel VM, Oller DK. Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning*. 1994; 44(2):221.
- Umbel VM, Pearson BZ, Fernandez MC, Oller DK. Measuring bilingual children's receptive vocabularies. *Child Development*. 1992; 63(4):1012–1020. [PubMed: 1505238]

- Wang, S.; Castilleja, N.; Sepulveda, M.; Daniel, MH. Importance of Conceptual Scoring to Language Assessment in Bilingual Children. Paper presented at the 2011 ASHA Convention; San Diego, CA. 2011.
- Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III Tests of Achievement. Rolling Meadows, IL: Riverside Publishing; 2001.
- Woodcock, RW.; McGrew, KS.; Schrank, FA.; Mather, N. Woodcock-Johnson III Normative Update. Rolling Meadows, IL: Riverside Publishing; 2007.
- Yan S, Nicoladis E. Finding le mot juste: Differences between bilingual and monolingual children's lexical access in comprehension and production. *Bilingualism-Language and Cognition*. 2009; 12(3):323–335.10.1017/s1366728909990101
- Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool Language Scales, Fifth Edition Spanish. San Antonio, TX: The Psychological Corporation; 2012.

Figure 1.

Standard scores for English vocabulary measures obtained using language-specific and conceptual scoring procedures. Values represent marginal means adjusted to control for group differences in SES. Error bars indicate two standard errors. Asterisks indicate the presence of a significant difference from the monolingual group.

** $p < .01$, *** $p < .001$

Table 1

Demographic Characteristics for Participant Groups

Characteristic	Monolingual <i>M</i> (SD)	Simultaneous <i>M</i> (SD)	Sequential <i>M</i> (SD)
n	40	39	19
Age (years)	6.5 (0.87)	6.2 (0.75)	6.2 (0.89)
Gender	26 boys, 14 girls	15 boys, 24 girls	7 boys, 12 girls
Nonverbal Intelligence (KBIT-2)	107 (14.5)	103 (13.2)	99 (9.2)
SES (parent years of education)	18.2 (3.1) ^a	17.2 (4.5) ^a	12.7 (6.6) ^b
Racial/Ethnic Background			
Caucasian	77%	8%	0%
Latino	0%	51%	84%
African-American	3%	0%	0%
Asian	3%	0%	0%
Multi-racial	17%	41%	16%

^{a,b} For SES, the sequential bilinguals differed significantly from both the monolinguals ($p < .001$) and the simultaneous bilinguals ($p < .01$), controlling for multiple comparisons (Bonferroni).

Table 2

Language Background Characteristics for Simultaneous and Sequential Bilingual Groups

Characteristic	Simultaneous (n=39)	Sequential (n=19)
Age of Acquisition (months) ^a		
English	20 (7.4)	43 (10.3) *
Spanish	19 (7.2)	22 (10.7)
Current % of Exposure (% waking hrs in typical week)		
English	49 (17.1)	35 (18.6) *
Spanish	51 (17.1)	65 (18.5) *
School Language ^b		
English-only	49%	63%
At least 50% instruction in Spanish	51%	37%
Home Language ^b		
Mostly English	39%	5% *
Mostly Spanish	46%	79% *
Both	15%	16%
Child's Preferred Language ^b		
English	62%	47%
Spanish	18%	32%
Both	20%	21%

^a Age of Acquisition was indexed by the age in months at which the child began producing two-word phrases in each language, according to parent report.

^b Percentages reflect the percent of the sample in each category.

* $p < .05$ (t -test for comparing means for age of acquisition and exposure; z -test for comparing proportions for school language, home language, and preferred language)

Table 3
Language-Specific and Conceptual Vocabulary Scores for Simultaneous and Sequential Bilinguals

	Simultaneous Bilinguals				Sequential Bilinguals			
	Language-Specific <i>M</i> (SE)	Conceptual <i>M</i> (SE)	Difference <i>M</i> (SE)		Language-Specific <i>M</i> (SE)	Conceptual <i>M</i> (SE)	Difference <i>M</i> (SE)	
English Receptive	103.0 (2.6)	110.9 (2.6)	7.8 (0.8)		91.3 (3.8)	101.3 (3.8)	10.0 (1.2)	
English Expressive	94.7 (2.4)	98.6 (1.6)	3.9 (1.5)		85.5 (3.4)	94.7 (2.2)	9.2 (2.2)	
Spanish Receptive	97.6 (3.0)	108.4 (2.7)	10.8 (1.1)		97.4 (4.3)	105.3 (3.8)	7.9 (1.6)	
Spanish Expressive	75.9 (3.6)	99.7 (2.5)	23.8 (2.9)		78.3 (5.2)	100.0 (3.7)	21.7 (4.2)	

Note. Values are marginal means adjusted for group differences in SES, with standard errors shown in parentheses.

Table 4

Proportion of Children Obtaining a Standard Score 85 Using Language-Specific vs. Conceptual Scoring

	Simultaneous Bilinguals (n=39)		Sequential Bilinguals (n=19)	
	Language-Specific	Conceptual	Language-Specific	Conceptual
English Receptive	90%	95%	53%	68%
English Expressive	82%	92%	42%	74%
Spanish Receptive	85%	97%	74%	90%
Spanish Expressive	31%	92%	53%	84%
Both Receptive ^a	100%	100%	79%	100%
Both Expressive ^b	92%	100%	74%	89%

Note. For each measure, a standard score of at least 85 reflects performance within or above the average range relative to monolingual norms.

^aThe proportion of children receiving a standard score of at least 85 on *either* the English or the Spanish receptive measure.

^bThe proportion of children receiving a standard score of at least 85 on *either* the English or the Spanish expressive measure.