On the Sensitivity and Specificity of Nonword Repetition and Sentence Recall to Language and Memory Impairments in Children

Lisa M. D. Archibald
Marc F. Joanisse
The University of Western Ontario

Purpose: The present study examined the utility of 2 measures proposed as markers of specific language impairment (SLI) in identifying specific impairments in language or working memory in school-age children.

Method: A group of 400 school-age children completed a 5-min screening consisting of nonword repetition and sentence recall. A subset of low (n = 52) and average (n = 38) scorers completed standardized tests of language, short-term and working memory, and nonverbal intelligence.

Results: Approximately equal numbers of children were identified with specific impairments in either language or working memory. A group about twice as large had deficits in both language and working memory. Sensitivity of the screening measure for both SLI and specific working memory impairments was 84% or greater, although specificity was closer to 50%. Sentence recall performance below the 10th percentile was associated with sensitivity and specificity values above 80% for SLI.

Conclusions: Developmental deficits may be specific to language or working memory, or include impairments in both areas. Sentence recall is a useful clinical marker of SLI and combined language and working memory impairments.

KEY WORDS: specific language impairment, working memory, short-term memory, nonword repetition, sentence recall

Specific language impairment (SLI) refers to an unexpected failure to develop language normally despite intact general cognitive, emotional, and hearing abilities and typical opportunities to acquire language. Part of the challenge presented by SLI relates both to the marked heterogeneity in the communication profiles of these children as well as the largely unexplained recovery by some proportion of the group. As a result, it has been difficult to determine the key deficit in SLI that will both contribute to researchers’ theoretical understanding of language learning and assist in reliable identification of children who require clinical services. In recent years, however, some progress has been made on this front: Several research groups have suggested clinical markers of the disorder, including nonword repetition (Bishop, North, & Donlan, 1996; Conti-Ramsden, Botting, & Faragher, 2001), sentence recall (Conti-Ramsden et al., 2001), and finite verb morphology (Bedore & Leonard, 1998; Leonard, Miller, & Gerber, 1999; Rice, 2003; Rice & Wexler, 1996). In the present study, we examined the utility as a screening measure of two of these markers: nonword repetition and sentence recall.
Clinical markers are phenotypic manifestations that characterize a particular disorder type (Bishop et al., 1996). The identification of a marker has both theoretical and clinical implications. Consistent behavioral traits provide clues as to the underlying cognitive processes that may give rise to the disorder, which in turn may influence theories aimed at explaining complex phenomena such as language. As well, such markers, or more specifically, the tests used to identify them, provide an important clinical tool for identifying individuals requiring further assessment, particularly if the measure is efficient and effective. This is especially true for uncommon disorders such as SLI that can be difficult and time-consuming to assess.

In order to be considered a “good” marker or test, the behavior in question must be present in individuals who have the disorder and absent in those who do not. Recognition that most tests tapping relevant markers are not entirely accurate has led to the development of several parameters for assessing and comparing test outcomes with results based on reference (“gold”) standards (i.e., measures widely accepted as identifying the disorder). Sensitivity refers to the accuracy of a test in detecting individuals with the disorder, and specificity is the accuracy in identifying those without the disorder. Although there are no generally accepted thresholds for sensitivity and specificity, some researchers have suggested that values above 80% for both should be considered minimal (Plante & Vance, 1994). Sensitivity and specificity do not tell the whole story, however. The practical implications of sensitivity and specificity rates vary depending on the prevalence of the disorder in the population. Similarly, predictive values for test results that are positive (i.e., positive predictive value [PPV] — percentage of positive results that are true positive) or negative (i.e., negative predictive value [NPV] — percentage of negative results that are true negatives) vary widely and nonlinearly, depending on the prevalence. For example, PPVs decrease sharply at lower prevalence levels even when sensitivity and specificity are both above 85% (Zenilman, Miller, Gaydos, Rogers, & Turner, 2003).

An alternative statistic for summarizing diagnostic accuracy of a test that is independent of the prevalence of the disorder is the likelihood ratio (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). Likelihood ratios (LRs) for positive test results reflect the ratio between the probability that a positive test comes from a person with the disorder and the probability of a positive test from a person without the disorder (i.e., sensitivity/1 — specificity). LRs incorporate sensitivity and specificity in a single measure that does not rely on the prevalence of the disorder in the population. As well, LRs can be calculated for different cut-points in a test rather than fixed at the rate of prevalence, allowing for the calculation of the best cut-off criterion. Large, positive LRs indicate that a positive result is much more likely to be associated with a person with the disorder. Generally, LRs above 10 are considered to indicate a strong and often conclusive increase in the likelihood of a disorder and ratios between 5 and 10 a moderately strong increase (Jaeschke, Guyatt, & Lijmer, 2002).

### Nonword Repetition

There has been considerable interest in identifying clinical markers for SLI both because of the potential such markers have to inform theories relating cognitive processes and language learning and to identify children requiring further clinical services. Nonword repetition was first suggested as a phenotypic marker of heritable language impairment based on a twin study by Bishop et al. (1996). This idea was bolstered by impressive findings that poor nonword repetition (a positive test result) was associated with a positive LR of 25 and almost perfectly (98%) discriminated children receiving clinical services for language impairment and typically developing peers (Dollaghan & Campbell, 1998). Further studies have indicated some need for caution, however. For example, poor nonword repetition had an LR of 6.5 for distinguishing children with and without language impairment in a population-based sample (Elliis Weismer et al., 2000), leading the authors to suggest that nonword repetition was a reliable but not sufficient marker of SLI. Mixed results have also been reported for sensitivity and specificity of nonword repetition; some authors have reported both values above 80% for monolingual English (Gray, 2003) and Italian preschoolers (Bortolini et al., 2006). However, Conti-Ramsden and colleagues have demonstrated very high specificity (85%–100%) but only fair sensitivity (66%–78%) for nonword repetition in 5-year-olds (Conti-Ramsden, 2003; Conti-Ramsden & Hesketh, 2003) and 11-year-olds (Conti-Ramsden et al., 2001). Good negative predictive value is reflected also in the finding for bilingual children that good nonword repetition ruled out language impairment, but poor nonword repetition did not necessarily rule in language impairment (Kohnert, Windsor, & Yin, 2006). Interestingly, nonword repetition was not found to discriminate language-impaired and typical groups of 5-year-old Cantonese speakers possibly due to the listlike nature of Cantonese compared with English nonwords (Stokes, Wong, Fletcher, & Leonard, 2006). Similar accuracy for language-impaired and typical English groups has been reported for one-syllable nonword list recall when adjusted for differences in short-term memory (Archibald & Gathercole, 2007). Taken together, these findings indicate that nonword...
replication is a useful clinical marker but insufficient on its own to identify SLI in English-speaking children.

The majority of studies in which nonword repetition has been investigated in SLI groups have used one of two tests: the Children’s Test of Nonword Repetition (CNRep; Gathercole & Baddeley, 1996) and the Nonword Repetition Test (NRT; Dollaghan & Campbell, 1998). These tests have been compared in detail elsewhere (Archibald & Gathercole, 2006b). Briefly, the CNRep has 40 items that are more “wordlike” than the NRT, and an online correct/incorrect item-level scoring system is used. The NRT has 16 items that were created to be unlike English words, and it is scored at the phoneme-level based on offline phonetic transcriptions. Percentage phonemes correct is often preferred for research purposes (e.g., Dollaghan & Campbell, 1998); however, online item-level scoring minimizes training needs and scoring time, making it more clinically practical (Gray, 2003). We recently compared item-level scoring based on online judgements and offline phonetic transcriptions for the 40 items of the CNRep and found very high agreement ($r = .90, p < .001$; Archibald, 2008). The feasibility of item-level scoring of the NRT has not been investigated previously, however.

### Sentence Recall

The sentence recall task (also sentence repetition, sentence imitation, recalling sentences) involves the immediate repetition of auditory sentences. Poor sentence recall performance in SLI versus typically developing groups has been reported in a number of studies (e.g., Briscoe, Bishop, & Norbury, 2001; Eadie, Fey, Douglas, & Parsons, 2002; Laws & Bishop, 2003; Norbury, Bishop, & Briscoe, 2001; Redmond, 2003). As well, sentence imitation tasks have been included as core subtests of language assessment batteries for years (e.g., the Clinical Evaluation of Language Fundamentals–4 [CELF-4]; Semel, Wiig, & Secord, 2003; Test of Language Development–3 [Newcomer & Hamill, 1997]). In a study of 160 eleven-year-old children in which the potential of four measures (nonword repetition, sentence recall, third-person singular, past tense) as psycholinguistic markers of SLI was examined, Conti-Ramsden et al. (2001) reported that sentence recall was the best clinical marker, with sensitivity and specificity at 90% and 85%, respectively. Interestingly, sentence recall (but not nonword repetition) discriminated language-impaired and unimpaired Cantonese-speaking children in the Stokes et al. (2006) study, with sensitivity and specificity at 77% and 97%, respectively. Despite this evidence, sentence recall has received comparatively little attention as a marker of SLI. Redmond (2003) noted that sentence recall tasks used in studies comparing impaired and unimpaired groups should include a graded scoring system (rather than correct/incorrect) and sufficiently complex (but not long) sentences in order to avoid ceiling effects in the typically developing groups.

It is no surprise that children with a language-learning impairment should score poorly on sentence recall measures. Explaining the deficit, however, presents more of a challenge because sentence recall presumably taps linguistic and memory systems at a variety of levels. Sentence recall has been found to be correlated with measures of phonological short-term memory, including nonword repetition (e.g., Bishop et al., 1996; Blake, Austin, Cannon, Lisus, & Vaughan, 1994; Conti-Ramsden et al., 2001; Kamhi & Catts, 1986; Willis & Gathercole, 2001). However, links exist also with linguistic skills such as grammar (Eadie et al., 2002) and vocabulary (Stokes et al., 2006). In the case of Cantonese, Stokes et al. (2006) found that whereas both nonword repetition and sentence recall were related to vocabulary skills, only sentence recall was associated with receptive grammar scores. This pattern of findings was taken to suggest that both sentence recall and nonword repetition tap phonological short-term memory, whereas only sentence recall is markedly influenced also by language ability. Other mediating skills, such as sequencing and temporal processing factors, also may provide possible explanations for the relationship between sentence recall and grammatical ability (Stokes et al., 2006).

### Identifying Language and Memory Impairments

It is clear from the preceding discussion that children may do poorly on nonword repetition and sentence recall tasks for language-related or memory-related reasons. This issue remains a matter of considerable debate among researchers in terms of how to explain the deficits in SLI. One suggestion has been that the low scores in nonword repetition and sentence recall reflect a core memory deficit in SLI that limits language learning (e.g., Gathercole & Baddeley, 1990). Others have argued that a core linguistic deficit in SLI constrains performance on these and all verbally mediated memory tasks (e.g., MacDonald & Christiansen, 2002). One problem in this line of research is that typical studies have focused on children with SLI and have not included sufficient measures to examine both the memory and language skills in good and poor repeaters. Thus, it is unknown whether there are children with low scores on nonword and sentence repetition who have memory but not language deficits, or vice versa. A detailed examination of
language and memory skills in relation to nonword repetition and sentence recall will help to determine whether these tasks are clinical markers of memory or language impairments, or both.

Determining the gold standard to be used for the purpose of identifying deficits is a complex yet crucial task. This is somewhat easier in the case of language than of memory impairment, however. Although individual clinicians and researchers may raise issues with any available test battery, most would agree that the CELF-4 (Semel et al., 2003) provides an acceptable measure of language deficit such that the test is widely used in clinical practice and research studies to identify children with SLI (e.g., Conti-Ramsden et al., 2001; O’Brien, Zhang, Nishimura, Tomblin, & Murray, 2003; Webster et al., 2006). The case for memory impairments is more difficult. Not only have very few studies examined specific developmental memory deficits (e.g., Hanten & Martin, 2001), but test batteries for memory skills appropriate for children are used less widely or have been published only recently (e.g., Alloway, 2007; Wilson, Cockburn, & Baddeley, 2003). In the present study, we chose to use a test battery of short-term and working memory as the reference standard for memory deficit for several reasons. First, short-term memory referring to the brief retention of information (J. Brown, 1958) without demands for additional cognitive processing (e.g., digit recall) has been linked to our clinical markers of interest, nonword repetition and sentence recall. Second, working memory (Baddeley & Hitch, 1974)—the ability to both briefly retain and process information (e.g., follow an instruction)—has been closely linked to academic learning (e.g., Gathercole, Alloway, Willis, & Adams, 2005; Swanson, 2004), indicating that it may be important in early skill development. Third, both verbal and visuospatial measures of short-term and working memory exist, providing (at least some) memory tests independent of language skill.

According to many theoretical models, working memory subsumes short-term memory and includes a central executive system that directs attention, coordinates cognitive processing, and suppresses irrelevant information (e.g., Baddeley, 1986; Baddeley & Hitch, 1974; Cowan, 1995, 2001; Engle, Kane, & Tuholski, 1999). Domain specificity within short-term memory is recognized within working memory models as distinct short-term memory stores (e.g., Bayliss, Jarrold, Gunn, & Baddeley, 2003; Kane et al., 2004) or a unitary store with interference occurring between same-modality representations (Cowan, 2001). Evidence of deficits in phonological but not visuospatial short-term memory tasks for SLI groups provides support for domain specificity within short-term memory (Archibald & Gathercole, 2006a). The processing component of working memory is typically considered a domain-general resource (e.g., Bayliss et al., 2003; Engle et al., 1999), although this tenet is by no means universally held (e.g., Caplan & Waters, 1999).

In the present study, we aimed to assess short-term and working memory for both verbal and visuospatial information and used the Automated Working Memory Assessment (AWMA; Alloway, 2007) as the gold standard for this purpose. This test battery is standardized for children aged 4–18 years and incorporates three different subtests to calculate each of four composite scores: phonological short-term memory, visuospatial short-term memory, verbal working memory, and visuospatial working memory. Short-term and working memory tasks are distinguished by whether they involve brief retention only (e.g., phonological: digit recall; visuospatial: dots on a grid) or additionally require information processing (e.g., verbal: judge the veracity of a sentence and recall the final word; visuospatial: judge the orientation of a shape and recall the direction it faced). On the basis of individual test results, domain-specific impairments in short-term memory could be identified in the present study. For working memory deficits, however, we required evidence of a domain-general deficit as reflected by poor scores on both the verbal and visuospatial working memory composites. Indications of a general impairment in working memory are consistent with the preponderance of evidence supporting a domain-general view of the processing resources in working memory (e.g., Alloway, Gathercole, & Pickering, 2006; Bayliss et al., 2003; Engle et al., 1999). In addition, the presence of both verbal and visuospatial working memory deficits would rule out the argument that low scores in one domain of working memory were secondary to a same-domain short-term memory impairment.

Overview of the Present Study

In the present study, we investigated the relationship between nonword repetition, sentence recall, language, short-term and working memory, and nonverbal cognition in a group of 400 children aged 5–9 years. One purpose was to examine the utility of nonword repetition and sentence recall as clinical markers of language and/or working memory impairments. The extent to which poor nonword and sentence repetition was sensitive and specific to language or working memory impairments was evaluated. A second aim of the study was to explore the occurrence of SLI, specific working memory impairment, and combined language and working memory deficits in a group of children identified as poor repeaters. In particular, we were interested in whether working memory deficits occurred independent of language impairment and whether nonword repetition and sentence recall acted as similar or better clinical markers of poor working memory.

902 Journal of Speech, Language, and Hearing Research • Vol. 52 • 899–914 • August 2009
Method

Participants

A total of nine schools in the southwest region of Ontario, Canada were recruited to the study. All children in senior kindergarten through Grade 3 in each of the schools were invited to participate. Approximately 1,255 consent forms were distributed, of which 412 were returned and signed by parents. Twelve children were unable to participate in the study because they were either absent from school on the screening day or they were too young (junior rather than senior kindergarten students). Thus, a total of 400 children (203 boys) ranging in age from 5.3 to 9.42 years participated in the study. By parent report, approximately 87% were right-handed, 94% spoke English as their first language, and 3% were learning a second language other than English. None of the children were diagnosed with hearing impairment or developmental or neurological disorders. Demographic statistics for all participants by age group are provided in Table 1.

Procedure

All participants completed a 5-min screening measure in a quiet room at their school, consisting of two tasks described below.

Nonword repetition. Items were presented from the NRT (Dollaghan & Campbell, 1998) and consisted of 16 nonwords, four stimuli each containing one, two, three, and four syllables. The nonwords were constructed from a limited set of phonemes (11 consonants, 9 vowels) excluding late-developing sounds. The nonwords followed an alternating consonant–vowel structure, and none of the syllables corresponded to English lexical items. Only tense vowels were used, and therefore the stress patterns of the nonwords were unlike typical English words in that they had no weak syllables. A detailed description of the criteria guiding the development of the NRT is provided in Dollaghan and Campbell (1998). The nonwords were presented auditorily via a digital audio recording of an adult female speaker following the phonetic transcription and pronunciation guidelines described by Dollaghan and Campbell (1998). For this study, the nonwords were presented in fixed random order across all nonword lengths rather than in randomized groups from shortest to longest, as in the Dollaghan and Campbell study. In addition, item-rather than phoneme-level scoring was used, with each nonword judged online as correct or incorrect by a trained research assistant. It must be acknowledged that although clinical utility is greater for online item-level than for phoneme-level scoring, some of the ability to discriminate differences in performance may be lost especially for a task with only 16 items.

Sentence recall. The sentences were taken from Redmond (2003) and consisted of 16 sentences each composed of 10 words (10–14 syllables), with an equal number of active and passive sentences. The Redmond study has previously shown that these sentences reliably differentiate SLI and typically developing groups while ensuring that performance of the typically developing group was not at ceiling. The sentences were presented in fixed order via a digital audio recording of an adult female speaker (rather than via live voice as in the Redmond study) immediately following the nonword repetition task. Sentences were scored online by the research assistant with either a 2 (correct), 1 (three or fewer errors), or 0 (more than four errors or no response).

Table 1 displays the descriptive statistics for the screening scores across age groups. There was a developmental trend for both tasks in one-way analyses of variance (ANOVAs) with Bonferroni correction, nonword repetition, $F(3, 396) = 13.529, p < .001, \eta^2 = .09$, and sentence recall, $F(3, 396) = 32.304, p < .001, \eta^2 = .20$, with significant improvements in nonword repetition between nonadjacent age groups ($p \leq .001$, all nonadjacent pairings) and between all age groups except age groups 7 and 8 for sentence recall ($p < .001$, all pairings except 8 vs. 9 years, $p < .05$ and 7 vs. 8, $p > .05$). In order to evaluate the accuracy with which the screening tasks
identified children with language and/or memory impairments, it would be ideal to evaluate in detail the skills of all of the participants; however, time and economical constraints precluded this. Thus, focus was directed toward two groups, those who scored either poorly or in the average range on the screening. To do this, percentile scores for each chronological age group were calculated on the basis of the present data. Participants were defined as low scorers if they scored below the 15th percentile for their age on both nonword and sentence repetition or below the 10th percentile on either nonword or sentence repetition (\(n = 53\)). Average scorers were defined as those who scored above the 35th percentile for their age on both nonword and sentence repetition (\(n = 197\)).

A subset (\(n = 88\)) of the original 400 children completed more detailed assessments, including the low scorers (\(n = 52\); 22 boys; 29 girls; 1 individual left the school and could not be located for further testing) and a group of average scorers selected with the constraints of choosing 5 boys and 5 girls from each school grade level and from the same schools as those of the low scorers. Of the 40 average scorers selected, 2 could not be tested (1 left the school, and 1 withdrew from the study). Of the remaining 38 average scorers tested (18 boys), raw score files for the memory measures for 2 of the boys were lost due to technological difficulties. Within 6 months of the original screening, the low- and average-scoring groups were tested in a quiet room in their school over two individual sessions occurring within 1 week of each other and each lasting approximately 40 min. Each child completed tests tapping the following areas in the order indicated, with the first session administered counterbalanced across participants: Session 1, receptive and expressive language, nonverbal intelligence; Session 2, verbal and visuospatial short-term and working memory.

**Receptive and expressive language.** The core subtests of the CELF-4 (Semel et al., 2003) were administered as the reference standard for language skills. The core subtests consisted of Concepts and Following Directions, Recalling Sentences, Formulating Sentences and, depending on the age of the child, Word Knowledge (under 9 years) or Word Classes (9 years or older). A composite core language score (CLS) was calculated for each participant on the basis of the test norms.

**Nonverbal intelligence.** The Test of Nonverbal Intelligence–3 (TONI-3; L. Brown, Sherbenou, & Johnsen, 1997) was administered as a measure of general nonverbal cognitive ability. The test involves completing a visual pattern. This test was included in order to describe the nonverbal abilities of the groups investigated in the present study. No inclusionary or exclusionary decisions were based on performance on the TONI-3. Although a nonverbal performance criterion continues to be used in the majority of studies investigating SLI, an exclusionary criterion of IQ scores below 70 has been adopted in several recent studies (e.g., Conti-Ramsden et al., 2001; Thordardottir, 2008) in light of equivocal findings between language-impaired groups with either typical or atypical nonverbal scores (e.g., Ellis Weismer et al., 2000; Tomblin & Zhang, 1999). Only 1 child in the present study scored below 70 on the TONI-3, and that child performed in the average range on all remaining tests.

**Verbal and visuospatial short-term and working memory.** The AWMA (Alloway, 2007), a computerized assessment tool including subtests tapping phonological short-term memory (digit recall, nonword recall, word recall), visuospatial short-term memory (dot matrix, block recall, mazes memory), verbal working memory (listening recall, counting recall, backwards digit recall), and visuospatial working memory (Mr. X, odd-one-out, spatial span), was administered. The short-term memory tasks required only the immediate recall of information that is phonological (digits, nonwords, words) or visuospatial (location, path). For example, the word (and nonword) recall tasks required repetition of common one-syllable words (or one-syllable nonwords) presented at a rate of one (non)word per second, whereas the dot matrix (and block recall) tasks required the repetition of the location of dots (blocks) on a matrix (or board with nine blocks). The mazes memory test required the repetition of a path through a maze. The working memory tasks, however, required cognitive processing of verbal (sentence veracity, counting, reversing order) or visuospatial (mental rotation, identifying differences) information as well as retention of same-modality material (i.e., phonological: words, digits; visuospatial: location). For example, the counting recall task required the counting of circles while ignoring triangles on each presented page, then the recall of the tallies, whereas the odd-one-out test required the identification of a different shape from three, then the recall of the location of the odd shapes. The Mr. X (and spatial span) tasks required judgment of the orientation of each Mr. X (or shape), then recalling where each Mr. X (or red dot on shape) had been pointing. Items were administered in a span procedure beginning at the easiest level, moving to the next level when four out of six items were completed correctly and discontinued when three errors occurred at one level. Composite scores for phonological short-term memory, visuospatial short-term memory, verbal working memory, and visuospatial working memory were calculated for each participant on the basis of test norms.

**Results**

**Screening results.** Table 2 summarizes the number of children who met the criteria for low and average...
scorers on the screening measures. Almost half the original group (49%) were average scorers, scoring above the 35th percentile for their age on both nonword repetition and sentence recall. In the low-scorers group of 53, 36% scored below the 15th percentile for their age on both nonword repetition and sentence recall, and an additional 64% scored below the 10th percentile on either task. For the purposes of the sensitivity and specificity analysis to follow, low scorers were considered to have “tested positive” and average scorers “negative” for the disorder. A substantial group (38%) could not be classified according to our definitions (outlined above) as either low or average scorers. It should be noted that the cut-off scores adopted for the screening measure were set arbitrarily in order to focus the more detailed assessments on low and average scorers, and thus the unclassified designation is simply a placeholder. We made no predictive hypothesis about this group and did not investigate this group further in the study.

Language and working memory profiles of low and average scorers. Descriptive statistics are shown in Table 3 for the standardized tests and subtests of language, nonverbal intelligence, and short-term and working memory completed with almost all of the low scorers (98%) and a representative sample of the average scorers. Consider first the areas tested as reflected by the nonverbal intelligence measure and the composite scores for language and memory. These scores were significantly higher for the average than for the low scorers for each participant. A deficit was defined as a score of more than 1 SD below the mean of the standardization sample corresponding to an effect size of 1.0, conventionally considered to be large in magnitude (Cohen, 1988). Thus, individuals with a composite language score on the CELF-4 below 86 were considered to have a language impairment. In order to have a working memory impairment, a child must have received a score below 86 on both the verbal and the visuospatial working memory composites. And conversely, a score above 86 on both

| Table 2. Number of low scorers, average scorers, and unclassified participants on screening measure. |
|---|---|---|---|---|---|---|
| Age group | NWRep & SR < 15th percentile | NWRep < 10th percentile | SR < 10th percentile | NWRep & SR > 35th percentile | Unclassified |
| 6 | 8 | 3 | 5 | 50 | 42 |
| 7 | 4 | 5 | 4 | 57 | 32 |
| 8 | 1 | 4 | 8 | 44 | 41 |
| 9 | 6 | 1 | 4 | 46 | 35 |
| Total | 19 | 13 | 21 | 197 | 150 |

Note. NWRep = nonword repetition; SR = sentence recall.
composites was required in order for us to consider the child not to have a working memory impairment. According to this definition, children with deficit scores in either verbal or visuospatial working memory but not both were unclassified in terms of their working memory status. These definitions gave rise to the following possible language and working memory profiles: (a) working memory and language impairment, (b) language impairment only, (c) language impairment and unclassified working memory, (d) working memory impairment only, (e) unclassified working memory only, and (f) no language or working memory impairments. In the following sections, we examine the presence of these profiles in our data, compare the standardized test performance of the groups, and analyze the sensitivity and specificity of the screening measure to language and/or working memory deficits, as described above.

All of the possible language and working memory profiles occurred in our sample, the distribution of which is shown in Figure 1a, b, and c for the average and low scorers and combined group, respectively. The majority of average scorers (61%) had no language or working memory impairments, although 31% fell into the unclassified working memory category. All of the children identified with language impairments regardless of working memory status (n = 28) were from the low-scorers group. Across the full group, the number of children with specific impairments of either language or working memory was equivalent, and just over half that of those with combined language and working memory deficits. The core language subtests of the CELF-4 include only one receptive measure. Of the children with language impairment (n = 28), 4 had scaled scores within 1 SD below the mean on this subtest, 3 in the combined language and working memory impairment group, and 1 in the language impairment-only groups.

Table 4 provides a comparison of the standardized test performance for the language and working memory groups described above. Some of the scores were necessarily low or high due to the definitions set for each group (i.e., low composite language scores for the three groups with language impairments, and low scores in language and working memory for the combined language and working memory impairment group). Three

### Table 3. Descriptive statistics for standardized tests (M = 100, SD = 15) of language, cognition, short-term memory (STM), and working memory (WM) for the low and average scorers.

<table>
<thead>
<tr>
<th>Area tested</th>
<th>Test</th>
<th>Composite Score/Subtest</th>
<th>Low scorers (n = 52) M (SD)</th>
<th>Average scorers (n = 38) M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonverbal intelligence</td>
<td>TONI-3</td>
<td>N/A</td>
<td>95.92 (13.21)</td>
<td>108.64 (14.72)</td>
</tr>
<tr>
<td>Language</td>
<td>CELF-4</td>
<td>CLS</td>
<td>85.33 (15.97)</td>
<td>106.05 (11.60)</td>
</tr>
<tr>
<td>Concepts and directions</td>
<td></td>
<td>7.78 (3.53)</td>
<td>10.81 (2.69)</td>
<td></td>
</tr>
<tr>
<td>Recalling sentences</td>
<td></td>
<td>6.42 (3.02)</td>
<td>11.55 (2.02)</td>
<td></td>
</tr>
<tr>
<td>Formulating sentences</td>
<td></td>
<td>8.50 (3.33)</td>
<td>11.53 (2.31)</td>
<td></td>
</tr>
<tr>
<td>Word structure</td>
<td></td>
<td>7.26 (2.97)</td>
<td>10.19 (2.85)</td>
<td></td>
</tr>
<tr>
<td>Word classes</td>
<td></td>
<td>9.00 (2.97)</td>
<td>10.57 (3.99)</td>
<td></td>
</tr>
<tr>
<td>Phonological STM</td>
<td>AWMA</td>
<td>Composite</td>
<td>86.12 (13.44)</td>
<td>105.83 (12.50)</td>
</tr>
<tr>
<td>Verbal WM</td>
<td>AWMA</td>
<td>Composite</td>
<td>87.13 (17.74)</td>
<td>104.50 (17.83)</td>
</tr>
<tr>
<td>Visuospatial STM</td>
<td>AWMA</td>
<td>Composite</td>
<td>93.54 (21.77)</td>
<td>109.03 (25.92)</td>
</tr>
<tr>
<td>Visuospatial WM</td>
<td>AWMA</td>
<td>Composite</td>
<td>88.90 (16.53)</td>
<td>101.28 (21.11)</td>
</tr>
</tbody>
</table>

Note. TONI-3 = Test of Nonverbal Intelligence–3; CLS = Composite Language Score; CELF-4 = Clinical Evaluation of Language Fundamentals–4; AWMA = Automated Working Memory Assessment.

*Scaled scores (M = 10, SD = 3). n = 41. *n = 29. *n = 29. *n = 7.
groups were characterized as having a language impairment with or without classifiable working memory deficits. A one-way ANOVA comparing the language level of these three groups marginally failed to reach significance, \(F(2, 25) = 3.247, p = .056, \eta_p^2 = .21\). Only the comparison between the language-impaired-only and the language-impaired-with-unclassifiable working memory deficits reached significance in pairwise \(t\) tests, \(t(14) = 2.605, p = .021\). Thus, the language skills of the combined working memory and language impairment group did not differ from those of the language-impaired group with unclassifiable working memory, although the latter group had relative strengths in visuospatial short-term and working memory.

Consider also the nonverbal intelligence and short-term memory scores for all profile groups shown in Table 4. The three groups with language impairment had lower mean scores in nonverbal intelligence and phonological short-term memory than the groups without language impairment. This difference was found to be significant when groups with or without language impairment (collapsed across working memory profiles) were compared in separate one-way ANOVAs on nonverbal intelligence, \(F(1, 88) = 27.277, p < .001, \eta_p^2 = .24\), and phonological short-term memory, \(F(1, 86) = 26.427, p < .001, \eta_p^2 = .24\). Verbal working memory deficits, however, were not always associated with a language impairment, as evidenced by the age-appropriate verbal working memory skills of the language impairment–only group. Nevertheless, it is interesting that the two groups with unclassified working memory deficits with or without language impairment were distinguished by the presence of verbal working memory deficits in the language-impaired group and visuospatial working memory deficits in the group without language impairment. Further inspection, however, revealed that one third \((n = 6)\) of those in the unclassified group had verbal (and not visuospatial) working memory deficits. The mean composite language score for this group of 6 was 107.50 \((SD = 10.75)\), with a verbal working memory average score of 80.33 \((SD = 4.84)\).

In terms of working memory impairments, there was evidence of a general short-term memory deficit across the verbal/phonological and visuospatial domains in the combined working memory and language impairment group (see Table 4). The deficits in the working memory impairment–only group did not extend to short-term memory. In addition, the distribution of short-term and working memory deficits across all participants who scored more than 1 \(SD\) below the standardized mean on at least one memory composite regardless of language status \((n = 56)\) was inspected further. This group included 11 children in the no-language or working memory impairments group \((n = 2);\) low scorers: \(n = 9)\) who had deficits limited to one \((n = 6; 11\%)\) or both
Table 4. Descriptive statistics for standardized tests (M = 100, SD = 1.5) of language, nonverbal intelligence, short-term memory (STM), and working memory (WM) for the language and working memory profile groups.

<table>
<thead>
<tr>
<th>Area tested</th>
<th>WM and LI</th>
<th>LI only</th>
<th>LI and unclass. WM</th>
<th>WMI only</th>
<th>Unclass. WM</th>
<th>No deficits</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n = 12)</td>
<td>(n = 7)</td>
<td>(n = 8)</td>
<td>(n = 7)</td>
<td>(n = 18)</td>
<td>(n = 36)</td>
<td></td>
</tr>
<tr>
<td>Nonverbal intelligence</td>
<td>90.00 (5.66)</td>
<td>92.57 (16.30)</td>
<td>88.38 (7.93)</td>
<td>100.43 (10.23)</td>
<td>100.50 (9.64)</td>
<td>109.28 (16.08)</td>
</tr>
<tr>
<td>Language</td>
<td>73.42 (9.65)</td>
<td>79.14 (6.78)</td>
<td>66.25 (10.49)</td>
<td>95.86 (8.24)</td>
<td>101.67 (11.08)</td>
<td>106.11 (11.17)</td>
</tr>
<tr>
<td>Phonological STM</td>
<td>84.00 (10.13)</td>
<td>80.71 (9.84)</td>
<td>82.12 (17.72)</td>
<td>92.57 (12.69)</td>
<td>99.61 (14.57)</td>
<td>101.64 (16.40)</td>
</tr>
<tr>
<td>Verbal WM</td>
<td>70.25 (11.51)</td>
<td>99.43 (11.73)</td>
<td>75.88 (9.37)</td>
<td>76.57 (8.83)</td>
<td>91.44 (9.66)</td>
<td>110.14 (14.56)</td>
</tr>
<tr>
<td>Visuospatial STM</td>
<td>80.92 (15.20)</td>
<td>98.14 (19.85)</td>
<td>91.62 (13.05)</td>
<td>98.86 (13.81)</td>
<td>102.94 (13.46)</td>
<td>107.03 (31.87)</td>
</tr>
<tr>
<td>Visuospatial WM</td>
<td>72.42 (9.69)</td>
<td>98.71 (9.34)</td>
<td>94.38 (8.83)</td>
<td>77.14 (5.46)</td>
<td>82.33 (15.71)</td>
<td>109.22 (14.91)</td>
</tr>
</tbody>
</table>

Note. LI = language impairment; unclass. = unclassified; WMI = working memory impairment.

(n = 5; 9%) short-term memory domains. A small number had domain-specific deficits in short-term and working memory for phonological (n = 4; 7%) or visuospatial (n = 1; 2%) information. The co-occurrence of working memory and language impairment is incorporated in Table 4. For children whose memory deficits were limited to short-term memory only, language impairment was present in 5/9 with phonological (with or without visuospatial deficits) but none (0/4) with visuospatial-only deficits.

Sensitivity and specificity of screening measure. Table 5 compares the sensitivity, specificity, and LRs for assessment outcomes, as defined above. In the first analysis, our criterion for defining low and average scorers on the screening measure was used to classify screening results as positive or negative, respectively. The sensitivity of this criterion was high-to-very high (84%–100%) for language, working memory, or combined language and working memory deficits. The corresponding LRs of 2.4, 1.6, and 1.9, respectively, indicated that a child who scored positive on the screening was more likely to have an impairment. Specificity, however, was low (47%–50%), suggesting that several children who scored positive on the screening were not necessarily found to have an impairment in the full assessment.

We next calculated the same values for each screening task separately, considering a positive screening result as any score below the 15th percentile for each measure. It should be noted that in this case, the present “low scorers” represent a partial set of the original 400 who would have met this criterion. For nonword repetition, specificity was improved (60%–63%), but sensitivity was considerably weaker (40%–46%), and LRs were near equivocal. The results were better for the sentence recall task, however. Sensitivity (96%–100%) was high for language impairment and for combined working memory and language impairments, although specificity was greater for the former (76% vs. 65%, respectively). Performance above the 15th percentile on the sentence recall portion of the screening effectively ruled out language or combined working memory and language impairments. Conversely, a child who scored positive on the screening according to this criterion was four times more likely to be in the language-impaired group. Indeed, LRs for sentence recall below the 15th percentile were two or more for all impairment groups.

We also determined whether an even more severe cut-off of below the 10th percentile on one of the screening tasks would improve the results. There was one instance of notable improvement for sentence recall: Sensitivity and specificity were both above 80% for identifying language impairment, with positive results over 8.5 times more likely to come from a child with the impairment. Corresponding values for identifying children with combined working memory and language impairments approached similar levels.

Discussion

Recent efforts to identify clinical markers of SLI have suggested nonword repetition and sentence recall as potential candidates. Although both nonword repetition and sentence recall may tap memory and linguistic skills, even population-based studies have included reference standard tests of language and not memory abilities (e.g., Ellis Weismer et al., 2000). The present study aimed to assess the potential of nonword repetition and sentence recall as clinical markers of developmental language and/or working memory impairments. A group of 400 school-age children completed a screening measure of nonword repetition and sentence recall. The majority of the low scorers and a representative sample of the average scorers on this screening completed tests of language, short-term and working memory, and nonverbal intelligence. An approximately equal number of
children were identified with either specific language or specific working memory impairments, and a group about twice as large with mixed language and working memory impairments. The language-impaired groups had phonological short-term memory impairments also but not necessarily verbal working memory deficits. Short-term memory impairments co-occurred with working memory deficits in the group with combined language and working memory but not specific working memory impairments.

Sensitivity of a screening measure, including both nonword repetition and sentence recall, or sentence recall alone, was high for specific language, specific working memory, or combined language and working memory impairments. Specificity, however, was lower. Levels of sensitivity and specificity above 80% and the highest likelihood ratio were found for sentence recall performance below the 10th percentile. Nonword repetition on its own was less sensitive to impairments, although specificity values were closer to those for sentence recall.

**Language and memory groupings.** The present work is the first to identify in the same study children with specific impairments in either working memory or language, or combined language and working memory impairments. Consistent with findings that the language performance of children with SLI is not qualitatively distinct from that of typically developing children (Dollaghan, 2004), we identified impairments on the basis of poor standardized test performance and found approximately equal numbers of specific language and working memory impairments.

### Table 5. Sensitivity and specificity of nonword repetition and sentence recall from screening measure.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Screening</th>
<th>+</th>
<th>-</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impairment</strong></td>
<td><strong>Language</strong></td>
<td>26</td>
<td>0</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td><strong>WM</strong></td>
<td>16</td>
<td>3</td>
<td>36</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>12</td>
<td>0</td>
<td>40</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td><strong>Any memory</strong></td>
<td>30</td>
<td>14</td>
<td>22</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

**Criteria: Both NWRep and sentence recall below 15th percentile or one below 10th**

<table>
<thead>
<tr>
<th>Impairment</th>
<th>Sensitivity(a)</th>
<th>Specificity(b)</th>
<th>Likelihood</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>26/26=100%</td>
<td>36/62=58.1%</td>
<td>2.4</td>
<td>.62</td>
</tr>
<tr>
<td>WM</td>
<td>12/14</td>
<td>26/22</td>
<td>1.9</td>
<td>.66</td>
</tr>
<tr>
<td>Combined</td>
<td>18/18</td>
<td>26/26</td>
<td>1.4</td>
<td>.58</td>
</tr>
</tbody>
</table>

**Criteria: NWRep below 15th percentile**

<table>
<thead>
<tr>
<th>Impairment</th>
<th>Sensitivity(a)</th>
<th>Specificity(b)</th>
<th>Likelihood</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>12/14</td>
<td>26/22</td>
<td>1.9</td>
<td>.66</td>
</tr>
<tr>
<td>WM</td>
<td>9/10</td>
<td>26/22</td>
<td>1.9</td>
<td>.66</td>
</tr>
<tr>
<td>Combined</td>
<td>5/7</td>
<td>30/46</td>
<td>1.1</td>
<td>.51</td>
</tr>
<tr>
<td>Any memory</td>
<td>18/26</td>
<td>17/27</td>
<td>1.1</td>
<td>.51</td>
</tr>
</tbody>
</table>

**Criteria: NWRep below 10th percentile**

<table>
<thead>
<tr>
<th>Impairment</th>
<th>Sensitivity(a)</th>
<th>Specificity(b)</th>
<th>Likelihood</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>5/14</td>
<td>21/41</td>
<td>1.2</td>
<td>.55</td>
</tr>
<tr>
<td>WM</td>
<td>6/20</td>
<td>20/42</td>
<td>1.3</td>
<td>.56</td>
</tr>
<tr>
<td>Combined</td>
<td>2/10</td>
<td>24/52</td>
<td>1.1</td>
<td>.51</td>
</tr>
<tr>
<td>Any memory</td>
<td>10/34</td>
<td>16/28</td>
<td>1.1</td>
<td>.51</td>
</tr>
</tbody>
</table>

**Criteria: Sentence recall below 15th percentile**

<table>
<thead>
<tr>
<th>Impairment</th>
<th>Sensitivity(a)</th>
<th>Specificity(b)</th>
<th>Likelihood</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>25/15</td>
<td>47/96.2</td>
<td>4.0</td>
<td>.80</td>
</tr>
<tr>
<td>WM</td>
<td>16/3</td>
<td>23/46</td>
<td>2.5</td>
<td>.71</td>
</tr>
<tr>
<td>Combined</td>
<td>12/0</td>
<td>27/49</td>
<td>2.8</td>
<td>.74</td>
</tr>
<tr>
<td>Any memory</td>
<td>26/18</td>
<td>13/31</td>
<td>2.0</td>
<td>.67</td>
</tr>
</tbody>
</table>

**Criteria: Sentence recall below 10th percentile**

<table>
<thead>
<tr>
<th>Impairment</th>
<th>Sensitivity(a)</th>
<th>Specificity(b)</th>
<th>Likelihood</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>22/4</td>
<td>6/56</td>
<td>8.7</td>
<td>.90</td>
</tr>
<tr>
<td>WM</td>
<td>11/8</td>
<td>17/52</td>
<td>2.3</td>
<td>.70</td>
</tr>
<tr>
<td>Combined</td>
<td>10/2</td>
<td>18/58</td>
<td>3.5</td>
<td>.78</td>
</tr>
<tr>
<td>Any memory</td>
<td>20/24</td>
<td>8/28</td>
<td>2.0</td>
<td>.67</td>
</tr>
</tbody>
</table>

**Note.** + = positive result (presence of impairment); – = negative result (absence of impairment); Combined = combined working memory and language impairment; Any memory = deficit score on any memory composite.

aTotal number with impairment. bTotal number without impairment.
skills. It may be that both language and working memory skills are normally distributed in the pediatric population, with the tails of each distribution representing children with specific impairments, and some overlap in the distributions forming the group with combined language and working memory deficits. One possible argument to explain the groupings in the present study could be that children with milder impairments received low scores on only a few tests and ended up in the specific groups, whereas those with more severe impairments scored poorly on the majority of tests, ending up in the combined deficit group. The present findings, however, do not support this conclusion. The children with SLI and those with combined working memory and language impairments had equivalent language deficits. And those with language and unclassified working memory impairments who would be expected to have intermediate scores on the basis of this prediction had the most severe language deficits overall.

The present results suggest the existence of specific working memory impairment (SWMI) as a discrete category of impairment in children. Although the study cannot speak strongly to its underlying nature, it appears that children with SWMI have intact basic linguistic skills in terms of grammatical expression, sentence formulation, and following directions. Such results, however, do not rule out more subtle communication disorders that might be more apparent following testing of a broader set of abilities. It may be that more complex skills, such as narration and discourse, place demands that exceed the working memory limitations of this group, for example. As well, several studies reporting strong associations between working memory deficits and learning disabilities (e.g., Gathercole et al., 2005; Swanson, 2004) would lead to the prediction that children with SWMI would face academic challenges. However, it is not known whether the children with SWMI in the present study were identified by their parents, teachers, or peers as different in terms of communication, learning, or other skills. We are presently engaged in further studies of children with SWMI to address these questions.

The present findings provide some insight into the possible nature of SLI. One finding consistent across all of the children with SLI in the present and many previous studies (e.g., Archibald & Gathercole, 2006b; Bishop et al., 1996; Conti-Ramsden et al., 2001) was the co-occurrence of phonological short-term memory deficits. The consistency of this finding across three groups of children with language impairment in the present study is impressive. It must be acknowledged, however, that the children with language impairment in the present study were selected on the basis of their low performance on two screening measures with high short-term memory demands: nonword repetition and sentence recall. The low scorers on the screening did perform particularly poorly on the standardized CELF-4 sub-test most closely resembling one of the screening tasks, recalling sentences. It is possible, then, that this selection scheme resulted in a higher degree of consistency in the phonological short-term memory impairments across the language-impaired groups in the present study. Nevertheless, there was considerable variability in phonological short-term memory in the language-impaired groups in the present work, suggesting that phonological short-term memory deficits are not always associated with language impairment.

Groups of children with language impairment in the present study were distinguished by whether or not working memory deficits co-occurred. The identification of separate groups specifically impaired in language or working memory, however, suggests an additive rather than a unidirectional, causal pathway. That is, the presence of working memory impairments do not always cause language learning deficits, and vice versa. It may be that important qualitative differences distinguish children with SLI and those with combined language and working memory impairments, and these differences could help to explain some of the heterogeneity reported among children with developmental language impairments.

A substantial group in the present study had working memory impairments that did not cross domains. It may be that a limited capacity in the ability to both briefly store and process information was not constraining performance in these groups. The unclassified working memory impairments were predominantly verbal in nature for those children with coexisting language deficits. These children had the most severely impaired language skills overall, suggesting that their performance on the verbal working memory tasks may have been constrained by their poor language skills. For those without language impairment, unclassified working memory deficits were characteristically visuospatial in nature. One possible explanation of these findings relates to the unimpaired group. The children in the present study completed several standardized tests, increasing the likelihood of poor performance on at least one subtest. The criterion requiring age-appropriate performance on the language and working memory measures in order to be considered unimpaired may have selected an “above average” group, as reflected by their standard scores between 106 and 110 on most measures. Thus, the unclassified working memory impairment group may have been a more average group, as reflected by their standard scores between 91 and 102 on most measures. A second possibility is that these children had specific difficulty with visuospatial information, which in turn constrained performance on the visuospatial working memory tasks. Further study is needed to examine the nature and impact of such deficits.
Interestingly, lower cognitive skills were a characteristic of all of the groups with language impairment in the present study despite being broadly within age expectations. These findings extend previous observations of this nature (e.g., Botting, 2005; Leonard, 1998). It is unclear why lower cognitive skills are associated specifically with language and not working memory deficits. Another common distinction in the SLI literature is between children whose deficits are limited only to expression (E-SLI) and those with both receptive and expressive deficits (R-SLI; e.g., Evans, 1996). It was difficult to make this distinction in the present study, as the core language subtests of the CELF-4 (Semel et al., 2001) included only one receptive measure on which the majority of children with language impairments scored in the deficit range. Nevertheless, the 4 children with possible E-SLI in the present study were not uniformly grouped. The presence of 3 of them in the combined working memory and language-impaired group provides further support for the notion that the groupings in the present work were not related to severity of deficits.

Nonword repetition and sentence recall as clinical markers of language and memory. In the present study, we used a 5-min screening measure administered by trained research assistants, but without expertise in communication disorders, and successfully identified all of the children with language impairment or combined working memory and language impairment, and 90% of those with working memory deficits only. The main problem was the reduced specificity, which is associated with the low positive predictive value reported for many similar screenings (Law, Boyle, Harris, Harkness, & Nye, 2000). Nevertheless, the low demands that the present screening technique places on training of administrators and testing times make it of considerable clinical interest as a measure for identifying children in need of further assessment. It should be noted as well that this screening measure may be sensitive to only a limited range of language differences, and thus average performance on the screening would not rule out a communication disorder.

As in many previous findings, poor nonword repetition characterized the impaired groups (e.g., Dollaghan & Campbell, 1998; Gathercole, Tiffany, Briscoe, Thorn, & the ALSPAC Team, 2005). More specific comparisons of the present nonword repetition findings to previous research, however, are limited. Past studies have used either Gathercole and Baddeley’s (1996) CNRep (e.g., Bishop et al., 1996; Conti-Ramsden, 2003; Conti-Ramsden et al., 2001) or phoneme-level scoring of the NRT (Dollaghan & Campbell, 1998; Ellis Weismen et al., 2000). In the present work, we administered the NRT and scored at the item level. It may be that item-level scoring over 16 items is not sufficiently sensitive to individual differences, as reflected by the lack of developmental differences across all age groups in the present study and the lower sensitivity and specificity of the measure. Nevertheless, unlike item-level scoring as used here, phoneme-level scoring lacks practicality as a screening instrument to be administered in the field. Thus, item-level scoring warrants further investigation.

Poor sentence recall, however, provided a useful clinical marker of both SLI and a combined working memory and language impairment. Performance below the 10th percentile was 8.7 times more likely to come from a child with a language impairment. Although sentence recall may provide a convenient screening method for identifying language impairment (Conti-Ramsden et al., 2001; Stokes et al., 2006), its theoretical contribution is less clear. It can be no surprise that children with language or working memory impairments will do poorly on the task due to its reliance on both language and memory abilities. Sentence recall necessarily taps a variety of linguistic and memory skills, and thus individuals may do poorly on the task for different reasons. There were some indications that sentence recall was relatively better at identifying language than working memory impairments presumably due to its significant linguistic demands.

As with many screenings for communication disorders (Law et al., 2000), there were a high number of false positives in the present study. Researchers have suggested that those who test positive on an initial screen be tested again with a different screening measure for confirmation (Zenilman et al., 2003). Although the combined nonword repetition and sentence recall criteria were no better than sentence recall alone at identifying children with language impairments in the present study, it may be that a nonword repetition test with more items or scored at the phoneme level could provide a suitable confirmatory test.

Limitations of the present study. In the present study, we adopted an epidemiological approach, relying on standardized tests for classification purposes. In addition, our decision to categorize working memory abilities on the basis of cross-domain performance was theoretically motivated but had some necessary pitfalls. In particular, it involved having two “unclassified” groups into which at least some individuals must fall, probabilistically speaking. Quite apart from the interpretational challenge of these groups, it also led to small subgroup sizes even when drawing from a relatively large initial sample. The study also focused on children who had scored either poorly or in the average range on a screening measure, thus excluding a portion of the naturally occurring distribution. There was up to a 6-month gap between administrations of the screening and standardized tests, leading to the possibility that different rates of development reduced the association
between the two performances. Together, these limitations provide reasons for cautious interpretation of the present results. More generally, the finding of specific and mixed language and working memory impairments, although interesting, needs to be investigated in a larger study, including the full range of performance and with clinical validation and description.

Conclusions. We examined in the present study whether a brief screening instrument could identify children with probable language and working memory impairments. We evaluated the sensitivity and specificity of nonword repetition and sentence recall to language and working memory impairments in school-age children. Sentence recall, but not nonword repetition, scored at the item level provided a useful clinical marker of SLI and deficits in both language and working memory. In addition, this work is the first to identify in the same study groups of children with SLI, SWMI, and combined working memory and language impairments. More important, deficits in language and working memory did not always co-occur, suggesting an additive rather than a causal connection. Further investigations will investigate the distinct nature and communication profiles of SWMI and SLI.

Acknowledgments

This work was supported by a National Science and Engineering Research Council of Canada postdoctoral fellowship awarded to the first author and by a Canadian Institutes of Health Research Operating Grant awarded to the second author. The valuable assistance of participating school personnel and families is gratefully acknowledged.

References


Received May 14, 2008
Revision received November 2, 2008
Accepted December 22, 2008
DOI: 10.1044/1092-4388(2009/08-0099)

Contact author: Lisa Archibald, School of Communication Sciences and Disorders, Elborn College, University of Western Ontario, London, Ontario N6G 1H1, Canada.
E-mail: larchiba@uwo.ca.
On the Sensitivity and Specificity of Nonword Repetition and Sentence Recall to Language and Memory Impairments in Children

Lisa M. D. Archibald, and Marc F. Joanisse

*J Speech Lang Hear Res* 2009;52;899-914; *originally published online Apr 29, 2009*;
DOI: 10.1044/1092-4388(2009/08-0099)

The references for this article include 18 HighWire-hosted articles which you can access for free at: [http://jslhr.asha.org/cgi/content/full/52/4/899#BIBL](http://jslhr.asha.org/cgi/content/full/52/4/899#BIBL)

This article has been cited by 7 HighWire-hosted article(s) which you can access for free at: [http://jslhr.asha.org/cgi/content/full/52/4/899#otherarticles](http://jslhr.asha.org/cgi/content/full/52/4/899#otherarticles)

This information is current as of December 26, 2012

This article, along with updated information and services, is located on the World Wide Web at: [http://jslhr.asha.org/cgi/content/full/52/4/899](http://jslhr.asha.org/cgi/content/full/52/4/899)