Published in final edited form as:

Am J Speech Lang Pathol. 2010 November ; 19(4): 298–310. doi:10.1044/1058-0360(2010/09-0064).

Cross-Language Nonword Repetition by Bilingual and Monolingual Children

Jennifer Windsor, **Kathryn Kohnert**, **Kelann F. Lobitz**, and **Giang T. Pham**

University of Minnesota, Minneapolis

Abstract

Purpose—Identifying children with primary or specific language impairment (LI) in languages other than English continues to present a diagnostic challenge. This study examined the utility of English and Spanish nonword repetition (NWR) to identify children known to have LI.

Method—Participants were 4 groups of school-age children (N = 187). There were 2 typically developing groups: proficient Spanish-English sequential bilinguals and monolingual English speakers. There were 2 groups of children with LI, one Spanish-English and the other monolingual English speakers. Children participated in both English and Spanish NWR.

Results—Children's NWR performance was significantly correlated across languages. In English NWR, the 2 groups with LI had lower accuracy at the longest syllable length than the 2 typically developing groups. In Spanish NWR, monolingual children with LI had lower repetition accuracy than bilingual children with LI and typical monolingual children, with all 3 groups outperformed by the typical bilingual group. Likelihood ratios indicated adequate diagnostic power only for English NWR in ruling out the typical bilingual children as showing LI.

Conclusion—The results demonstrate that NWR performance relies on the dual influences of LI and native language experience. However, it remains possible that NWR is useful in a composite marker for LI.

Keywords

language disorders; sequential bilingualism; assessment procedures; specific language impairment

Children's nonword repetition (NWR) in their native language has received a great deal of attention as a significant marker of monolingual children's overall language ability. Although the nature and direction of the relation during development have been interpreted in different ways, several studies have shown a link between NWR and vocabulary size for typically developing children (Gathercole, Hitch, Service, & Martin, 1997; Gathercole, Willis, Emslie, & Baddeley, 1992) and for children identified as showing a primary or specific language impairment, conventionally termed SLI (Edwards, Beckman, & Munson, 2004; Munson, Kurtz, & Windsor, 2005). Correlations between NWR and grammatical measures have also been demonstrated (Archibald, Joanisse, & Shepherd, 2008; Girbau &

Contact author: Jennifer Windsor, Department of Speech-Language-Hearing Sciences, 115 Shevlin Hall, University of Minnesota, 164 Pillsbury Drive SE, Minneapolis, MN 55455. windsor@umn.edu.

Schwartz, 2007; Thordardottir, 2008). However, the link with complex sentence comprehension has been less compelling (Montgomery, Magimairaj, & O'Malley, 2008). Because of its consistent association with at least some aspects of language performance, NWR has been proposed as a useful clinical tool to assist in the identification of monolingual children with SLI. The primary purpose of the current study is to examine the value of English and Spanish NWR in assessing the integrity of bilingual children's language-learning system.

Theories of NWR

Originally, NWR was conceived of as a measure of phonological short-term memory capacity (Gathercole & Baddeley, 1990a, 1990b) and as such was thought to implicate phonological short-term storage as a pivotal construct in the language performance of SLI (Montgomery, 1995). Later work has shown that NWR indeed is correlated with other short-term memory measures, such as serial digit recall (Archibald et al., 2008). However, it is also clear that NWR performance reflects factors in addition to phonological short-term memory alone (see Gathercole, 2006, and associated commentaries). Key factors that have been proposed include the integrity of phonological awareness and representation, lexical knowledge, and speech output (Archibald & Gathercole, 2006, 2007; Edwards et al., 2004; Graf-Estes, Evans, & Else-Quest, 2007; Kovács & Racsmány, 2008; Metsala, Stavrinos, & Walley, 2009; Snowling, Chiat, & Hulme, 1991). Gupta and colleagues (Gupta, 2006; Gupta, Lipinski, Abbs, & Lin, 2005) suggested that NWR performance can be framed within a computational model of serial order that includes both lexical and sublexical processing. MacDonald and Christiansen (2002) proposed a connectionist framework in which verbal working memory capacity is not distinct from the representation of other linguistic knowledge. From this perspective, NWR may correlate with measures of language performance not because short-term memory and working memory are distinct constructs underlying language performance but because NWR is another type of language task.

Despite these theoretical differences about the underlying nature of NWR, a robust finding is that children with typical language and children with SLI have lower NWR accuracy on longer nonwords than on shorter nonwords. This finding is consistent across studies that have contrasted different English NWR tasks (Archibald & Gathercole, 2006) and different dialects (Oetting & Cleveland, 2006). This length effect also has been shown in languages other than English (Girbau & Schwartz, 2007). A meta-analysis of 23 studies of English-speaking children with SLI found only small NWR differences to age peers in repeating shorter syllables, with larger group differences for longer syllables (Graf Estes et al., 2007). Increased syllable length may be associated with increased segmental and suprasegmental demands. However, there is general consensus that the length effect does speak to phonological storage capacity in some form as a significant variable in NWR performance, especially when children's phonological systems are not yet developed fully or the non-word stimuli are unfamiliar (Bowey, 2006; Gupta et al., 2005).

The relation between NWR and lexical knowledge that has been identified for typically developing children is not as clear for children with SLI. It also has been argued that what makes NWR relevant to SLI is not phonological storage capacity but other linguistic and

cognitive aspects of NWR tasks (Montgomery & Evans, 2006). However, NWR has been considered a promising clinical marker of SLI for several years (Conti-Ramsden, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000). The way in which NWR is assessed introduces some variability to this finding, with English-speaking children with SLI tending to show particular difficulty with nonwords that are longer or include consonant clusters (Archibald & Gathercole, 2006; Graf Estes et al., 2007). Most research has replicated early findings of significant group differences in NWR accuracy for monolingual children with and without SLI. Notably, Graf Estes et al.'s (2007) meta-analysis found that the SLI groups had NWR accuracy that was an average of 1.3 *SD*s below age-matched control groups. While this is a substantive difference, review of the original studies indicates that the overall language skills of the SLI groups often were defined a priori as being substantially poorer than those of the typical age-matched groups.

Clinical Use of Nonword Repetition

Results from recent studies vary in the extent to which NWR accuracy by itself has sufficiently high sensitivity and specificity at any given cutoff to be a reliable diagnostic or screening tool for SLI. For example, in a study of 160 monolingual 11-year-olds with a history of SLI, Conti-Ramsden, Botting, and Faragher (2001) found high specificity (92%) but only moderate sensitivity (74%) for NWR (using a cutoff point of the 10th percentile of the control group's performance). That is, children who scored above this cutoff point could reliably be identified as typical language learners, but children scoring below this point could not reliably be ruled in as showing SLI. On the other hand, using discriminant function analysis, Gray (2003) found very high sensitivity (95%) and specificity (100%) for 44 monolingual 4–6-year-olds (22 children each with SLI and typical language).

Compounding the issue of sensitivity to diagnostic category is that the exclusionary and broad inclusionary criteria for SLI provide inherently heterogeneous samples (Leonard, 1998). This heterogeneity makes it difficult to generalize across samples and to identify salient group characteristics, such as relatively low NWR accuracy, unless effect sizes are reasonably large. This issue is central to the long-standing debate as to whether SLI is indeed a discrete clinical category, separable from both typical language development (Leonard, 1991) and other types of developmental language impairments (Catts, Adlof, Hogan, & Ellis Weismer, 2005; Hayiou-Thomas, Oliver, & Plomin, 2005).

Finally, lower NWR performance relative to controls has been shown for several other groups of children with speech-language difficulties, including children with phonological disorders, autism, focal lesions, reading impairment, and stuttering (Anderson, Wagovich, & Hall, 2006; Gupta, MacWhinney, Feldman, & Sacco, 2003; Munson, Edwards, & Beckman, 2005; Roodenrys & Stokes, 2001; Smith Gabig, 2008). Relatively low accuracy also has been found for adolescents with residual speech errors (Preston & Edwards, 12007). Thus, NWR may be better viewed as a reliable correlate of monolingual groups' speech-language abilities rather than as a diagnostic marker of any particular clinical group with language difficulties.

Even if it does not have adequate power to meet the gold standard of a diagnostic tool, there are at least three advantages of NWR as one part of a screening protocol or diagnostic battery for children's language difficulties. First, compared to standardized language tests, NWR tasks are quick to administer and relatively easy to score. For example, the Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996) requires children to repeat 40 English nonwords of varying syllable lengths. The Nonword Repetition Task (Dollaghan & Campbell, 1998) involves repetition of only 16 English nonwords of varying lengths. Both tasks are scored by calculating the percentage of phonemes produced correctly.

The second advantage is that NWR has been shown to reduce the cultural bias associated with racial differences, with most comparisons being between African American and Caucasian children (Campbell, Dollaghan, Needleman, & Janosky, 1997; Ellis Weismer et al., 2000; Rodekohr & Haynes, 2001). For instance, Ellis Weismer et al. (2000) showed that school-age children from majority backgrounds received higher scores on a battery of standardized language tests than children from minority backgrounds. However, reflecting the fact that NWR appears to be less dependent on language experience than standardized tests, there was no group difference in NWR performance. NWR also has been shown to be unaffected by race when nonmainstream dialects—southern White and southern African American English dialects— have been examined (Oetting & Cleveland, 2006). Similarly, NWR has been found to be robust in the face of gender differences (Chiat & Roy, 2007; Roy & Chiat, 2004) and, among typically developing children, socioeconomic differences (Chiat & Roy, 2007).

Finally, NWR has been found to be useful with young children, when screening would be particularly valuable to identify children at risk for language-learning difficulties. Roy and Chiat (2004) successfully implemented an NWR task with typical 2- to 4-year-olds, with the task capturing developmental differences in repetition accuracy. Chiat and Roy (2007) demonstrated the task's efficacy with a group of young children referred for speech-language assessment. Stokes and Klee (2009) also showed a robust predictive relation between typically developing 2-year-olds' NWR and their reported vocabulary knowledge. As noted above, Gray (2003) found that preschool-age children with SLI could be identified reliably using an NWR task.

Nonword Repetition in Both Languages by Typical Bilingual Speakers

A disadvantage of NWR tasks for clinical purposes is that NWR does not eliminate the role of linguistic experience or bias, but rather reflects language-specific influences (Masoura & Gathercole, 1999; Thorn & Gathercole, 1999). About one in five children in the United States who are 5 years of age and older speak a language other than English as their native language (U.S. Census Bureau, 2006). This presents a pressing assessment challenge given that most speech-language pathologists in the United States are monolingual English speakers. The research base on bilingual language development also is not yet as well established as it is for monolingual children (Bedore & Peña, 2008; Kohnert & Medina, 2009). Moreover, typical bilingual language acquisition is distributed across two (or more) languages and is associated with even greater variability than is found in monolingual acquisition (Kan & Kohnert, 2005; Peña, Bedore, & Zlatic-Giunta, 2002; for reviews, see

Kohnert, 2008; Kohnert, Windsor, & Ebert, 2009). Overall, identifying which bilingual children truly have language-learning difficulties has become a key issue for many practicing clinicians.

The current study examined the NWR performance of bilingual Spanish-English and monolingual English-speaking children with and without primary language impairment. To date, studies of NWR in bilingual children have focused mainly on typically developing children, and most of the attention has been on links between NWR and vocabulary skill in the second language (L2). There is now a substantial literature on NWR for children who are novice learners of an L2, typically exposed to the L2 in a classroom context. The overall finding for these L2 learners has been that NWR accuracy is higher in the first language (L1) than in the L2. It is also the case that NWR and vocabulary knowledge in the L2 are related, at least for children with smaller vocabularies in their newer and substantially less developed language. For example, Service (1992) and Service and Kohonen (1995) showed that native Finnish-speaking school-age children's NWR in their L2, English, predicted their overall English language outcomes. Cheung (1996) studied Cantonese-speaking children who were learning English as a foreign language at school. In line with Service's (1992) results, the children's English nonword span (repetition of an increasing number of nonwords) predicted the number of trials needed to learn unfamiliar English words. However, the relation held only for children with smaller English vocabularies. This finding was interpreted as consistent with the idea that phonological short-term memory does not contribute to word learning when long-term phonological knowledge has been acquired (i.e., when the nonword stimuli resemble real words or have high "wordlikeness").

Masoura and Gathercole (1999) examined the Greek and English NWR of Greek children learning English in school. NWR accuracy was moderately correlated across languages, with stronger NWR in the L1, Greek. Vocabulary ability also was correlated across languages, even after accounting for the contribution of NWR. The authors interpreted this finding as indicating that phonological short-term memory alone was sufficient to explain vocabulary acquisition. In another study, Masoura and Gathercole (2005) examined the NWR skills of Greek children learning English as an L2, but at a point at which children had at least 3 years of exposure to English and relatively large English vocabularies. NWR performance again was found to be correlated across languages; however, reminiscent of Cheung's (1996) conclusion, NWR was correlated with current English vocabulary knowledge but not with the number of trials needed to learn new words. That is, phonological short-term memory apparently was more important for initial vocabulary learning, while later word learning could be lexically mediated by existing vocabulary skills.

Thorn and Gathercole (1999) studied three groups of children with English as their L1: monolingual English-speaking children, English-French bilinguals exposed to both languages before the age of 3 years, and English-speaking children who began to learn French in school after the age of 3 years. The latter two groups had similar French vocabulary knowledge. The results indicated that the three groups had equivalent English NWR accuracy. The bilingual English-French group who had been exposed to both languages before 3 years of age showed equivalent NWR accuracy across languages. The two English-French groups were comparable in French NWR and outperformed the English-

speaking group in French NWR. Thorn and Gathercole interpreted the similarity between the two English-French groups despite their nonequivalent French experience as evidence for the link between NWR and vocabulary knowledge.

Links between children's NWR and grammatical ability have also been examined. Notably, French and O'Brien (2008) examined the link between French-speaking children's NWR and a written test of grammatical knowledge in their L2, English. Although there was not a strong link with vocabulary knowledge, children's English NWR was associated with later grammatical knowledge. Indicating that the association between NWR and grammar was not language-specific, the same children's NWR in an unfamiliar third language, Arabic, also predicted their English grammatical knowledge.

Nonword Repetition by Bilingual Speakers With SLI

To our knowledge, only four studies have examined the NWR of developing bilinguals with language difficulties (Calderón, 2003; Girbau & Schwartz, 2008; Gutiérrez-Clellen & Simon-Cereijido, 2010; Palladino & Cornoldi, 2004). Calderón (2003) developed a set of 22 nonwords adhering to the phonotactic constraints of Spanish. Stimuli were two, three, and four syllables long. The NWR task was administered to 16 Spanish-English children with SLI (mean age of 5 years) and 16 age-matched typically developing bilingual peers. All participants lived in California and had greater proficiency in Spanish at the time of testing. Results showed that the children with SLI were significantly less accurate in repeating the Spanish nonwords than the typical peers at all three word lengths; however, there was no interaction between group and word length.

Palladino and Cornoldi (2004) studied NWR in Italian-speaking children who showed and did not show difficulties learning English as an L2 in school. The results showed that the group with English language difficulties had poorer NWR performance in their L1, Italian, than the other group of children. Some of the children who had difficulties learning English also were reported to have a history of language difficulties in Italian. Thus, this study also may have included children with SLI.

Girbau and Schwartz (2008) found that Spanish NWR separated 7- to 10-year-old Spanish-speaking children with and without SLI living in New York who were sequential bilinguals learning English as their L2. Participants were 22 children in a dual language school, 11 with SLI and 11 age-matched typically developing peers. The children repeated 20 Spanish nonwords (across five syllable lengths). For both groups, longer words were repeated with lower accuracy, with the SLI group making a greater number of consonant substitution and omission errors than their typical peers. Likelihood ratios based on the percentage of nonwords repeated correctly indicated that the Spanish NWR task was useful as a screening measure for SLI. For this relatively small group, a score below 33% correct repetition of three-, four-, and five-syllable nonsense words was sufficient to rule in a diagnosis of SLI. NWR performance also was positively correlated with subtests in a standardized language test that had an auditory working memory component, including auditory association and grammatical integration. (See Girbau & Schwartz, 2007, for similar results with Spanish-speaking children living in Spain who also understood but rarely used Catalan.)

The studies by Calderón (2003) and Girbau and Schwartz (2008) assessed NWR in a single language, the children's L1. There also were no monolingual comparison groups to further understand the utility of NWR as a screening tool. Kohnert, Windsor, and Yim (2006) examined Spanish-English bilingual children's NWR in their L2, English. There were three groups of school-age participants: typical monolingual English-speaking children, typical bilingual Spanish-English sequential bilingual children who were proficient in both languages, and monolingual English-speaking children with SLI (identified as children with a primary language impairment [LI] by the authors). Kohnert et al. (2006) found that the bilingual children performed with significantly lower NWR accuracy than the typical monolingual English-speaking age peers. This result was in line with Masoura and Gathercole's (1999) finding of an L1 advantage. However, the typical bilingual group did have significantly higher NWR accuracy than the monolingual English-only group with SLI, suggesting some screening utility for the English NWR task with bilingual children. Comparing all three groups using a cutoff score of 93% phonemes correct, there was intermediate-high specificity (85%) but only moderate sensitivity (66%) to SLI. That is, monolingual and bilingual children who scored above this cutoff point could reliably be identified as typical language learners with fairly high accuracy, but children scoring below this point could not reliably be ruled in as showing SLI.

Recently, Gutiérrez-Clellen and Simon-Cereijido (2010) examined the Spanish and English NWR performance of 4- to 7-year-old Spanish-English bilingual children with and without SLI. The typically developing children had higher NWR accuracy in both languages than children with SLI. However, NWR in each language was found to have only moderate specificity and low sensitivity to SLI. Specificity was high when children's NWR in both languages was considered together. Gutiérrez-Clellen and Simon-Cereijido concluded that children's NWR performance could be related to individual differences in language exposure and usage.

The current study extends the work of Kohnert et al. (2006) and Gutiérrez-Clellen and Simon-Cereijido (2010) by investigating monolingual and bilingual children's NWR in both languages. Both children with and without LI participated. Consistent with our previous work, we refer to the children with language impairment as showing a "primary" rather than "specific" language impairment (hereafter termed LI). Our first question was the relative value of the English and Spanish NWR tasks in identifying children at risk for LI at the group and individual levels. The second question was whether NWR performance was correlated across languages for both the typical groups and groups with LI.

Given the extant literature, we anticipated that there would be NWR differences in accuracy between children with and without LI. However, we also expected that the influence of language-specific knowledge would lower the clinical utility of NWR for sequential bilingual children when the NWR task was administered in their L2. Following the research for typical children who were naïve learners of an L2, we expected that the bilingual children's NWR would be correlated across languages. We also anticipated that the English-speaking children's NWR would be correlated across languages, with the Spanish nonwords acting as less familiar, less wordlike stimuli than the English nonwords.

Method

Participant Recruitment

This study was part of a larger project that included typical and atypical learners of English or Spanish as their primary home language. To recruit participants for this study, we worked with personnel in several Minneapolis-St. Paul schools to send study consent forms to the families of students age 6;0 (years;months) to 11;6. This included children in regular education and children receiving special education services. A total of 228 children returned completed consent forms. Children were excluded from the current study if sensory, neurological, or socioemotional concerns were reported by parents or school personnel. Children also were excluded if they had a primary diagnosis other than primary language-learning difficulties (e.g., developmental delay, autism spectrum disorder, or attention deficit disorder) or did not pass hearing screening carried out by study personnel at 25 dB at 1000, 2000, and 4000 Hz. This hearing screening criterion was less conservative than American National Standards Institute standards but not atypical when screening children's hearing in school settings. Bilingual children were excluded if their L1 or home language was not Spanish. Monolingual English-speaking children were excluded if they had been exposed to an L2 through language immersion educational programs or had close family members who spoke languages other than English. Forty-one children were excluded, leaving 187 children. We did not obtain information on socioeconomic background from children. However, almost all children in the schools from which children were recruited qualified for free/reduced lunch. To this extent, the children came from fairly similar socioeconomic backgrounds. There was no a priori reason to expect group differences in parents' income or education.

Participant Groups

The 187 children each participated in one of four groups. There were 69 typically developing children who spoke English as their native and only language (EO group) and 34 native English-speaking children with LI (EO-LI group). The other two groups included 65 typically developing sequential bilinguals who had learned Spanish as their L1 and English as their L2 (BI group), and 19 Spanish-English bilinguals with LI (BI-LI group).

As part of determining group membership, the monolingual children in the EO group were administered the Recalling Sentences and Concepts and Directions subtests of the English version of the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF–4; Semel, Wiig, & Secord, 2003). Children repeat sentences of varying length and grammatical complexity in Recalling Sentences, and they carry out directions of increasing complexity in Concepts and Directions. Children in the monolingual EO-LI group were administered all CELF–4 subtests. Bilingual children in the BI and BI-LI groups were administered the CELF–4 Recalling Sentences and Concepts and Directions subtests in English as well as the parallel Spanish versions of these two subtests (Wiig, Secord, & Semel, 2006). Children in all four groups were administered the Test of Nonverbal Intelligence, Third Edition (TONI–3; Brown, Sherbenou, & Johnsen, 1997). Each group is described below. Table 1 summarizes the groups' language and cognitive test scores.

Monolingual speakers with typical development (EO group)—The 69 children in the EO group had a mean age of 8;7 (*SD* = 1;7). About two thirds of the group (*n* = 46) reported that they were Caucasian, with 12 African American, two Hispanic, one Native American, and eight biracial children. All children in the EO group were reported by parents or school personnel to have age-appropriate language. No child had repeated a school grade or received special educational services for language difficulties. The average of the two English CELF–4 subtest standard scores ranged from 6 to 16 across children (where a subtest score of 10 is average). Each child in the EO group had a nonverbal IQ at or above –1 *SD* on the TONI–3.

Monolingual speakers with primary LI (EO-LI group)—The 34 children in the EO-LI group had a mean age of 8;9 (*SD* = 1;5). There were 11 Caucasian, 11 African American, one Hispanic, and 11 biracial children. Each of the children in the EO-LI group had a history of language-learning difficulties, and there were reported concerns about language abilities. All but three children currently were receiving special educational services for language in their schools. The children in the EO-LI group met conventional criteria for SLI. Children showed a discrepancy between nonverbal IQ and (lower) language performance and had no reported history of sensory, neurological, or socioemotional difficulties. We included two children in the EO-LI group who had a nonverbal IQ below –1.3 *SD*s on the TONI–3 but an IQ score above that associated with a classification of mild mental disability. These children could be considered to show a nonspecific language impairment because both cognitive and language scores were outside the normal range (Tomblin & Zhang, 1999). However, there is increasing awareness that the profiles of children with specific and nonspecific language impairment tend to be the same and, further, that measurement error in test scores can obscure this group distinction (Hayiou-Thomas et al., 2005). Both children showed a discrepancy between their language and nonverbal IQ test scores, and they were not otherwise different from the larger EO-LI group in their language abilities. The mean English CELF–4 composite standard score of the group was 64.7 (*SD* = 12.8). The mean expressive score was 65.5 (*SD* = 12.3), and the mean receptive score was 72.4 (*SD* = 11.8).

Bilingual speakers with typical development (BI group)—The 65 typically developing bilingual Spanish-English children had a mean age of 8;0 (*SD* = 1;4). All children in the BI group were Hispanic, with four children reporting that they were biracial. As with the EO group, children in the BI group were reported to have language skills consistent with their age and language experiences, had never repeated a grade or received special educational services for language, and scored at or above –1 *SD* on the TONI–3. All participants in the BI group learned Spanish as their L1 at home and English as their L2 in their school as well as through interactions with the broader English-speaking community. By definition, bilingual children with LI have lower skills in both their L1 and L2 as compared with age- and experience-matched peers (American Speech-Language-Hearing Association, 2004; see also Kohnert, 2008; Peña & Bedore, 2009). Thus, to be included in the typical BI group, children needed to score within the normal range on either the Spanish or English CELF–4 Recalling Sentences and Concepts and Directions subtests, but not necessarily both. Children showed average performance on the Spanish CELF–4 subtests. As anticipated for these young sequential bilinguals, some children received lower scores on

the English subtests (see Table 1). The average of the two Spanish subtest standard scores ranged from 5 to 15 across children. The average of the two English subtest scores ranged from 1 to 16.

This within-group variation on language test scores is consistent with a growing literature investigating ability in both languages of developing bilinguals. That is, even when the ages and contexts of L1 and L2 acquisition are similar among minority L1 children learning a majority L2, there is considerable variation in performance within any well-defined group of typically developing bilingual learners on most language outcome measures (Kohnert, Kan, & Conboy, 2010). This is also true to some extent for typically developing monolingual children (see Marchman & Thal, 2005, for review) but is exacerbated in children exposed to two languages.

Bilingual speakers with primary LI (BI-LI group)—The 19 children in this group had a mean age of 8;5 (*SD* = 1;7). All children were Hispanic, with one biracial child. To be included in the BI-LI group, children had to show a significant difficulty in both languages, defined as receiving an average standard score across the two Spanish CELF–4 subtests no greater than 5 and an average standard score across the two English subtests also no greater than 5. As in the EO-LI group, there were two children in the BI-LI group with TONI–3 standard scores below –1.3 *SD*s but who otherwise resembled other children in the BI-LI group in their language performance. Fifteen of the 19 children had been diagnosed with LI by certified speech-language pathologists and were receiving speech-language pathology services in their schools. Parents reported significant concerns about the language skills of two of the other children. Table 1 shows that the BI-LI group had similar CELF–4 subtest scores in their L1, Spanish, as the EO-LI group in their only language, English. That is, the two LI groups appeared to show approximately equivalent language difficulties on average.

Nonword Repetition Tasks

Each of the 187 children participated in both English and Spanish NWR tasks in either their school or our research laboratory. For English NWR, we used Dollaghan and Campbell's (1998) Nonword Repetition Task, which includes 96 phonemes in 16 nonwords, four at each length from one to four syllables (e.g., /naɪb/, /teɪvak/, /dɔɪtaʊvaeb/, and /taevatʃinaɪg/). The items follow English phonotactics and a CVC syllable template but exclude late-developing consonants, consonant clusters, and lax vowels, leading to stimuli that do not follow the metrical stress pattern of English and do not have a close resemblance to real English words. For Spanish NWR, we used the task developed by Ebert, Kalanek, Cordero, and Kohnert (2008). This task includes 120 phonemes in four nonwords at each length from one to five syllables and follows a CV template, which is most common in Spanish (e.g., /gi/, /bopa/, /nutife/, ditujaβu/, and /betetʃoδupe/). Spanish words are longer, on average, than English words; therefore, stimuli of up to five syllables are included in this Spanish NWR task. The nonwords follow the phonotactic constraints of Spanish, use only early-developing consonants and tense vowels, and do not bear a close resemblance to Spanish words. The nonwords do maintain the Spanish pattern of stress on the penultimate syllable. Although the two tasks are parallel, they are not designed to be directly equivalent in item difficulty (see Ebert et al., 2008, for a full description and stimuli list of the Spanish nonwords).

The nonwords in each task were spoken by native speakers and recorded onto a CD. The two NWR tasks were administered in counterbalanced order by trained research assistants. Instructions for the Spanish NWR task were provided in Spanish for the bilingual children; English instructions were provided for the English NWR task. The two tasks were administered to most children in the EO-LI and BI-LI groups in a quiet university room, and to most other children in a quiet room at their schools. The nonwords in both tasks were presented under headphones to all children, and children provided one imitation of each nonword. Children's productions were audio-recorded, and each task was scored for percentage of phonemes produced correctly.

Reliability

For the purpose of reliability checking, 20% of the children's NWR tasks in both languages were transcribed by a second trained judge. Seventy-four of the children's NWR tasks, 37 in each language, were transcribed by two scorers to obtain a measure of interjudge reliability. Average percentage agreement across children in transcribing each phoneme for English NWR was 89.2% (i.e., there was agreement on 3,169 of the 3,552 phonemes). Average percentage agreement for Spanish NWR was 91.0% (i.e., agreement on 4,040 of the 4,440 phonemes). In each language, reliability for one child's NWR responses was not as high as for other children, affecting the average reliability. Modal reliability for English and Spanish NWR was, respectively, 93.8% and 96.7%.

Analyses

Group comparisons were first performed for each NWR task using analysis of variance (ANOVA). NWR task correlations across languages for each of the monolingual and bilingual groups were performed using Pearson's rho, with separate Bonferroni-adjusted correlations for the typically developing children and children with LI in each language. NWR task sensitivity and specificity values were used to calculate the likelihood that each NWR task could robustly identify children with and without LI.

Results

Group Differences

Group differences were examined in two ways, by overall task accuracy and by syllable length in each task. Figure 1 shows the English NWR accuracy (percentage of phonemes correct) for each of the four groups of children across the four syllable lengths. Table 2 shows the mean total percentage of phonemes correct for each group. There was a clear effect of native language experience in addition to effects of typical or LI group status and nonword syllable length. A two-way ANOVA (Group × Syllable Length) showed that there were significant main effects of group, $F(3, 183) = 16.70$, $p < .001$, $\eta_p^2 = .215$, and syllable length, $F(3, 747) = 82.88$, $p < .001$, $\eta_p^2 = .58$, and a significant Group × Syllable Length interaction effect, $F(9, 747) = 4.13$, $p < .001$, $\eta_p^2 = .063$. These effect sizes indicate that there was a moderate effect of group, relatively large effect of syllable length, and a small interaction effect. Post hoc comparisons (using $\alpha = .05$) to identify the source of the main

and interaction effects showed that the EO group had significantly higher total English NWR than all other groups. The typical BI group had overlapping performance with the EO-LI group and significantly higher accuracy than the BI-LI group. All groups showed significantly lower accuracy with the longest, four-syllable nonwords. Most of the group difference was attributable to performance on the three- and four-syllable nonwords. At the three-syllable length, the EO group significantly outperformed all groups. At the four-syllable level, the two typical groups both significantly outperformed the two groups with LI.

Figure 2 shows the Spanish NWR accuracy of each of the four groups of children across the five syllable lengths in this task; Table 2 shows the total percentage of phonemes correct. There were significant main effects of group, $F(3, 183) = 20.01$, $p < .00$, $\eta_p^2 = .247$, and syllable length, $F(4, 934) = 82.92$, $p < .001$, $\eta_p^2 = .312$, as well as a significant Group × Syllable Length interaction effect, $F(12, 934) = 4.16$, $p < .001$, $\eta_p^2 = .08$. These effect sizes indicate that there were moderate effects of group and syllable length and a relatively small interaction effect. The typical BI group had higher Spanish NWR accuracy than the typical EO group and the BI-LI group. The EO-LI group had lower accuracy than the BI-LI group and much lower accuracy than the typical EO group (see Figure 2). Post hoc comparisons to clarify the main effects showed that, as for English NWR, performance at the two longest syllable lengths most clearly separated the groups. The BI group showed decreased NWR accuracy only with five-syllable nonwords and otherwise had close to ceiling performance on this task. The EO and BI-LI groups had significantly lower performance than the typical BI group for four-syllable nonwords. The EO-LI group had significantly lower performance than all other groups for four- and five-syllable nonwords. There was the same trend for the BI-LI group, although this difference failed to reach statistical significance for five-syllable nonwords.

Task Correlations

The overall group results show the advantage for native-language speakers in the NWR tasks. The monolingual English-speaking children performed with higher accuracy in English NWR than did the bilingual English-Spanish-speaking children. The bilingual children had higher NWR accuracy in their L1, Spanish, than the English-only speaking children. We next examined within-group correlations between the two NWR tasks in separate analyses for each of the EO, BI, EO-LI, and BI-LI groups (see Tables 3 and 4). For each group, we were interested in the correlations between the two NWR tasks and the language subtest standard scores beyond any shared relation with chronological age and nonverbal IQ (TONI–3 standard score). Because group differences were most apparent at the longer syllable lengths, and with potential ceiling effects at the shorter syllable lengths, we used the longest syllable length in each NWR task as the dependent measure for these tasks.

Tables 3 and 4 show the Pearson partial correlations (adjusted for age and nonverbal IQ) among the NWR tasks and language subtests. For the EO group, there was a moderate correlation between English and Spanish NWR and between each of the NWR tasks and the expressive English CELF–4 subtest, Recalling Sentences. There was a slightly lower but still

significant correlation between the two NWR tasks and the receptive English CELF–4 subtest, Concepts and Directions. The same pattern was evident for the EO-LI group, except that the correlation between Spanish NWR and Concepts and Directions was negligible (see Table 3).

Table 4 shows that for the typical BI group, English and Spanish NWR again were robustly correlated. English NWR also was significantly correlated with the two English CELF–4 subtests but not with the Spanish CELF–4 subtests. Spanish NWR had negligible correlations with any of the four CELF–4 subtests. The relatively small BI-LI group precluded all but the most robust correlations from being statistical significant. For this group, Spanish NWR was significantly correlated with English Concepts and Directions and tended to be associated with Spanish Recalling Sentences. English NWR tended to be associated with English and Spanish Concepts and Directions. The two NWR tasks were not significantly correlated, and the value of $r = .197$ was much lower than found for the typical bilingual group.

Task Sensitivity and Specificity

A central question was whether accuracy in either English or Spanish NWR would be sufficient to separate children with LI from their typically developing peers, regardless of the children's native language. For English NWR, the group comparisons indicated that repetition of the four-syllable nonwords was most likely to separate both the EO-LI and BI-LI groups from the other two groups (see also Figure 1). A cutoff score of 78% accuracy at this syllable length gave the largest separation between groups. Fifty of the 69 children in the EO group and 37 of the 65 children in the BI group achieved at least this level of accuracy on the English NWR task. Only eight of the 34 children in the EO-LI group and one of the 19 children in the BI-LI group did so. This led to sensitivity values of 0.76 and 0.94 for the EO-LI and BI-LI groups, respectively, and specificity values of 0.73 and 0.57 for the typical monolingual and bilingual speakers, respectively. The moderate sensitivity and specificity for the monolingual English speakers was associated with a positive likelihood ratio of 2.78 and a negative likelihood ratio of 0.32. The high sensitivity and relatively low specificity for the bilingual speakers led to an equivalent positive likelihood ratio for the monolingual speakers (2.20) and a lower negative likelihood ratio (0.09).

Following conventional interpretation of ratios outlined by Mant (1999), only the negative likelihood ratio for the bilingual group can be considered to have adequate diagnostic power. That is, bilingual children who repeated greater than 78% of phonemes correctly in four-syllable English nonwords could be reliably ruled out from other bilingual children as having LI. Combining the monolingual and bilingual children as one larger group did not change these findings overall but did reduce the ability to rule out typical children as showing LI. A total of 87 of the 134 typically developing children achieved at least 78% accuracy, compared with nine of the 53 children with LI, providing sensitivity of 0.83 and specificity of 0.65. The positive likelihood ratio was 2.37, and the negative likelihood ratio was more modest at 0.26.

For Spanish NWR, the group comparisons indicated that repetition of both four- and five-syllable nonwords potentially could separate the EO-LI group from the EO and BI groups,

and that NWR of five-syllable words could be informative for the BI-LI group (see Figure 2). Using a cutoff score of 80% correct, similar diagnostic power was found for five-syllable Spanish nonwords as for English NWR. Of the 134 typically developing children, 97 children had at least 80% accuracy (44 EO and 53 BI). Of the 53 children with LI, 16 children reached this level of accuracy (8 EO-LI and 8 BI-LI). This provided sensitivity to LI of 0.77 and 0.58 for the monolingual and bilingual groups, and specificity of 0.64 and 0.82, respectively. These moderate values were associated with a positive likelihood ratio of 2.11 and a negative likelihood ratio of 0.37 for monolingual children, and a positive ratio of 3.14 and negative ratio of 0.52 for the bilingual groups. None of these values can be considered to demonstrate sufficient diagnostic power by conventional standards. Combining the monolingual and bilingual groups does not affect diagnostic power significantly, with overall sensitivity of 0.70 and specificity of 0.72. This provides a positive likelihood ratio of 2.53 and a negative likelihood ratio of 0.42.

Discussion

This study is the first to compare the NWR performance of Spanish-English bilingual and monolingual English children with LI in both Spanish and English. The monolingual children's NWR is discussed below, followed by the bilingual children's performance and a comparison of relative performance across groups.

Monolingual Children's NWR

Confirming previous group results, we found that monolingual English-speaking children with LI performed below their monolingual age peers in accurately repeating three- and four-syllable English nonsense words. The children with LI also performed below their monolingual age peers in NWR in an unfamiliar language, Spanish. Again, the significant group difference was for longer, four- and five-syllable nonwords. Our results are insufficient to identify short-term memory as the only factor influencing this performance profile (see also MacDonald & Christiansen, 2002). However, the finding that it is repetition of only longer nonwords which separates the groups with both more and less familiar linguistic material is well in line with an explanation of LI performance that emphasizes short-term memory difficulties.

In anticipation of this word length effect and possible ceiling effects for shorter nonwords, we analyzed the correlation between English and Spanish NWR using the longest syllable length in each task. The two NWR tasks were not designed specifically to have equivalent difficulty, and the Spanish nonwords contained a few allophonic variations not found in the English NWR task. However, for each of the EO and EO-LI groups, overall task accuracy in percentage of phonemes correct was essentially the same across languages (see Table 2). Beyond the contributions of age and nonverbal IQ, children's English NWR performance was correlated with their Spanish NWR for both the EO and EO-LI groups. NWR performance also was correlated with the standardized language subtests for both the EO and EO-LI groups. The one exception to this was that the EO-LI group's Spanish NWR was not associated with their performance on the English Concepts and Directions subtest.

Certainly, the task demands presented by NWR were sensitive to the monolingual EO-LI profile, at least at the group level. To some extent, however, these results are unsurprising. While we did not obtain a separate measure of short-term memory, the EO-LI and typical EO groups already had been identified a priori based on their ability in the native language. The group results would carry greater importance if they could be generalized to the level of individual monolingual children with LI. The positive likelihood ratio of 2.78 obtained here for English NWR (using a cutoff score of 78%) identifies that it is almost three times more likely that a child with a score below the cutoff has LI rather than typical development. The negative likelihood ratio of 0.32 for English NWR indicates that it is about one third as likely that a child scoring above this cutoff has LI rather than typical development. Similar likelihood ratios were found for the English-speaking children's performance in Spanish NWR (using a cutoff score of 80%), with a positive likelihood ratio of 2.11 and a negative likelihood ratio of 0.37.

These kinds of low-to-intermediate likelihood ratios are not unlike those found in previous studies of English NWR with other samples of English-speaking children (e.g., Conti-Ramsden et al., 2001; Kohnert et al., 2006). The group of school-age monolingual children with LI examined here was heterogeneous with respect to their relative skills in expressive and receptive language. It is possible that this variation diminished the clinical significance of the NWR effect. However, the children with LI tended to show marked language difficulties overall, with a group mean language test score of –2.3 *SD*s. That English NWR accuracy was not powerful enough to have diagnostic salience for this group suggests that NWR in isolation is unlikely to be a valuable clinical marker for children in the school years. It remains possible that children's NWR performance could be combined with their performance on other tasks and found to be part of a composite clinical marker of LI for monolingual children.

Bilingual Children's NWR

As found by Girbau and Schwartz (2008), the typical BI group had higher Spanish NWR accuracy than the BI-LI group. As with the monolingual English children in their native language, this difference was most apparent in the current study at longer syllable lengths, with the greatest difference for four-syllable nonwords. The same pattern of significantly lower accuracy was found for the BI-LI group's English NWR. On this task, the typical BI group outperformed the BI-LI group in repeating four-syllable nonwords. Thus, the bilingual children who were diagnosed with LI demonstrated poorer NWR performance at longer syllable lengths in both their L1 and L2.

Table 2 shows that, unlike the monolingual English children, the bilingual children with and without LI consistently had higher NWR accuracy in Spanish (their L1) than in English (their L2). This is in line with Masoura and Gathercole's (1999) finding for higher NWR accuracy in the L1 and presumably reflects the role of language experience in NWR in addition to children's short-term memory capacities. This result is also consistent with the higher CELF–4 performance in Spanish than English, on average, for the BI and BI-LI groups. However, it should be noted that the differences in standardization for the Spanish

and English versions of the CELF–4 mean that these scores cannot be used to interpret relative levels of ability in the L1 and L2.

As with the monolingual children, the bilingual children's NWR was moderately correlated across languages beyond any shared variance due to chronological age and nonverbal IQ. Like the monolingual children, the two NWR tasks were significantly correlated for the typical bilingual group, in this case with a moderately high association between the two tasks. There also were significant correlations between English NWR and the two English CELF–4 subtests for these typical bilingual speakers. To the extent that the subtests rely on grammatical skill, this correlation in L2 is similar to that found by French and O'Brien (2008) between current NWR and later grammatical skills. Spanish NWR was not correlated with Spanish or English CELF–4 subtest performance. Correlations were less predictable for the BI-LI group, including a finding of no significant correlation between Spanish and English NWR. This may be an artifact of heterogeneity within the relatively small sample of 19 children and warrants additional investigation.

Likelihood ratios for the bilingual children's L1 (Spanish) were similar to those found for the monolingual children's only language (English). The positive likelihood ratio of 2.78 and negative likelihood ratio of 0.32 suggest that Spanish NWR has some potential value as part of a composite marker to identify LI but is not particularly powerful alone. While an equivalent positive likelihood ratio of 2.20 was found for the bilingual children's English NWR, the negative likelihood ratio of 0.09 does reach the conventional level for diagnostic accuracy. In this case, a score of over 80% in English NWR was 11 times more likely to come from a bilingual child with typical development than from a bilingual child with LI. The BI-LI group was an average of 5 months older than the BI group. It is possible that this age difference diminished the clinical significance of the NWR effect in our sample of children. However, like the EO-LI group, children in the BI-LI group also had relatively severe language difficulties that warranted clinical services (see Table 1). Even so, higher positive likelihood ratios would be necessary to identify this group of children with LI using NWR alone. However, the negative likelihood ratio indicates that English NWR is sufficient to rule out typical bilingual children as having LI, when their performance is compared with other bilingual children with similar social and language-learning opportunities. These results also underscore the challenge in using standardized, experience-dependent measures administered separately in the L1 and L2 to identify LI in children who are in a highly dynamic period of language development, often characterized by shifts in relative proficiency levels in home and school/community languages (e.g., Kohnert et al., 2009).

Comparing Monolingual and Bilingual Children's NWR

The study results replicate Thorn and Gathercole's (1999) finding that young French-English bilingual children who have English as their L2 outperform monolingual English-speaking children in the bilingual children's L1. The results also replicate Kohnert et al.'s (2006) finding that typical bilingual Spanish-English children's English NWR falls above that of monolingual English speakers with LI but below that of typical monolingual English speakers (see Figure 1). Kohnert et al. interpreted this finding as the interplay between decreased integrity of the underlying language system for the children with LI and

differential English experience for the bilingual children—that is, that both language-learning ability/disability and (length of) language experience influence NWR performance.

The novel contribution of the current study is the examination of both monolingual and bilingual children's NWR across two languages. Speaking to the influence of language exposure or experience, the typical Spanish-English bilingual speakers had higher accuracy than the typical monolingual English speakers in Spanish NWR, and the reverse was true in English NWR. The same effect of language-specific experience held for monolingual and bilingual children with LI. Thus, the effects of relative language experience and underlying language-learning integrity played out in the same way across the two languages.

The effect of language experience on NWR performance appeared to be independent of the effect of underlying language ability (typical or impaired group status). The pattern of decreasing accuracy with increasing syllable length was evident for all groups in both English and Spanish. However, both monolingual and bilingual children with LI were separated from their typical peers only at longer syllable lengths. Thus, while differential language experiences affected the performance of all groups, short-term memory capacity appeared to be a significant factor separating children with and without LI. Where language-specific experience and underlying language ability intersected was in comparisons across monolingual and bilingual children. The typical bilingual children performed similarly to the monolingual English children with LI in English NWR, and the typical English monolinguals performed similarly to the bilingual children with LI in Spanish NWR.

Conclusion

The primary study finding was that NWR performance is influenced by experience with the target language. As part of a composite marker, NWR performance may be sensitive to LI within a given language. However, for these school-age monolingual and bilingual children, there was adequate diagnostic power only in ruling out the typical bilingual children as showing LI when administered an English NWR task. There was a broad age range across participants, and the bilingual children had varying levels of English and Spanish proficiency as measured by standardized language subtests. As noted by Gutiérrez-Clellen and Simon-Cereijido (2010), individual variation in language experience may have influenced the bilingual children's NWR performance. Given the variation in children's performance on the standardized language measures, such individual differences likely also would be evident in NWR. Overall, the current study supports an increasing body of literature demonstrating that NWR in a single language is not sufficient to act as a clinical marker of LI in linguistically diverse populations.

Acknowledgments

This study was supported by Grant R21 HD053222 to the first two authors from the National Institute of Child Health and Human Development and by an award to the third author from the Undergraduate Research Opportunities Program at the University of Minnesota. We are grateful to the following research assistants for help with data collection and scoring: Heather Bauer, Olga Campbell, Kelly Nett Cordero, Kerry Ebert, Rachel Hopp, Erika Hoyt, Hannah Julien, Pui Fong Kan, Chao Khang, Alaina Kelley, Laura Kluge, Lacey LeBarge, Amelia Medina, Vanessa Monge, Abby Pomeroy, Marjorie Southward, Lacey Thomas, Karyn Wagner, and Jamie Wennblom. Special thanks to participants and their families as well as to Frank Cirrin and the faculty and speech-language pathologists at Minneapolis Public Schools and Aurora Charter School.

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

Anderson JD, Wagovich SA, Hall NE. Non-word repetition skills in young children who do and do not stutter. Journal of Fluency Disorders. 2006; 31:177–199. [PubMed: 16814376]

Archibald LMD, Gathercole SE. Nonword repetition: A comparison of tests. Journal of Speech, Language, and Hearing Research. 2006; 49:970–983.

Archibald LMD, Gathercole SE. Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. Psychonomic Bulletin & Review. 2007; 14:919–924. [PubMed: 18087960]

Archibald LMD, Joanisse MF, Shepherd M. Associations between key language-related measures in typically developing school-age children. Journal of Psychology. 2008; 216:162–172.

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

Bowey JA. Clarifying the phonological processing account of nonword repetition. Applied Psycholinguistics. 2006; 27:548–552.

Brown, L., Sherbenou, R., Johnsen, S. Test of Nonverbal Intelligence, Third Edition. Austin, TX: Pro-Ed; 1997.

Calderón, J. Working memory in Spanish-English bilinguals with language impairment (Unpublished doctoral dissertation). University of California, San Diego and San Diego State University; 2003.

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

Catts HW, Adlof SM, Hogan TP, Ellis Weismer S. Are specific language impairment and dyslexia distinct disorders? Journal of Speech, Language, and Hearing Research. 2005; 48:1378–1396.

Cheung H. Nonword span as a unique predictor of second language vocabulary learning. Developmental Psychology. 1996; 32:867–873.

Chiat S, Roy P. The Preschool Repetition Test: An evaluation of performance in typically developing and clinically referred children. Journal of Speech, Language, and Hearing Research. 2007; 50:429–443.

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

Conti-Ramsden G, Botting N, Faragher B. Psycholinguistic markers for specific language impairment (SLI). Journal of Clinical Psychology and Psychiatry. 2001; 42:741–748.

Dollaghan C, Campbell TF. Nonword repetition and child language impairment. Journal of Speech, Language, and Hearing Research. 1998; 41:1136–1146.

Ebert KD, Kalanek J, Cordero KN, Kohnert K. Spanish nonword repetition: Stimuli development and preliminary results. Communication Disorders Quarterly. 2008; 29:69–74.

Edwards J, Beckman ME, Munson B. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in non-word repetition. Journal of Speech, Language, and Hearing Research. 2004; 47:421–436.

Ellis Weismer S, Tomblin JB, Zhang X, Buchwalter P, Chynoweth JG, Jones M. Nonword repetition performance in school-age children with and without language impairment. Journal of Speech, Language, and Hearing Research. 2000; 43:865–878.

French LM, O'Brien I. Phonological memory and children's second language grammar learning. Applied Psycholinguistics. 2008; 29:463–487.

Gathercole SE. Nonword repetition and word learning: The nature of the relationship. Applied Psycholinguistics. 2006; 27:513–543.

Gathercole SE, Baddeley AD. Phonological memory deficits in language disordered children: Is there a causal connection? Journal of Memory and Language. 1990a; 29:336–360.

Gathercole SE, Baddeley AD. The role of phonological memory in vocabulary acquisition: A study of young children learning new names. British Journal of Psychology. 1990b; 81:439–454.

Gathercole, SE., Baddeley, AD. The Children's Test of Nonword Repetition. London, England: The Psychological Corporation; 1996.

Gathercole SE, Hitch GJ, Service E, Martin AJ. Short-term memory and new word learning in children. Developmental Psychology. 1997; 33:966–979. [PubMed: 9383619]

Gathercole SE, Willis C, Emslie H, Baddeley AD. Phonological memory and vocabulary development during the early school years: A longitudinal study. Developmental Psychology. 1992; 28:887–898.

Girbau D, Schwartz RG. Non-word repetition in Spanish-speaking children with specific language impairment (SLI). International Journal of Language & Communication Disorders. 2007; 42:59–75. [PubMed: 17365086]

Girbau D, Schwartz RG. Phonological working memory in Spanish-English bilingual children with and without specific language impairment. Journal of Communication Disorders. 2008; 41:124–145. [PubMed: 17765916]

Graf Estes K, Evans JL, Else-Quest NM. Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. Journal of Speech, Language, and Hearing Research. 2007; 50:177–195.

Gray S. Diagnostic accuracy and test-retest reliability of nonword repetition and digit span tasks administered to preschool children with specific language impairment. Journal of Communication Disorders. 2003; 36:129–151. [PubMed: 12609578]

Gupta P. Commentaries: Nonword repetition, phonological storage, and multiple determination. Applied Psycholinguistics. 2006; 27:564–568.

Gupta P, Lipinski J, Abbs B, Lin P. Serial position effects in nonword repetition. Journal of Memory and Language. 2005; 53:141–162.

Gupta P, MacWhinney B, Feldman HM, Sacco K. Phonological memory and vocabulary learning in children with focal lesions. Brain and Language. 2003; 87:241–252. [PubMed: 14585293]

Gutiérrez-Clellen VF, Simon-Cereijido G. Using nonword repetition tasks for the identification of language impairment in Spanish-English-speaking children: Does the language of assessment matter? Learning Disabilities Research & Practice. 2010; 25:48–58. [PubMed: 22707854]

Hayiou-Thomas ME, Oliver B, Plomin R. Genetic influences on specific versus nonspecific language impairment in 4-year-old twins. Journal of Learning Disabilities. 2005; 38:222–232. [PubMed: 15940960]

Kan PF, Kohnert K. Preschoolers learning Hmong and English: Lexical-semantic skills in L1 and L2. Journal of Speech, Language, and Hearing Research. 2005; 48:372–383.

Kohnert, K. Language disorders in bilingual children and adults. San Diego, CA: Plural; 2008.

Kohnert K, Kan PF, Conboy B. Lexical and grammatical associations in sequential bilingual preschoolers. Journal of Speech, Language, and Hearing Research. 2010; 53:1–15.

Kohnert K, Medina A. Bilingual children and communication disorders: A 30-year research retrospective. Seminars in Speech and Language. 2009; 30:219–233. [PubMed: 19851950]

Kohnert K, Windsor J, Ebert KD. Primary or "specific" language impairment and children learning a second language. Brain and Language. 2009; 109:101–111. [PubMed: 18313136]

Kohnert K, Windsor J, Yim D. Do language-processing tasks separate children with language impairment from typical bilinguals? Learning Disabilities Research & Practice. 2006; 21:19–29.

Kovács G, Racsmány M. Handling L2 input in phonological STM: The effect of non-L1 phonetic segments and non-L1 phonotactics on nonword repetition. Language Learning. 2008; 58:597–624.

Leonard L. Specific language impairment as a clinical category. Language, Speech, and Hearing Services in Schools. 1991; 22:66–68.

Leonard, L. Children with specific language impairment. Cambridge, MA: MIT Press; 1998.

MacDonald MC, Christiansen MH. Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1992). Psychological Review. 2002; 109:35–54. [PubMed: 11863041]

Mant, J. Studies assessing diagnostic tests. In: Dawes, M.Davies, P.Gray, A.Mant, J.Seers, K., Snowball, R., editors. Evidence-based practice: A primer for health care professionals. New York, NY: Churchill Livingstone; 1999. p. 67-78.

Marchman, V., Thal, D. Words and grammar. In: Tomasello, M., Slobin, DI., editors. Beyond nature and nurture: Essays in honor of Elizabeth Bates. Mahwah, NJ: Erlbaum; 2005. p. 141-164.

Masoura EV, Gathercole SE. Phonological short-term memory and foreign language learning. International Journal of Psychology. 1999; 34:383–388.

Masoura EV, Gathercole SE. Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. Memory. 2005; 13:422–429. [PubMed: 15948628]

Metsala JL, Stavrinos D, Walley AC. Children's spoken word recognition and contributions to phonological awareness and nonword repetition: A 1-year follow-up. Applied Psycholinguistics. 2009; 30:101–121.

Montgomery JW. Examination of phonological working memory in specifically language-impaired children. Applied Psycholinguistics. 1995; 16:355–378.

Montgomery J, Evans J. Commentaries: Commentary on keynote. Applied Psycholinguistics. 2006; 27:573–577.

Montgomery JW, Magimairaj BM, O'Malley MH. Role of working memory in typically developing children's complex sentence comprehension. Journal of Psycholinguistic Research. 2008; 37:331–354. [PubMed: 18521752]

Munson B, Edwards J, Beckman ME. Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. Journal of Speech, Language, and Hearing Research. 2005; 48:61–78.

Munson B, Kurtz BA, Windsor J. The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. Journal of Speech, Language, and Hearing Research. 2005; 48:1033–1047.

Oetting JB, Cleveland LH. The clinical utility of nonword repetition for children living in the rural south of the US. Clinical Linguistics & Phonetics. 2006; 20:553–561. [PubMed: 17056486]

Palladino P, Cornoldi C. Working memory performance of Italian students with foreign language learning difficulties. Learning and Individual Differences. 2004; 14:137–151.

Peña, E., Bedore, L. Bilingualism in child language disorders. In: Schwartz, RG., editor. Handbook of child language disorders. New York, NY: Psychology Press; 2009. p. 281-307.

Peña E, 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; 45:938–947.

Preston JL, Edwards ML. Phonological processing skills of adolescents with residual speech sound errors. Language, Speech, and Hearing Services in Schools. 2007; 38:297–308.

Rodekohr R, Haynes W. Differentiating dialect from disorder: A comparison of two processing tasks and a standardized test. Journal of Communication Disorders. 2001; 34:255–272. [PubMed: 11409607]

Roodenrys S, Stokes J. Serial recall and nonword repetition in reading disabled children. Reading and Writing. 2001; 14:379–394.

Roy P, Chiat S. A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: Evidence from typically developing children. Journal of Speech, Language, and Hearing Research. 2004; 47:223–234.

Semel, E., Wiig, EH., Secord, W. Clinical Evaluation of Language Fundamentals, Fourth Edition. San Antonio, TX: The Psychological Corporation; 2003.

Service E. Phonology, working memory, and foreign-language learning. The Quarterly Journal of Experimental Psychology: Human Experimental Psychology. 1992; 45(A):21–50. [PubMed: 1636010]

Service E, Kohonen V. Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? Applied Psycholinguistics. 1995; 16:155–172.

Smith Gabig C. Verbal working memory and story retelling in school-age children with autism. Language, Speech, and Hearing Services in Schools. 2008; 93:498–511.

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

Stokes SF, Klee T. Factors that influence vocabulary development in two-year-old children. Journal of Child Psychology and Psychiatry. 2009; 50:498–505. [PubMed: 19017366]

Thordardottir E. Language-specific effects of task demands on the manifestation of specific language impairment: A comparison of English and Icelandic. Journal of Speech, Language, and Hearing Research. 2008; 51:922–937.

Thorn ASC, Gathercole SE. Language-specific knowledge and short-term memory in bilingual and non-bilingual children. The Quarterly Journal of Experimental Psychology: Human Experimental Psychology. 1999; 52(A):303–324. [PubMed: 10371873]

Tomblin, JB., Zhang, X. Language patterns and etiology in children with specific language impairment. In: Tager-Flusberg, H., editor. Neurodevelopmental disorders. Cambridge, MA: MIT Press; 1999. p. 361-382.

U.S. Census Bureau. Language spoken at home by ability to speak English for the population 5 years and over (American Community Survey Table No. B16001). 2006. Retrieved from http://factfinder.census.gov/

Wiig, EH., Secord, W., Semel, E. Clinical Evaluation of Language Fundamentals, Fourth Edition: Spanish Edition. San Antonio, TX: The Psychological Corporation; 2006.

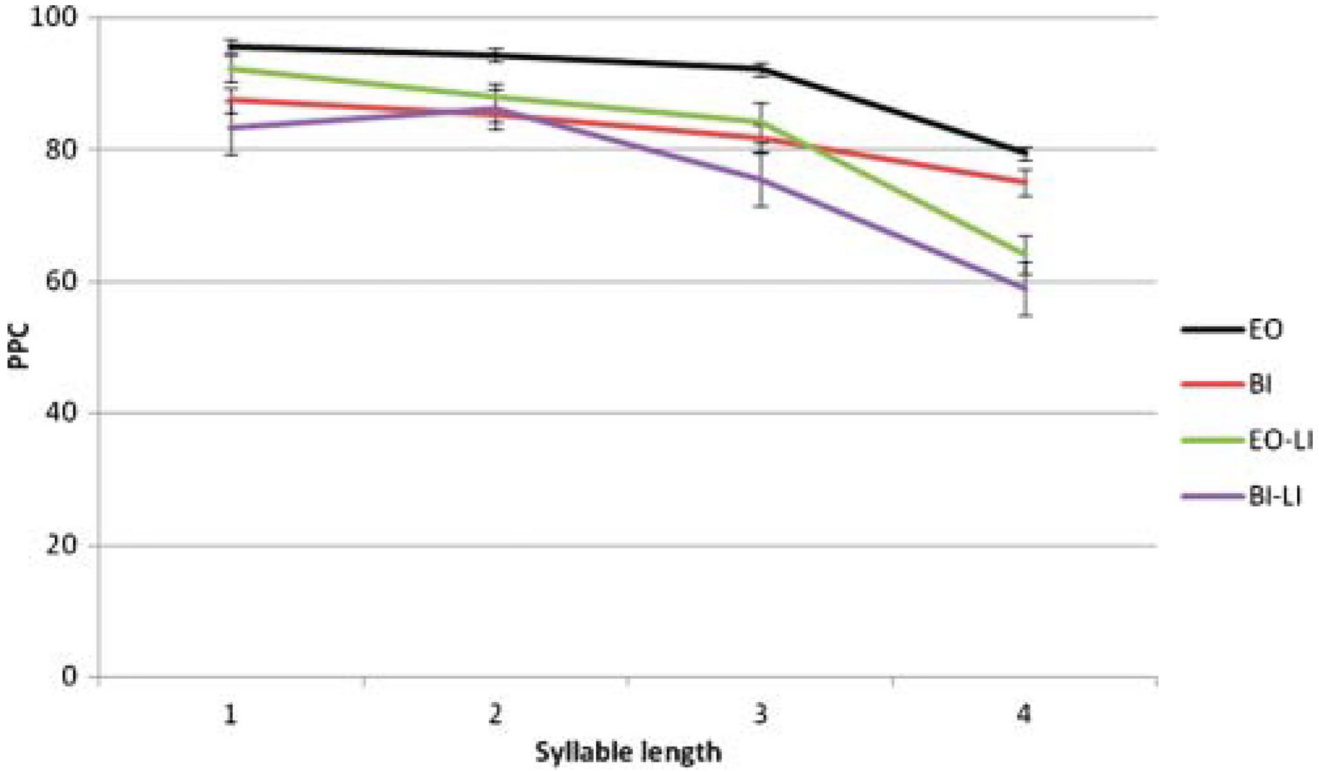

FIGURE 1.

Group mean English nonword repetition (NWR) performance across syllable length. Error bars indicate standard error. PPC = percentage of phonemes correct; EO = typical English-only; BI = typical bilingual; EO-LI = English-only with language impairment; BI-LI = bilingual with language impairment.

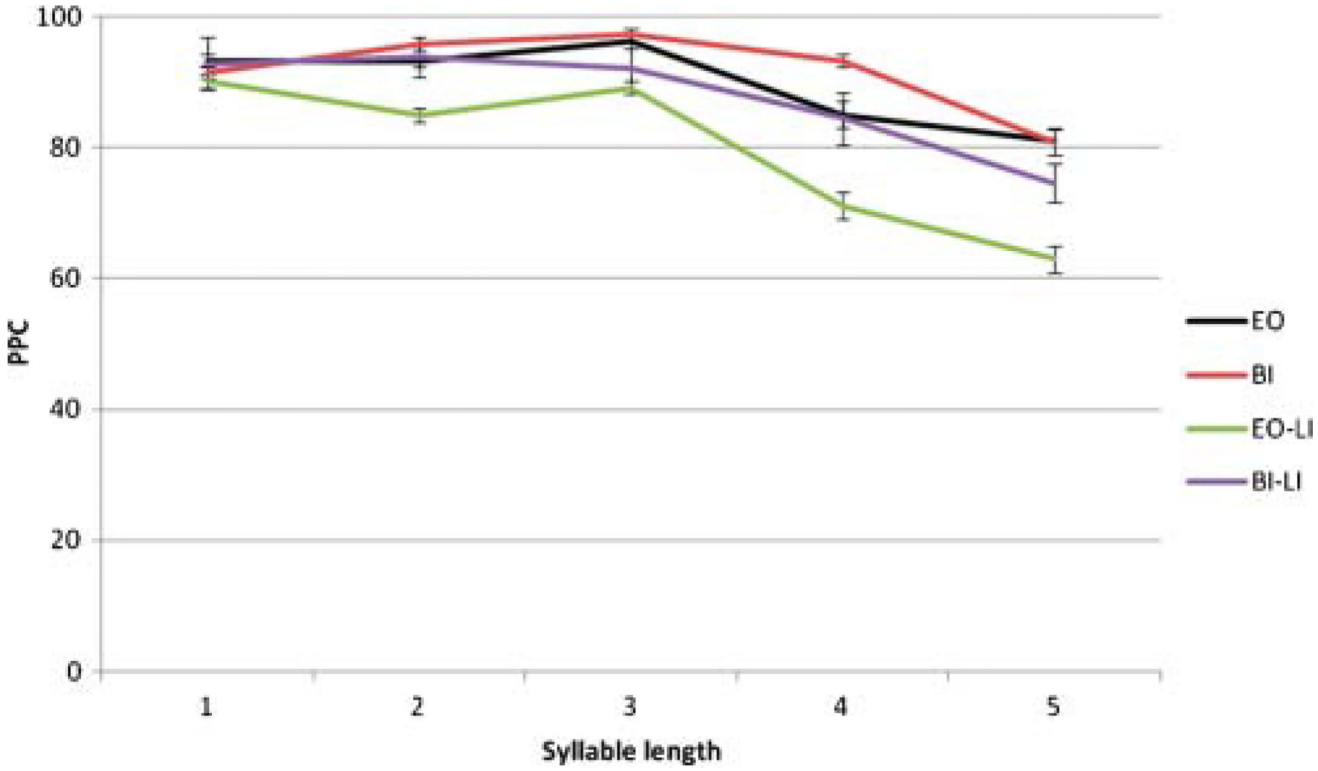

FIGURE 2.
Group mean Spanish NWR performance across syllable length. Error bars indicate standard error.

TABLE 1

Group mean language and nonverbal IQ test scores.

Group	*n*	Gender (M/F)	TONI–3	CELF–4			
				E-RS	E-CD	S-RS	S-CD
EO	69	35/34	110.4 (15.0)	11.6 (2.6)	10.5 (3.0)		
BI	65	28/37	103.5 (10.7)	5.6 (4.3)	6.4 (3.6)	9.5 (2.8)	9.0 (2.7)
EO-LI	34	20/14	93.4 (11.7)	3.9 (2.4)	3.7 (2.6)		
BI-LI	19	16/3	95.5 (15.0)	2.7 (1.3)	2.4 (1.5)	3.5 (2.0)	3.8 (1.8)

Note. TONI–3 = Test of Nonverbal Intelligence, Third Edition (Brown et al., 1997); CELF–4 = Clinical Evaluation of Language Fundamentals, Fourth Edition (Semel et al., 2003); E-RS and S-RS = English and Spanish versions of the CELF–4 Recalling Sentences subtest; E-CD and S-CD = English and Spanish versions of the CELF–4 Concepts and Directions subtest; EO = typical English-only; BI = typical bilingual; EO-LI = English-only with language impairment; BI-LI = bilingual with language impairment. Standard deviations are in parentheses. Excluding the two children in each of the EO-LI and BI-LI groups with lower IQ scores marginally changes the mean TONI–3 scores to 94.7 (*SD* = 10.9) and 98.9 (*SD* = 12.1), respectively. Excluding these four children does not noticeably alter the mean CELF–4 subtest scores.

Am J Speech Lang Pathol. Author manuscript; available in PMC 2018 April 17.

TABLE 2

Group mean total nonword repetition (NWR) scores (percentage of phonemes correct).

Group	English	Spanish
EO	88.3 (8.5)	87.5 (7.5)
BI	80.6 (1.3)	91.6 (1.0)
EO-LI	78.4 (2.0)	75.1 (2.5)
BI-LI	72.4 (2.9)	84.5 (1.9)

Note. Standard deviations are in parentheses.

TABLE 3

Pearson partial correlations (adjusted for age and IQ) among NWR performance and language test scores for the English-speaking children.

	E-NWR	S-NWR	E-CD	E-RS
		EO (n = 69)		
E-NWR	—	.375**	.285*	.421**
S-NWR		—	.290*	.377**
E-CD			—	.437**
E-RS				—
		EO-LI (n = 34)		
E-NWR	—	.322*	.378*	.539**
S-NWR		—	.033	.336*
E-CD			—	.034
E-RS				—

Note. E-NWR = four-syllable nonwords; S-NWR = five-syllable nonwords.

*Significant at $p < .05$.

**Significant at $p < .01$ (with Bonferroni adjustment).

TABLE 4

Pearson partial correlations (adjusted for age and IQ) among NWR performance and language test scores for the Spanish-English-speaking children.

	E-NWR	S-NWR	E-CD	E-RS	S-CD	S-RS
			BI (n = 65)			
E-NWR	—	.713**	.241*	.284*	.057	.108
S-NWR		—	.127	.068	.003	.042
E-CD			—	.031	.329**	.086
E-RS				—	.189	.234*
S-CD					—	.389**
S-RS						—
			BI-LI (n = 19)			
E-NWR	—	.197	.337	.082	.258	.138
S-NWR		—	.429*	.082	.081	.308
E-CD			—	.068	.529*	.063
E-RS				—	.105	.063
S-CD					—	.313
S-RS						—

*Significant at $p < .05$.

**Significant at $p < .01$ (with Bonferroni adjustment).

Am J Speech Lang Pathol. Author manuscript; available in PMC 2018 April 17.