It Takes Two: Improving Assessment Accuracy in Bilingual Children

by Elizabeth D. Peña & Lisa M. Bedore

As the population of U.S. school children becomes more diverse, speech-language pathologists increasingly will face the challenge of accurately determining if bilingual children have expected language skills given their age and language experiences.

Assessment of bilingual children—that is, differentiating the performance of children with and without language impairment—is difficult at many levels, due to a number of factors: lack of developmental information about the many languages families speak, lack of information about the rate and order of bilingual language development, and lack of data identifying the most problematic language features for bilingual children with language impairment.

We know that we should consider both languages of a bilingual child when conducting an assessment (Bedore & Peña, 2008; Kohnert, 2010). However, only about 8% of school clinicians report having training in bilingual assessment (ASHA Schools Survey, 2008), and only 5% of SLPs report that they use a language other than English (ASHA, n.d.). Even a clinician who has the ability to test in two languages needs assessment materials and the information base on which to make assessment decisions.

In the Human Abilities in Language Acquisition (HABLA) Lab, we have systematically developed test materials to classify the language abilities of bilingual children. As part of our research on bilingualism and language impairment in children, we also have proposed models for clinical decision-making. In the past 15 years, we have faced a number of challenges in our efforts to develop bilingual measures. To help other clinicians assess bilingual children, we also have developed a combined, bilingual approach that may work in making diagnostic decisions.

Challenges

An early task identified in the field of bilingual assessment was understanding the developmental milestones of language-learning in bilingual children, and developing measures that accurately captured problematic forms for bilingual children with language impairment. As part of a contract from the National Institute on Deafness and Other Communication Disorders, we began to study—with collaborators Iglesias, Gutiérrez-Clellen, and Goldstein—a large set of items in English and Spanish that would potentially differentiate the performance of children with and without language impairment. The selection of items was informed by a growing literature on language development and language impairment in Spanish and in English. One question was the extent to which items that differentiate monolingual children with language impairment also differentiate bilingual children (Gutiérrez-Clellen, Simon-Cereijido, & Wagner, 2008).

Regardless of the domain, the kinds of items that will differentiate language impairment differ by language. In syntax, for example, tense-marking in English is hard—but easy in Spanish (Bedore & Leonard, 2000). In semantics, functions are easier in Spanish, but similarities and differences are easier in English (Peña, Bedore, & Rappazzo, 2003; Restrepo & Silverman, 2001).

In bilinguals, however, it appears that item difficulty is similar—but, in some instances, item difficulty varies between bilingual and monolingual children. For example, in Peña et al. (2003), monolingual English speakers scored 48% correct on expressive characteristic properties items (e.g., "Tell me three things about this truck"), but bilinguals scored 35%. When scores are distributed in this way, bilingual children's performance is more likely to be in the low-average or impaired range.

Assessment Development

We used an item-analysis approach to select test items to build test tasks, such as the Bilingual English Spanish Assessment (BESA; Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, in development) and the Bilingual English Spanish Oral Screeners (BESOS; Peña, Bedore, Iglesias, Gutiérrez-Clellen, & Goldstein, in development). Specifically, we selected those items that demonstrated evidence to
differentiate children with and without impairment. The BESOS, for example, consists of four short tests of Spanish semantics, English semantics, Spanish morphosyntax, and English morphosyntax. Each test contains 10–18 items that are sensitive to development and also separate children with and without impairment. Thus far, the BESA and BESOS have been administered in one or both languages to about 2,500 children across several projects, including 1,600 children tested in both Spanish and English. These data can help us determine whether testing children in two languages makes a difference.

Analysis of our findings shows that there is developmental change associated with the tests in each language. We explored several ways of using Spanish and/or English to differentiate among children with and without language impairment. A long-standing assumption is that language impairment can be identified more accurately in the child's first or dominant language. To test this assumption, we compared two groups of bilingual children (English-dominant and Spanish-dominant) in three different ways. First, we evaluated their performance using their English scores alone and then their Spanish scores alone. Next, we examined their performance in their dominant language. Finally, we examined performance in combining their best language in each of two areas (semantics and morphosyntax).

We used BESOS screening data from the Diagnostic Markers of Language Impairment grant (Peña, Gillam, Bedore, & Bohman, 2006) to illustrate these approaches. Children were all screened pre-kindergarten and were on average 5 years, 3 months (SD = 5 months). For this example, 68 children were included: 32 were English-dominant bilinguals (use and exposure to English 60%–80% of the time) and 36 were Spanish-dominant bilinguals (use and exposure to Spanish 60%–80% of the time). A total of 10 children were identified with language impairment.

We used a cutoff score of one standard deviation below the mean on either semantics or morphosyntax subtests of the BESOS across each analysis. We calculated the number of children correctly identified as impaired by the BESOS procedure (sensitivity) and those correctly identified as typically developing (specificity). Sensitivity and specificity rates of 80% are considered fair; 90% or more are considered good (Plante & Vance, 1994).

When we used the English version of the BESOS only with English-dominant and Spanish-dominant children together, sensitivity/specificity was 100%/34%. When we used the Spanish version of the test for both groups of children, sensitivity/specificity was 100%/50%. Although we expect that screeners will over-identify typical children as having risk for impairment, the English version of the test with both Spanish- and English-dominant children incorrectly identified 65% of the children as impaired. The Spanish version of the test incorrectly identified 50% of the children as impaired. These false-positive rates are unacceptably high.

We can improve classification by examining performance in a child's dominant language. Our results show we can achieve sensitivity/specificity of 100%/57% in English for English-dominant children and 100%/70% in Spanish for Spanish-dominant children. Thus, the value of testing in the child's better language is apparent. But there is room for improvement.

Cross-Language Testing

Even testing in both languages, however, presents a challenge: how to use the test results systematically to make decisions. We have begun to explore how performance across these tests can be combined to make decisions. It is relatively easy to make a decision when children score very low or very high in both languages, but it is relatively common for young children to do well in one language but achieve a marginal score in the other language or to achieve marginal scores in both. How should these mixed dominance scores be interpreted?

We have tested different ways of combining test scores across languages. Some possibilities are to combine scores across the two languages in a composite or to select combinations of better task or language performance to use as the basis for decision-making. Work in our lab with combinations of the two languages shows that classification can be more accurate when scores in both languages are used systematically for decision-making. We have explored taking the best score in each subtest (e.g., best syntax and best semantics), regardless of the language in which they occur. Preliminary analysis using a –1SD threshold with the same data set described above demonstrates that this approach yields a
sensitivity/specificity of 100%/81% for this subset of children. These preliminary results are highly promising.

Bilingual children need to be tested in both of their languages to improve classification accuracy. Through research with bilingual children with and without language impairment, we are beginning to find that we can improve the accuracy of assessments for bilingual children. If we test in only one language without regard to dominance we risk inappropriate over-identification of a large number of children. This over-identification is lessened when we test in the dominant language. But, if we test both languages and use them together, we reduce the high rate of over-identification. It is important to note that children can demonstrate mixed dominance across the different domains of language. Testing in one language, even the language the child uses the most, may not yield the most accurate and reliable results. Combining children's best performance across domains is promising for improving assessment practices for bilingual children.

Elizabeth D. Peña, PhD, CCC-SLP, is a professor of communication sciences and disorders at the University of Texas at Austin. Her research focuses on differentiating language impairment from language difference through dynamic assessment and development of assessment protocols for bilinguals. Contact her at lizp@mail.utexas.edu.

Lisa M. Bedore, PhD, CCC-SLP, is a professor of communication sciences and disorders at the University of Texas at Austin. Her research focuses on child language and phonological development and disorders with a special interest in Spanish-English bilingual children. She is a member of Special Interest Group 14, Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations. Contact her at lbedore@mail.utexas.edu.


References


