Learning Disabilities Research & Practice, *21*(1), 19–29
©2006, The Division for Learning Disabilities of the Council for Exceptional Children

Do Language-Based Processing Tasks Separate Children with Language Impairment from Typical Bilinguals?

Kathryn Kohnert, Jennifer Windsor, and Dongsun Yim
University of Minnesota

We report results from 2 language-based processing tasks designed to investigate the performance of linguistically diverse learners. The tasks were the Competing Language Processing Task (CLPT) and Non-Word Repetition (NWR). Participants were 100 school-age children in 1 of 3 different experimental groups: monolingual English-speaking children with specific or primary language impairment (LI), typical English-only-speaking children (EO), or typical Spanish–English bilingual children (BI). On both CLPT and NWR, EO group performance was best and LI group performance was poorest, with BI group performance falling in between. Likelihood ratios indicated that performance on these tasks does not provide compelling diagnostic power for separating typically developing bilinguals from monolingual children with LI. One exception is that children who obtained an NWR score of 93 percent or higher could be ruled out of the LI group with a high degree of confidence.

Primary or specific language impairment (LI) is present when there is a disproportionate deficit in language, in the face of otherwise normal development. This language deficit is determined relative to age peers who share similar cultural, linguistic, and educational experiences. LI is a high incidence developmental disorder that is not attributed to frank problems in motor, sensory, cognitive, neurological, social–emotional, or environmental systems (e.g., Leonard, 1998; Tomblin et al., 1997). A robust literature indicates that children with LI are at significant risk for concurrent and later difficulties with word identification or text comprehension or both (e.g., Bishop & Snowling, 2004; Tomblin, Zhang, Buckwalter, & Catts, 2000; Vellutino, Fletcher, Snowling, & Scanlon, 2004).

In both the case of spoken language and reading difficulties, it has been claimed that a deficit in phonological processing may be central, at least for languages like English that have an opaque orthography (Kohnert, Windsor, & Miller, 2004; see also Vellutino et al., 2004, for review). However, other subtle cognitive deficits in linguistic and nonlinguistic domains, including less efficient processing speed, working memory, and temporal integration, also have been raised as candidates underlying both spoken language and literacy problems for monolingual speakers (Kohnert & Windsor, 2004; Miller, Kail, Leonard, & Tomblin, 2001; Scarborough, 1998; Windsor & Hwang, 1999) and sequential bilingual language learners (Swanson, Sáez, Gerber, & Leafstedt, 2004).

For children with a diagnosis of LI, the increasing theoretical emphasis on underlying cognitive constructs has coincided with an increasing focus on language assessment procedures that also emphasize these cognitive–linguistic underpinnings rather than language knowledge per se. One presumed educational advantage of this shift in emphasis, and the focus of this study, is that these nontraditional linguistic processing measures may be less biased than traditional performance measures when applied to children from culturally and linguistically diverse populations (Laing & Kamhi, 2003).

Traditional measures used to identify LI include language sample analysis and norm-referenced standardized tests designed to gauge a child's expressive or receptive language skills compared to his or her peers. Performance on these tasks is critically dependent on a child's opportunities or experiences with the test language. As such, traditional language measures are sometimes considered knowledge- or experience-dependent (Campbell, Dollaghan, Needleman, & Janosky, 1997). Performance on many experience-dependent language assessment measures is effective in identifying LI among middle-income children from the dominant culture who are monolingual speakers of the language or dialect of the majority community and well represented in the available normative database. However, performance on experience-dependent tasks does not seem to be sensitive to the underlying differences between children with LI and typically developing children with diverse cultural or language backgrounds (e.g., Paradis & Crago, 2000; Dollaghan & Campbell, 1998). As a result, children from culturally or linguistically diverse backgrounds may be misidentified as learning disabled and underidentified for LI (Silliman, Wilkinson, & Brea-Spahn, 2004).

Language-based processing measures have been proposed as potentially nonbiased alternatives to traditional experience-dependent measures for the purpose of identifying LI in culturally or linguistically diverse learners. Language-based processing tasks deemphasize the role of

Requests for reprints should be sent to Kathryn Kohnert or Jennifer Windsor in the Department of Speech–Language–Hearing Sciences, 115 Shevlin Hall, 164 Pillsbury Dr. S.E., University of Minnesota, Minneapolis, MN 55455. Electronic inquiries may be sent to kohne005@umn.edu or windsor@umn.edu.

prior knowledge or experience by using linguistic units intended to be equally familiar to participants (such as high-frequency vocabulary) or equally unfamiliar to participants (such as nonsense words that do not exist in the test language) (Kohnert, 2004). The idea is to level the playing field and minimize the role that prior language experience may have on performance (Campbell et al., 1997). If the role of previous experience on task performance can be minimized at the same time that sensitivity to the presence of LI is maintained, then language-based processing tasks present a viable option for reducing assessment bias in a culturally or linguistically diverse population.

However, a recent study by Windsor and Kohnert (2004) suggests that not all language-based processing tasks differentiate the language profiles of children with LI and typically developing children learning a second language. Windsor and Kohnert administered spoken word recognition and picture naming tasks to monolingual English-speaking children with and without LI and to Spanish–English bilingual children without LI. Even for very common stimulus items, such as *pig*, *house*, or *hair*, neither word recognition accuracy nor response time separated monolingual children with LI from typical bilingual children. Similarly, there was equivalent performance between these two groups of children in the picture naming task, with both groups performing below that of their typical monolingual peers. These findings suggest that not all processing tasks that incorporate linguistic stimuli are equally effective in overcoming assessment bias.

The purpose of the current study is to determine if two prominent language-based processing measures separate children with LI from typically developing Spanish–English bilingual peers. The two tasks examined in this study are the Competing Language Processing Task (CLPT) and Non-Word Repetition (NWR). These tasks have received considerable attention in the literature for their potential to separate monolingual children with LI from their typically developing age and language-matched peers (e.g., Conti-Ramsden, 2003; Ellis Weismer, 1996; Ellis Weismer, Evans, & Hesketh, 1999; Gathercole & Baddeley, 1990). Performance on both CLPT and NWR tasks presumably relies on the processing of very basic linguistic information. There are differences, however, in the nature of the stimuli and the type of processing to be done across these two measures.

The CLPT is a listening span measure for children developed by Gaulin and Campbell (1994), based on the adult reading and listening span tasks of Daneman and Carpenter (1980). In CLPT, children are asked to recall linguistic information (repeat the last words in an increasing number of sentences) after responding to the veracity of each sentence (e.g., “Pumpkins are purple”). The two dependent variables are accuracy of word recall and yes/no sentence judgments. Stimulus items are high frequency vocabulary words embedded in simple sentences (see Campbell et al., 1997, for review). CLPT has been most closely linked to models of functional working memory because children are required to perform two operations simultaneously—to store increasing numbers of lexical items at the same time they process incoming information. As noted by Montgomery (2002), however, the CLPT task emphasizes storage rather than processing because the yes/no sentence judgments do not increase in difficulty throughout the task.

In NWR, children imitate strings of phonemes or nonsense words. These nonwords adhere to the phonotactic constraints of the test language in that they use conventional speech sounds combined in permissible ways, but have no semantic value. The dependent variable in NWR is the number or percent of phonemes correctly produced. The NWR task is most often associated with theoretical models of phonological working memory (e.g., Gathercole & Baddely, 1990; Gathercole, Willis, Baddeley, & Emslie, 1994), although phonological discrimination skills, articulatory output, and long-term lexical knowledge can also affect task performance (e.g., James, van Steenbrugge, & Chiveralls, 1994; Snowling, Chiat, & Hulme, 1991). NWR is assumed to emphasize the ability to keep phonological material available long enough for higher-level lexical and sentence processing to take place. The specific NWR task we employ is the task developed by Dollaghan and Campbell (1998), which has received most attention in the LI literature. For CLPT as well as NWR, our primary concern here is with the potential value of NWR and CLPT tasks as nonbiased assessment measures for linguistically diverse learners.

CLPT and NWR Performance by Diverse Learners

An increasing literature base suggests that CLPT and NWR may be robust assessment measures in some culturally and linguistically diverse populations. Campbell et al. (1997) compared the performance of “majority” (White) children and “minority” (primarily African American) children on an experience-dependent measure of language, the Oral Language Scales (OLS) from the Woodcock Language Proficiency Battery–Revised (Woodcock, 1991), along with CLPT and NWR. Participants were 156 typically developing 11- to 14-year-old monolingual English-speaking boys. Between-group comparisons revealed significant differences favoring majority children on the OLS, the experience-dependent measure of language. In contrast, there were no differences in performance between these typically developing majority and minority children on the two processing-dependent measures. Dialectal variation between majority and minority groups was not directly assessed; therefore, the viability of performance on CLPT and NWR tasks to separate dialectally or culturally diverse children was inferred from the race-based group assignment.

In a subsequent study, Dollaghan and Campbell (1998) investigated the sensitivity and specificity of NWR performance to LI in 6- to 9-year-old African American and White children. Children were either typically developing or had a diagnosis of LI. The LI group was a clinically referred sample, with eligibility for language intervention used as the sole criterion for LI. All children were monolingual English speakers, and potential dialectal differences were not reported. For present purposes, the most significant finding was that NWR performance was both sensitive and specific to LI, independent of participant race. Researchers found a likelihood ratio of 25.15 and a posttest probability of a little greater than 95 percent that children scoring less than or equal to 70 percent correct on NWR had LI. That is, children with 70 percent or lower accuracy were 25 times more likely to come from the LI group as compared to their typically developing peers, indicating compelling diagnostic value for this task. In contrast to the clear separation between LI and typically developing children on NWR, there was significant overlap between LI and typical performance for African American children on the Test of Language Development–2 (TOLD-2; Hammill & Newcomer, 1988), a standardized experience-based language measure. These results clearly indicated that NWR was a much more sensitive and specific measure of LI in racially, and presumably culturally, diverse children in this age range.

Ellis Weismer et al. (2000) investigated NWR performance in 581 school-age children (mean age = 7 years 11 months). Participants were all monolingual speakers of English, but represented various ethnic/racial and economic backgrounds. Children of color made up approximately 15 percent of the sample. Results indicated a clear advantage for majority children on the standardized experience-dependent language measures, which included the Peabody Picture Vocabulary Test–Revised (PPVT-R; Dunn & Dunn, 1981), the Comprehensive Receptive and Expressive Test (CREVT; Wallace & Hammill, 1994) and portions of the Clinical Evaluation of Language Fundamentals -3 (CELF-3; Semel, Wiig, & Secord, 1995). In contrast, there were no significant differences between majority and minority students in their NWR performance. These results further supported the notion that the NWR is a "culturally nonbiased measure of language processing" (p. 865).

Rodekohr and Haynes (2001) investigated performance on the CLPT, NWR, and TOLD-2P (Hammill & Newcomer, 1988) in 40 7-year-old children. Participants were placed into one of four groups, each with 10 children: typically developing African American children, typically developing White children, African American children with LI, and White children with LI. Importantly, the 20 African American participants in this study were confirmed speakers of African American English (AAE), reflecting dialectal differences from White participants. For all three tasks, children with LI performed more poorly than the typically developing children. Comparisons between the two groups of typically developing children revealed significantly higher scores on the TOLD-2P for the White group. In contrast, there were no statistically reliable differences between typical AAE speakers and their typical White peers on the CLPT or NWR. These results led researchers to suggest that processing-based tasks may be useful tools in reducing assessment bias with diverse learners. It is important to note, however, that there were some observable group differences in mean CLPT and NWR scores. For example, the typical AAE group had an average score of 36 ($SD = 14$) on CLPT word recall compared with 48 ($SD = 11$) for the typical White group. These differences did not reach the conventional level of statistical significance, which may be attributed to the relatively small sample sizes (10 children per group) and the large variability in performance within each group. Task sensitivity and specificity to LI at the individual level was not investigated.

The combined results from these studies clearly indicate that CLPT and NWR hold diagnostic promise for identifying LI among some groups of culturally or linguistically diverse children. What remains to be determined is the degree to which performance on the CLPT or NWR is affected by experience in other languages. For example, will typically developing children who have learned Spanish as a first language and English as a second language (L2) perform as well on the CLPT or NWR as their monolingual peers? Or will performance on these tasks be influenced by previous language experience? There is some evidence to suggest that the latter scenario may be the case. Thorn and Gathercole (1999) compared NWR performance by children learning an L2 and monolingual peers. As is conventional on these tasks, the nonsense words were created to adhere to the phonological rules of the test languages (French and English in this case). Monolingual children performed significantly better than did the typical L2 learners. This advantage for monolingual children over L2 learners indicates that NWR performance is dependent, to some extent, on previous language experience. However, it may be that the poorer performance on the NWR was due to the relatively little experience in the test language for the L2 learners, and that a critical minimum of experience in the L2 had not been attained. In contrast to the relatively novice L2 learners included in the Thorn and Gathercole study, bilingual participants in the current study are proficient speakers of Spanish as well as English, which provides a more rigorous test of the hypothesis that performance on the NWR or CLPT is not influenced by ability or experience in another language.

Gutiérrez-Clellen, Calderón, and Ellis Weismer (2004) developed Spanish versions of the CLPT and the Dual Processing Comprehension Task (DPCT; Ellis Weismer, 1996). The DPCT requires children to process two different sentences, presented at the same time. Thus, this task was assumed to emphasize attention inhibition rather than storage demands. These tasks, along with their English versions, were used to investigate performance by 44 typically developing second-grade Latino children. Language proficiency was determined on the bais of parent and teacher report, along with samples of children's spoken narratives. No standardized tests were administered in either English or Spanish. Experimental tasks were administered to half of the participants in only one language (either Spanish or English) as these children were considered to have limited skills in the other language. The remaining 22 children were considered proficient bilinguals and are therefore most comparable to participants in the current study. This proficient bilingual group was administered Spanish as well as English versions of the CLPT and DPCT. Analysis at the group level revealed no differences in performance between Spanish and English, indicating that these tasks did not exceed the language skills or processing capacity of these children in either language. That is, there was no disadvantage on English CLPT or DPCT, as compared to the Spanish version, for the proficient bilingual group. Gutiérrez-Clellen and colleagues also found no differences between CLPT and DPCT performance for the proficient bilingual group as compared to English- or Spanish-dominant participants. However, in comparing their results with previous studies with monolingual English-speaking children, there were large differences in CLPT word recall between bilingual participants ($M = 48$ in Gutiérrez-Clellen et al., 1999) and monolingual participants ($M = 60$ in Ellis Weismer et al., 1999, as well as in Gaulin & Campbell, 2004). In discussing these cross-study differences, Gutiérrez-Clellen and colleagues noted that studies that provided external measures of language proficiency for bilingual participants were needed to clarify the role of previous linguistic experience on processing-dependent measures.

In this study, we investigate performance on English CLPT and NWR tasks by English-speaking children with LI and two groups of typically developing peers with diverse language experience, proficient bilingual (Spanish–English) children (BI), and monolingual English-only children (EO). Children with LI were identified using conventional criteria. Independent measures of language proficiency for bilingual participants in the current study included performance on English and Spanish versions of the CELF-3 (Semel et al., 1995; Semel, Wiig, & Secord, 1997). The primary study goal is to determine if performance on these tasks separates typical learners from children with LI, independent of the number of languages spoken, at both group and individual levels. There are two competing hypotheses. If performance on the CLPT or NWR relies more heavily on the integrity of the underlying information-processing system, and not on previous experience in a given language, then we would expect performance by the BI and EO groups to be comparable and significantly better than that by the LI group. The alternate hypothesis is that performance on these tasks is influenced by experience in another language. In this case, performance on the two experimental tasks will separate the typical EO group from both the BI and LI groups. Given the different processing demands between CLPT and NWR, it is also possible that there will be task-related differences in performance by these three groups of children.

METHOD

Participants

Participants were 100 children, aged 7 years 10 months to 13 years 11 months recruited through advertising in local newspapers and elementary schools. Each child was assigned to one of three groups: 28 monolingual English-speaking children with primary language impairment (LI group, 16 male, 12 female, *M* age = 10 years 6 months); 50 monolingual English-only-speaking children with typical language (EO group, 22 male, 28 female, *M* age = 10 years 7 months), and 22 bilingual Spanish–English-speaking children with typical language (BI group, 9 male, 13 female, *M* age = 9 years 9 months). All children in the EO group were Caucasian. The majority of children in the LI group were Caucasian, with the remainder being Native American, African American, or Latino. All children in the BI group were Hispanic/Latino. Families' income level or socioeconomic status (SES) was determined using a 5-point scale in which 1 represents higher SES (business/professional workers) and 5 represents lower SES (unskilled workers) (Hollingshead, 1975). Mean group ratings and standard deviations were 2.39 ($SD = 0.83$) for LI, 2.00 ($SD = 0.97$) for EO, and 3.27 ($SD = 1.08$) for the BI group. This study was part of a larger experimental protocol and complete details about participants are available in Windsor and Kohnert (2004).

All children passed hearing, vision, and oral-motor screening, and had no history of neurological or social–behavioral concerns. Children in the LI group had Receptive or Expressive standard scores (or both) on the Clinical Evaluation of Language Fundamentals–3 (CELF; Semel et al., 1995) that were greater than 1 *SD* below the mean, with most children scoring more than 1.25 *SD* below the mean on at least one section of the CELF ($M = 80$, $SD = 9$). Children with LI scored no more than 1 *SD* below the mean on the Test of Nonverbal Intelligence–3 (TONI; Brown, Sherbenou, & Johnsen, 1997, $M = 102$, $SD = 12$). Many of the children were reported by parents to have a diagnosed learning or reading disability, and 17 of the 28 children had standard scores that were greater than 1 *SD* below the mean on either or both of the Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests–Revised (Woodcock, 1987). All but two children in the LI group had a history of speech–language or reading intervention (or both) services. Children in the EO group had no history of special education services (other than articulation treatment for isolated errors), and all children scored no lower than 1 *SD* below the mean on the Listening and Speaking sections of the CELF ($M = 114$, $SD = 13$) and on the TONI ($M = 110$, $SD = 12$). As is often typical in the LI literature (Swisher, Plante, & Lowell, 1994), the mean TONI score of the LI group fell below that of the two typical groups, $F(2, 99) = 6.14$, $p < .01$, $f = 0.4$.

Although children in the LI and EO groups learned English from birth and had no other languages spoken in the home, the BI children learned Spanish as a first or primary language in the home. The children had 4 to 8 years of experience in learning English, and attended English-only educational programs. Children in the BI group were selected conservatively, to ensure that the children were typical language learners. Each child scored no lower than 1 *SD* below the mean on both English ($M = 102$, $SD = 14$) and Spanish (Semel et al., 1997) versions of the CELF ($M = 116$, $SD = 12$). Children were excluded from the BI group if they were receiving speech–language services or if parents expressed concern about their Spanish language development. Each BI child also scored no more than 1 *SD* below the mean on the TONI ($M = 114$, $SD = 13$). Each child in all three groups showed age-appropriate performance on the Goldman–Fristoe Test of Articulation–2 (Goldman & Fristoe, 2000) when dialectal variations consistent with Spanish-influenced English were considered.

Experimental Tasks and Procedures

Competing Language Processing Task

The stimuli and procedures used for the CLPT were taken from Gaulin and Campbell (1994) and Ellis Weismer et al. (1999). Children were instructed to recall the sentence-final word in a series of sentences after judging the truth value of each sentence. The CLPT has six levels, with two groups of sentences at each level. The number of sentences in a group ranges from 1 to 6 across levels, for a total of 42 test sentences. Each sentence is three words long and contains vocabulary designed to be easy for young school-age children to understand. For example, at the second level, children hear "Pumpkins are purple" and then say yes or no about the truth of the sentence. Children then hear "Buses have wheels" and again make a yes/no judgment. Children are then asked to recall the two sentence-final words (purple, wheels). The yes/no judgments are used to ensure that the task requires online processing as well as recall. All children were administered all items, regardless of performance. That is, there was no discontinuation performance criterion.

Non-Word Repetition

The stimuli and procedures employed for the NWR task were those developed by Dollaghan and Campbell (1998). Children were given one opportunity to imitate 16 nonwords varying in length from one syllable to four syllables (e.g., "naib," "chovag," "chinotaub," "veitachaidoyp"). The 16 nonwords provided a total of 96 target phonemes to be imitated; 12 in one-syllable, 20 in two-syllable, 28 in three-syllable, and 36 in four-syllable nonwords. All nonwords included sounds and sound combinations common in English. Syllables were consonant–vowel–consonant or consonant–vowel combinations. Nonwords were presented from shortest to longest and children were instructed to repeat each nonword immediately after it was presented. As with CLPT, all children were administered all NWR items.

Stimulus items for both tasks were spoken by a native English-speaking female adult. Practice items, test items, and instructions were recorded on compact discs and then presented under headphones to children via computer. If a child appeared to not understand the instructions, the examiner explained these again in live voice. The children were tested individually in a quiet room by a trained research assistant. All responses were audio-recorded for later transcription and scoring by a trained research assistant.

Scoring, Reliability, and Statistical Analyses

The responses to CLPT were scored as percentage correct for Comprehension (yes/no answers) and Recall (word recall). The order of word recall did not have to match the sequence of sentence presentation. The responses of 10 children (10 percent, one third from each participant group) were randomly selected for reanalysis from the audiotapes. There was 100 percent agreement between two independent judges across the 420 scoring opportunities for both Comprehension and Recall scores. NWR responses were scored as percentage of phonemes correctly produced at each syllable length. Each phoneme (consonant or vowel) was scored as produced correctly or incorrectly in relation to the target sound. Following Dolloghan and Campbell (1998) and Ellis Weismer et al. (2000), phonemic substitutions and omissions were scored as incorrect, and sound distortions or additions were not scored as errors. Ten percent of responses from each participant group were randomly selected for reliability checking. Percentage agreement ranged from 94 percent (90/96) to 99 percent (95/96) across children, with an average of 96 percent. The results for NWR for one child from the BI group could not be scored because of poor audiotape quality.

For each task, there were two analyses conducted. The first analysis was a between-group comparison using a one-way analysis of variance (ANOVA). For the CLPT, age was included as a covariate. The children's ages spanned 5 years, and a preliminary regression analysis showed that age in months accounted for 17 percent of the total variance ($p < .01$) in children's mean CLPT Recall (cf. Gaulin & Campbell, 1994). In contrast, for the NWR, age did not account for significant variation in performance ($p = .2$) and, therefore, was not a factor in the overall ANOVA. Post hoc pairwise comparisons using a Bonferroni correction were used to investigate significant main effects for each task. Effect sizes are reported for statistically significant results, following Cohen (1988).

The second analysis was the calculation of likelihood ratios to explore the value of CLPT and NWR scores to identify individual children with LI. Likelihood ratios were used to identify the odds that a particular task score would be expected in a child with LI, versus the probability of a child belonging to either group of typical children. Thus, likelihood ratios provided information about task sensitivity (probability that a child with LI would be identified correctly) and specificity (probability that a typically developing child would be identified correctly). This type of individual analysis presents an important complement to the overall group analysis. Likelihood ratios were obtained following the procedures of Sackette, Haynes, Guyatt, and Tugwell (1991). For any given cutoff score on a task, a likelihood ratio of 20 or more is considered high for ruling an individual into the LI group, with this ratio corresponding to a posttest probability of 95 percent or greater that LI is present. Similarly, likelihood ratios of 0.8 or lower for ruling out a typical child from the LI group correspond to a posttest probability of less than 4 percent. This likelihood ratio is considered sufficient to rule out the presence of LI with high confidence. Likelihood ratios between these two levels are considered to have lower diagnostic value.

For each task, we calculated likelihood ratios in two waves. First, we determined the likelihood ratios to identify LI among monolingual English-only-speaking participants. Results from this analysis would provide an index of task sensitivity and specificity for monolingual speakers of English and would thus be directly comparable with results from previous studies. We then added typical bilingual children into the equation and calculated likelihood ratios for each task considering all participants. This allowed us to determine if sensitivity and specificity levels were altered by the inclusion of children with more diverse language experiences, thereby extending or qualifying comparisons previously reported in the literature.

RESULTS

Group Comparisons

The means, standard deviations, and range for each group on the CLPT are shown in Table 1. As expected, the LI group's mean CLPT Comprehension score was comparable to that of typical peers. Of greater interest was Recall accuracy, as this score has been shown to be most sensitive to differences between children with LI and typical peers (Ellis Weismer et al., 1999). There was a main effect of group for Recall, $F(2, 96) = 10.46, p < .001, f = 0.5$. Pairwise comparisons

TABLE 1
Percent Accuracy Scores for CLPT and NWR

	LI	*BI*	*EO*
CLPT Comprehension	97.9	98.3	99.3
(42 items)	(4.0)	(2.7)	(1.1)
CLPT Recall (42 items)	81–100	93–100	98–100
	56.5	63.2	73.0
	(16.1)	(20.3)	(15.0)
NWR (96 phonemes)	23–88	15–93	30–100
	80.7	86.7	91.9
	(6.6)	(8.1)	(5.2)
	68–94	70–96	75–100

Note. The means, standard deviations, and ranges are provided separately for each of the three groups: LI = (English-only) group with primary language impairment; BI = typical bilingual group; EO = typical English-only group.

showed significantly better performance for the EO group than the LI group, $F(1, 75) = 20.46$, $p < .001$, with a large effect size, $d = 1.07$. Differences between EO and BI groups did not reach statistical significance, $F(1, 69) = 2.03, p = .16$. The difference between LI and BI scores approached, but did not reach, the conventional level for statistical significance, $F(1,47) = 3.27, p = .08$.

Table 1 also shows overall mean group performance, standard deviations, and accuracy range for NWR. Group means for the percent phonemes correct at each of the four syllable lengths are shown in Figure 1. The ANOVA revealed robust differences among groups, $F(2, 96) = 28.39, p < .001$, $f = 0.7$. The performance of the EO group was significantly greater than that of LI, $F(1,76) = 67.98$, $p < .001$, $d = 1.9$, and BI groups, $F(1,69) = 7.71$, $p = .007$, $d = 0.84$. Effect sizes for between-group differences were large, indicating robust practical as well as statistical differences. The performance of the BI group was significantly greater than that of the LI group, $F(1, 47) = 8.28$, $p = .006$, $d = 0.78$. As shown in Figure 1, there was significant overlap between the three groups for one-, two-, and three-syllable nonwords. In contrast, performance on four-syllable nonwords appeared to clearly separate groups.

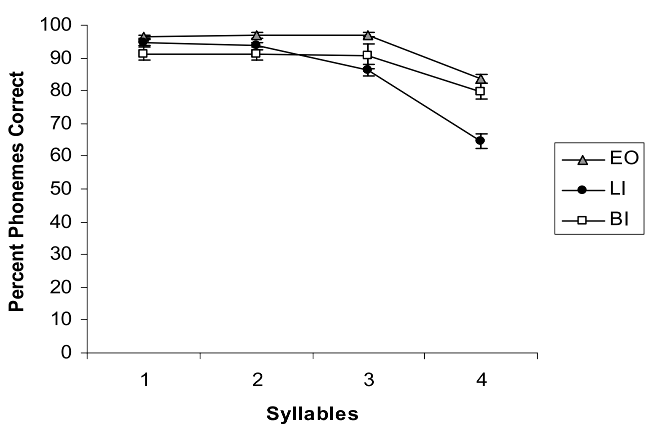

FIGURE 1 Mean percent phonemes correct across syllable lengths for each group. Error bars indicate standard error. LI = (English-only) group with language impairment; BI = typical bilingual group; EO = typical English-only group.

Likelihood Ratios

Likelihood ratios were first calculated for CLPT Recall, drawing on only the monolingual LI and EO groups ($n = 78$). A cutoff raw task score of 12 or lower (out of 42 items) was identified as the point at which the task had maximum sensitivity. At this cutoff score, the likelihood ratio was 3.57, that is, children with a score of 12 or lower were about 3 times more likely to come from the LI group than from the EO group. This ratio is associated with a posttest probability of 67 percent, well below the 95 percent considered clinically informative. A cutoff score of 35 or higher was the point at which the CLPT had highest specificity. At this cutoff, a likelihood ratio of 0.15 was obtained. That is, children with these scores were about 1/6 times more likely to come from a child in the LI group than the EO group. The posttest probability of 7 percent again represented an intermediate probability, with limited diagnostic power.

Likelihood ratios were next calculated drawing on all three groups of children: EO, LI, and BI ($N = 100$). Table 2 shows likelihood ratios for four cutoff levels (scores $\leq$11, 12–22, 23–34, and $\geq$35), which represented the points of maximum diagnostic change. Children with scores $\leq$11 were 2.54 times more likely to come from a child in the LI group than from the typical group, corresponding to an intermediate posttest probability of 49 percent. A score $\geq$35 was 0.20 times as likely to come from a child in the LI group as from the typical group. This corresponds to an intermediate posttest probability of 7 percent. Thus, both sets of analyses indicated that CLPT Recall performance was not highly informative for separating individual children with and without LI in this sample.

The same two sets of likelihood ratios were calculated for NWR, using percentage of phonemes produced correctly. For the monolingual groups, the cutoff score $\leq$76 percent was 10.7 times more likely to come from a child in the LI group than from a child in the EO group. The associated posttest probability of 85 percent was considered to have "intermediate high" diagnostic power. The likelihood ratio for a score $\geq$93 percent was 0.08, with a posttest of 4 percent. Thus, a monolingual English-speaking child who obtained a score of 93 percent or more could be ruled out of the LI group with a high degree of confidence. As shown in Table 3, including the BI group lowered task sensitivity, but not specificity. As for CLPT, four cutoff levels were of interest. Children with a score $\leq$72 percent were 5.07 times more likely to come from the LI group than the typical group, a posttest probability of 66 percent. As in the sample of monolingual children, children with a score $\geq$93 percent were again 0.08 times (about 1/12 times) as likely to come from the LI group as from the typical group. Thus, NWR appeared to be informative for ruling out the presence of LI in this sample of linguistically diverse children but not for ruling in, or identifying, LI.

TABLE 2
Likelihood Ratios for CLPT Recall

	Children With LI (n = 28)		*Children Without LI (EO and BI) (n = 72)*		
Raw Score	*Number*	*Proportion*	*Number*	*Proportion*	*Likelihood Ratio*
≤11	2	.0714	2	.0281	.0714/.0281 = 2.54
12–22	14	.5000	10	.1408	.5000/.1408 = 3.55
23–34	11	.3928	46	.6478	.3928/.6478 = 0.61
≥35	1	.0357	14	.1830	.0357/.1830 = 0.20

Note. Likelihood ratios were calculated for the CLPT Recall performance for the three groups of children: EO = typical English-only group, LI = English-only with language impairment, and BI = typical bilingual group. The likelihood ratios for four cutoff levels representing the points of maximum diagnostic change are shown.

DISCUSSION

The primary goal of this study was to determine if performance on two prominent language-based processing tasks, CLPT and NWR, was effective in separating typically developing Spanish–English (BI) and English-only (EO) speakers from their English-speaking peers with LI. Our LI participants were selected according to conventional criteria (delays in language despite cognitive, motor, and sensory skills within the normal range). Furthermore, over half of the children in the LI group showed below-average reading performance. Our BI children were proficient in English as well as Spanish, as indicated by standardized test performance within the average range for monolingual children, as well as performance at or above grade level in English educational programs. The hypothesis was that if English CLPT and NWR tested the integrity of the underlying information processing system, yet remained impervious to the influence of proficiency in another language, performance by typical monolingual and bilingual learners would be comparable, and significantly greater than that of LI. Because of the different processing demands imposed by the CLPT and NWR, it was also possible that results would not be uniform across these two measures. In the following paragraphs we discuss group, individual, and task-related differences in CLPT and NWR performance.

Consistent with previous research (e.g., Dollaghan & Campbell, 1998; Ellis Weismer et al., 1999), accuracy on both CLPT Recall and NWR tasks statistically separated monolingual EO and LI groups. This replication of previous work for monolingual English speakers with and without LI validates our methodology and reinforces previous study findings. The point of departure from previous research is in the performance of our BI group. Specifically, the mean score for the BI group was somewhere in between scores for EO and LI groups on each task. For CLPT, differences in group means did not clearly separate BI performance from either EO or LI. For NWR, accuracy for the BI group was significantly greater than that of the LI group, and significantly less than their EO peers. It is to be recalled, however, that other studies found that performance on these same CLPT and NWR tasks was effective in separating normal variation from disorders among African American children (e.g., Dollaghan & Campbell, 1998; Rodekohr & Haynes, 2001). At the very least, these combined study results show that the common designation of both African American and bilingual children under the broad rubric of "culturally or linguistically diverse" (CLD) is an oversimplification of the skills and experiences that must be considered in language assessments in general, and on language-based processing measures in particular.

Although group comparisons are important starting points, at the practical level assessment and intervention decisions are always made for individual children. It is, therefore, important to complement group comparisons with likelihood ratios that investigate the diagnostic sensitivity and specificity of these tasks for individual children. Results from CLPT likelihood ratios revealed significant overlap between LI and BI groups. At the most sensitive cutoff point, children

TABLE 3
Likelihood Ratios for NWR

	Children With LI (n = 28)		*Children Without LI (EO & BI) (n = 71)*		
Percent Correct	*Number*	*Proportion*	*Number*	*Proportion*	*Likelihood Ratio*
≤72	2	.0714	1	.0140	.0714/.0140 = 5.07
73–81	14	.5000	6	.0845	.5000/.0845 = 5.92
82–92	11	.3928	34	.4788	.3928/.4788 = 0.82
≥93	1	.0357	30	.4225	.0357/.4225 = 0.08

Note. Likelihood ratios were calculated for the NWR performance for the three groups of children: EO = typical English-only group; LI = English-only with language impairment; and BI = typical bilingual group. The likelihood ratios for four cutoff levels representing the points of maximum diagnostic change are shown.

with CLPT scores ≤11 (out of 42) were almost 3 times more likely to come from a child in the LI group than from the typical groups. Conversely, children with scores ≥35 were 1/5 times as likely to come from children in the LI group as from the typical groups. However, these trends are of insufficient diagnostic power to use the CLPT to clearly identify individual children at risk for LI in linguistically diverse populations. Given that the BI group included only children who had both several years of exposure to English and strong performance on standardized English and Spanish language tests, the practical utility of the CLPT appears particularly limited for differential diagnosis. This finding is consistent with the cross-study comparison between bilingual and monolingual children reported by Gutiérrez-Clellen et al. (2004).

The NWR appears to be more informative, although still limited in its diagnostic utility. Despite broad variability in accuracy within each of the three groups, on average the BI group significantly outperformed the LI group on this task (a mean of 87 percent vs. 81 percent accuracy). Although the BI group was, in turn, outperformed by the EO group (a mean of 92 percent accuracy), it may be that a robust cutoff score exists above which it is highly probable that the score comes from a child who does not have LI. In our study, children with scores ≥93 percent were only 1/12 times as likely to come from the LI group as from either of the two typical groups. As shown in Table 3, only 1 of the 22 children with LI had a score this high, yet more than half of the typical children achieved this score or higher. Given the caveat that our BI group was proficient in English, this finding suggests that NWR may have appropriately high specificity to be of diagnostic value to rule out LI in linguistically diverse learners. On the other hand, NWR may not be useful to identify or rule in LI among linguistically diverse children. Although children with scores ≤72 percent were about 5 times more likely to come from the LI group than the typical groups, this result is not sufficient for NWR to be regarded as a highly sensitive measure.

These results indicate that performance on versions of the CLPT and NWR used in the current study are not independent of previous language experience. It seems that relatively poor performance on these language-based processing tasks may occur for different reasons. English-only-speaking children with LI may perform more poorly than typical English-only age peers on these tasks due to subtle inefficiencies in their basic cognitive–linguistic processing systems. These inefficiencies may be characterized as specific weaknesses in phonological memory, discrimination, or representation (e.g., Edwards & Lahey, 1998; Gathercole & Baddely, 1990; Gathercole et al., 1994; MacDonald & Christiansen, 2002; Montgomery, 2002) or a more general slowing in the processing of different information types (e.g., Hayiou-Thomas, Bishop, & Plunkett, 2004; Kail, 1994; Windsor, 2002). The finding of relatively greater sensitivity to LI for the NWR over CLPT may indicate that phonological processing is a more basic construct in the cognitive–linguistic architecture than lexical or sentence processing (cf. MacDonald & Christiansen, 2002). It also may indicate that the central deficit in LI is better captured by measures that emphasize processing as in NWR versus measures that emphasize mainly storage as in CLPT (cf. Montgomery, 2002). Typical bilingual children may perform more poorly on these tasks because their performance is affected by differences in the long-term knowledge that is accumulated from diverse language-learning experiences. That is, despite their apparent sensitivity to LI, CLPT and NWR appear biased against children who have diverse language-learning experiences. These findings are particularly striking, given that the bilingual children included in this study were sophisticated in English as well as Spanish. It is likely that bilingual children with less experience in the test language would be at a further disadvantage on these language-based processing measures.

Results of cross-linguistic influence on language processing performance for developing bilinguals is consistent with findings from the adult literature, which demonstrates different patterns of sentence processing for bilinguals in each language, as compared to monolingual speakers. Differences in performance between bilingual and monolingual speakers have been discussed in terms of an amalgamation of cues and strategies from two different languages (Hernandez, Bates, & Avila, 1994) or language convergence (e.g., Montrul, 2004). In the current study, the dependent variable for both the CLPT and NWR was accuracy. A more fine-grained qualitative analysis that included the types of errors children made on the NWR, for example, might provide direct evidence of Spanish-influenced performance for the bilingual children. It is also possible that contrastive analysis of NWR errors made by EO and LI groups may reveal differences beyond the quantitative between-group differences reported here and in previous studies.

Given these results, as well as previous findings by Windsor and Kohnert (2004) showing overlap between typical bilingual and monolingual LI performance on lexical processing tasks, we might ask: Are all language-based processing tasks inherently biased against children who have diverse experiences in language? This question requires additional systematic attention, across different groups of language learners as well as across task type and difficulty level. For example, it may be the case that integrity of the underlying language system weighs more heavily than experience or proficiency in the test language on tasks that involve more complex novel language stimuli. In the current study, there was substantial overlap in the NWR task among all three groups on the shorter nonwords. The greatest degree of separation between LI and BI groups was on the repetition of four-syllable nonwords. At the same time, the performance of the BI and EO groups on these stimuli was essentially equal (see Figure 1). Although the small number of four-syllable items in this task condition provides little range to investigate sensitivity and specificity, it may be that an NWR task with longer stimuli is a more robust diagnostic tool. On the other hand, the tremendous variability in language experiences of children learning two languages cannot be overlooked. As noted above, our bilingual participants were sophisticated in English as well as Spanish. Although this conservative inclusionary criterion provided a rigorous test of our study question, it is unlikely that other bilingual children with less experience or skill in the test language would reach the same level of performance. That is, other typically developing bilingual children with less experience in English may have performed as poorly on the repetition of four-syllable nonsense words following the phonotactic constraints of English as English-only children with LI.

It will be important for future research to investigate performance on NWR tasks in both languages of developing bilinguals who vary in their experience and skill in each language. Such a within-subjects design would allow researchers to more precisely investigate the relative roles of general processing efficiency and language-specific proficiency. Given the current results for four-syllable words, such comparisons between bilinguals who speak different languages would also be interesting. Spanish word length in syllables is, on average, considerably longer than English word length. In contrast, Hmong is a predominantly monosyllabic language. Comparisons on English NWR tasks between Hmong–English bilingual speakers as compared to bilingual Spanish–English speakers would further help to identify the nature of cross-linguistic influence on language processing performance.

It also may be that we can take advantage of the subtle cognitive deficits accompanying LI to develop nonlinguistic processing measures that will reduce or eliminate assessment bias. For example, Kohnert and Windsor (2004) found that the same cohort of children as in this study could be separated in a predictable way at the group level on a two-choice visual detection task, that is, a task that emphasized perceptual–motor response time. Both EO and BI groups were able to very quickly identify target colors presented on a computer screen, and these two groups were equivalently faster than the LI group who performed the task significantly more slowly. Importantly, Kohnert and Windsor found that other nonlinguistic tasks (e.g., a parallel auditory detection task) did not have the same potential diagnostic power. These findings clearly show that all processing-dependent tasks should not be expected to be equally sensitive or specific to the presence of LI among diverse learners. Future investigations are needed to develop and refine measures that will be most effective in identifying children who have true difficulties in acquiring or using language as distinct from those children who are in the normal, prolonged process of acquiring a second language (Kohnert, 2002; Kohnert & Bates, 2002). Important considerations in this research will be the dependent variables used to index processing efficiency, the nature of the stimuli to be processed, as well as the conditions under which this processing is measured.

In summary, a critical finding here is that proficiency in the test language and integrity of the general language processing system are necessary, but insufficient to explain performance on language-based processing measures. Performance on both the CLPT and NWR by typically developing Spanish–English children fell somewhere in between that of monolingual English-speaking children with and without LI. Results of likelihood ratios indicated that neither the CLPT nor NWR held compelling diagnostic potential for identifying LI among linguistically diverse learners at the individual level. These results suggest that performance on language-based processing measures are not independent from previous language experience and cannot, without further refinement, be considered nonbiased assessment measures when used with bilingual children. One important exception is that performance on NWR did provide compelling evidence to rule out (as opposed to rule in) LI for bilingual as well as monolingual participants. Further research is needed to develop and refine measures that are both sensitive and specific to LI in linguistically diverse learners.

ACKNOWLEDGMENTS

Funding for this project was provided by two grants from the National Institute of Deafness and Other Communication Disorders, awarded to Jennifer Windsor (R01DC04437) and Kathryn Kohnert (R03DC05542). Portions of this study were presented at the 2003 Symposium on Research in Child Language Disorders and at the 2004 American Speech-Language-Hearing Association Convention. We gratefully acknowledge the contributions of the following individuals who assisted with task creation, data collection, and response scoring: Meredith Bailey-Orr, Krista Bowe, Tara Cassidy, Kerry Danahy, Mark DeRuiter, Caren Dorman, Pui Fong Kan, Ruth Miller, and Sara Turman. We thank Elaine Silliman, Cheryl Scott, and an anonymous reviewer for their very helpful comments on a previous draft of this study.

REFERENCES

Bishop, D., & Snowling, M. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130,* 858–886.

Brown, L., Sherbenou, R., & Johnsen, S. (1997). *Test of Nonverbal Intelligence-3*. Austin, TX: PRO-ED.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing dependent measures. *Journal of Speech, Language, and Hearing Research, 40,* 519–525.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 46,* 1029–1037.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Dollaghan, C., & Campbell, T. (1998). Non-word repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Dunn, L., & Dunn, L. (1981). *Peabody picture vocabulary test-revised (PPVT-R)*. Circle Pines, MN: American Guidance Service.

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19,* 279–309.

Ellis Weismer, S. (1996). Capacity limitations in working memory: The impact on lexical and morphological learning by children with language impairment. *Topics in Language Disorders, 17,* 33–44.

Ellis Weismer, S., Evans, J., & Hesketh, L. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1249–1260.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buchwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43,* 865–878.

Gathercole, S., & Baddeley, A. (1990). Phonological memory deficits in language impaired children: Is there a causal connection? *Journal of Memory and Language, 29,* 336–360.

Gathercole, S., Willis, C., Baddeley, A., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2,* 103–127.

Gaulin, C., & Campbell, T. (1994). Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual and Motor Skills, 79,* 55–64.

Goldman, R., & Fristoe, M. (2000). *Goldman-fristoe test of articulation-2*. Circle Pines, MN: American Guidance Service.

Gutiérrez-Clellen, V., Calderón, J., & Ellis Weismer, S. (2004). Verbal working memory in bilingual children. *Journal of Speech, Language, and Hearing Research, 47,* 863–876.

Hammill, D., & Newcomer, P. (1988). *Test of Language Development 2-Intermediate*. Austin, TX: PRO-ED.

Hayiou-Thomas, M. E., Bishop, D. V. M., & Plunkett, K. (2004). Simulating SLI: General cognitive processing stressors can produce a specific linguistic profile. *Journal of Speech, Language, and Hearing Research, 47,* 1347–1362.

Hernandez, A. E., Bates, E., & Avila, L. (1994). Sentence interpretation in Spanish-English bilinguals: What does it mean to be in-between? *Applied Psycholinguistics, 15,* 417–466.

Hollingshead, A. (1975). *Four factor index of social status*. New Haven, CT: Department of Sociology, Yale University.

James, D., van Steenbrugge, W., & Chiveralls, K. (1994). Underlying deficits in language-disordered children with central auditory processing difficulties. *Applied Psycholinguistics, 15,* 311–328.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 37,* 418–421.

Kohnert, K. (2002). Picture naming in early sequential bilinguals: A 1-year follow-up. *Journal of Speech, Language, and Hearing Research, 45,* 759–771.

Kohnert, K. (2004). Children learning a second language: Processing skills in early sequential bilinguals. In B. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (pp. 53–76). Baltimore: Brookes.

Kohnert, K., & Bates, E. (2002). Balancing bilinguals II: Lexical comprehension and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research, 45,* 347–359.

Kohnert, K., & Windsor, J. (2004). The search for common ground—Part II: Nonlinguistic performance by linguistically diverse learners. *Journal of Speech, Language, and Hearing Research, 47,* 891–903.

Kohnert, K., Windsor, J., & Miller, R. (2004). Crossing borders: Recognition of Spanish words by English-speaking children with LI. *Applied Psycholinguistics, 25,* 543–564.

Laing, S., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools, 34,* 44–55.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109,* 35–54.

Miller, C., Kail, R., Leonard, L., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44,* 416–433.

Montgomery, J. W. (2002). Understanding the language difficulties of children with specific language impairments: Does verbal working memory matter? *American Journal of Speech-Language Pathology, 11,* 77–91.

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition, 7,* 125–142.

Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech, Language, and Hearing Research, 43,* 834–847.

Rodekohr, R., & Haynes, W. (2001). Differentiating dialect from disorder: A comparison of two processing tasks and a standardized test. *Journal of Communication Disorders, 34,* 255–272.

Sackett, D., Haynes, R., Guyatt, G., & Tugwell, P. (1991). *Clinical epidemiology: A basic science for clinical medicine* (2nd ed.). Boston: Little, Brown and Company.

Scarborough, H. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia, 48,* 115–136.

Semel, E., Wiig, E., & Secord, W. (1997). *Clinical evaluation of language fundamentals-3: Spanish edition*. San Antonio, TX: The Psychological Corporation.

Semel, E., Wiig, E., & Secord, W. (1995). *Clinical evaluation of language fundamentals-3*. San Antonio, TX: The Psychological Corporation.

Silliman, E. R., Wilkinson, L. C., & Brea-Spahn, M. R. (2004). Policy and practices imperatives for language and literacy learning. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy* (pp. 97–129). New York: Guildford.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12,* 369–373.

Swanson, H. L., Sáez, L., Gerber, M., & Leafstedt, J. (2004). Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *Journal of Educational Psychology, 96,* 3–18.

Swisher, L., Plante, E., & Lowell, S. (1994). Nonlinguistic deficits of children with language disorders complicate the interpretation of their nonverbal IQ scores. *Language, Speech, and Hearing Services in the Schools, 25,* 235–240.

Thorn, A., & Gathercole, S. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology, 52,* 303–324.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40,* 1245–1260.

Tomblin, J. B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioral disorders, and language impairment among second grade children. *Journal of Child Psychology and Psychiatry, 41,* 473–482.

Vellutino, F., Fletcher, F., Snowling, M., & Scanlon, D. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2–40.

Wallace, G., & Hammill, D. (1994). *Comprehensive Receptive and Expressive Vocabulary Test (CREVT)*. Austin, TX: PRO-ED.

Windsor, J. (2002). Contrasting general and process-specific slowing in language impairment. *Topics in Language Disorders, 22,* 49–61.

Windsor, J., & Hwang, M. (1999). Children's auditory lexical decisions: A limited processing capacity account of language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 990–1002.

Windsor, J., & Kohnert, K. (2004). The search for common ground—Part I: Lexical performance by linguistically diverse learners. *Journal of Speech, Language, and Hearing Research, 47,* 877–890.

Woodcock, R. (1987). *Woodcock Reading Mastery Tests—Revised*. Circle Pines, MN: American Guidance Service.

Woodcock, R. (1991). *Woodcock Language Proficiency Battery—Revised*. Allen, TX: DLM Teaching Resources.

About the Authors

Kathryn Kohnert, Ph.D., CCC-SLP is Associate Professor in the Department of Speech-Language-Hearing Sciences, affiliate member of the graduate faculty in English as a Second Language, and Codirector of the *Center for Cognitive and Social Processes in Language* at the University of Minnesota. She was awarded a McKnight Professorship in 2003 for her research program, which investigates typical language learning and language impairments in diverse populations from a general cognitive processing perspective.

Jennifer Windsor, Ph.D., is Professor and Chair in the Department of Speech-Language-Hearing Sciences and Codirector of the *Center for Cognitive and Social Processes in Language* at the University of Minnesota. Her current research interests are in the cognitive processes that contribute to children's language development and language disorders; and in the relation between spoken and written language. Her primary focus is on children with primary language impairments and learning disabilities.

Dongsun Yim, M.A., is an advanced doctoral student in the Department of Speech-Language-Hearing Sciences at the University of Minnesota. Her main research interests focus on language and cognitive processes in children with primary language impairments.