Nonword Repetition: A Comparison of Tests

Lisa M. D. Archibald
Susan E. Gathercole
University of Durham,
Durham, United Kingdom

Purpose: This study compared performance of children on 2 tests of nonword repetition to investigate the factors that may contribute to the well-documented nonword repetition deficit in specific language impairment (SLI).

Method: Twelve children with SLI age 7 to 11 years, 12 age-matched control children, and 12 control children matched for language ability completed 2 tests of nonword repetition: the Children's Test of Nonword Repetition (CNRep) and the Nonword Repetition Test (NRT).

Results: The children with SLI performed significantly more poorly on both tests than typically developing children of the same age. The SLI group was impaired on the CNRep but not the NRT relative to younger children with similar language abilities when adjustments were made for differences in general cognitive ability. The children with SLI repeated the lengthiest nonwords and the nonwords containing consonant clusters significantly less accurately than the control groups.

Conclusion: The evidence suggests that the nonword repetition deficit in SLI may arise from a number of factors, including verbal short-term memory, lexical knowledge, and output processes.

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

Children with specific language impairment (SLI) experience particular difficulties in repeating multisyllabic nonwords such as /wʊgəˈlæmɪk/ or /nɔɪtaʊf/. This finding has led to widespread interest in both the cognitive processes tapped by nonword repetition and its diagnostic utility. The purpose of the present study was to compare the performance of a group of children with SLI on two of the most widely used tests of nonword repetition: the Children's Test of Nonword Repetition (CNRep; Gathercole & Baddeley, 1996) and the Nonword Repetition Test (NRT; Dollaghan & Campbell, 1998).

Nonword repetition deficits in SLI have been established in several research studies over the past 2 decades (e.g., Botting & Conti-Ramsden, 2001; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990; Gray, 2003; Kamhi & Catts, 1986; Kamhi, Catts, Maurer, Apel, & Gentry, 1988; Marton & Schwartz, 2003; Montgomery, 2004; Norbury, Bishop, & Briscoe, 2002; Sahlen, Reuterskiold-Wagner, Nettelbladt, & Radeborg, 1999; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). Difficulty repeating nonwords has been found in children with SLI relative to younger, typically developing children with matched language abilities (Edwards & Lahey, 1998; Gathercole & Baddeley, 1990; Montgomery, 1995), and in children with a history of SLI whose oral language is no longer distinguishable from age peers (Bishop, North, & Donlan, 1996; Botting & Conti-Ramsden, 2001). Multisyllabic nonword repetition has been found to be one task more impaired in children with

SLI than children with reading impairments, although these groups perform similarly on many other tasks (Kahmi & Catts, 1986; Kamhi et al., 1988).

Many studies have used one of two particular tests of nonword repetition. The CNRep (Gathercole & Baddeley, 1996) is widely used in the United Kingdom, and the NRT (Dollaghan & Campbell, 1998) has been used in many U.S.-based studies of SLI. Consider first findings from studies using the CNRep. Deficits on the CNRep in children with SLI have been found to have a strong genetic basis. In two twin studies, Bishop and colleagues demonstrated that the characteristic CNRep deficit in SLI is highly heritable and distinguishable from the auditory temporal processing difficulties that are also characteristic of the disorder (Bishop, Bishop, et al., 1999; Bishop, North, & Donlan, 1996). Recently, nonword repetition deficits in SLI assessed by a preliminary version of the CNRep (Gathercole, Willis, Emslie, & Baddeley, 1994) have been linked in particular with abnormalities of chromosome 16q (SLI Consortium, 2002, 2004). Based on the pattern of results from their twin study, Bishop et al. (1996) suggested that the CNRep provides an effective phenotypic marker of SLI. Conti-Ramsden and colleagues included CNRep in an evaluation of potential clinical markers of SLI in a group of 5-year-old children (Conti-Ramsden, 2003) and a group of 11-year-old children with a previous history of SLI (Conti-Ramsden, Botting, & Faragher, 2001). Results indicated that nonword repetition provided a useful clinical marker, although the more difficult task of sentence repetition was a more useful marker in the older age group.

The NRT has also been shown to be an excellent discriminator of children with language impairment. Dollaghan and Campbell (1998) reported that poor NRT performance was 25 times more likely to occur in a clinically referred group of children receiving language intervention than children with typical language development. Diagnostic accuracy of the NRT in this study surpassed that of the Spoken Language Quotient of the Test of Language Development—Primary, Second Edition (Newcomer & Hammill, 1988). Ellis Weismer et al. (2000) used the NRT to examine nonword repetition in a population-based sample of school-age children and reported that poor scores were 6.5 times more likely to occur in children receiving language intervention. An important feature of the NRT is that it has been found to be less culturally biased than typical standardized language tests in that scores have not been found to distinguish typically developing White American from African American children (Campbell, Dollaghan, Needleman, & Janosky, 1997). Further, scores on both the NRT and CNRep are reported to be largely independent of performance IQ in children with both typical and atypical language development (Conti-Ramsden et al., 2001; Ellis Weismer et al., 2000; Gathercole et al., 1994).

Although there is substantial evidence that children with SLI show similar patterns of deficit on both the CNRep and NRT, no direct comparisons of performance profiles on the two tasks have been made as yet for a common sample, and the purpose of the present study was to do this. One aim of the study was to examine whether children with SLI and children with typically developing language skills differ equally strongly on both measures. The second motivation for the study was to explore factors that may account for any features of performance specific to the individual tests. The two tests differ in their composition in several ways that are directly relevant to current theoretical accounts of the nonword repetition deficit in SLI. The CNRep (Gathercole & Baddeley, 1996) contains 40 nonwords that range in length from two to five syllables. Some of the stimuli contain consonant clusters (e.g., /ˈblɒntəˈsteɪpɪŋ/, /ˈprɪndl̩/), the majority contain weak syllables with a reduced vowel (e.g., /ˈhæm*pənt*/, /ˈtæ*fləst*/), and many include lexical components and morphemes (e.g., ′pen′ in /ˈ*pen*l̩/ or /ing/ in /ˈblɒntəˈsteɪp*ɪŋ*/). Nonwords are spoken with a natural prosodic pattern characteristic of English words of that particular length. Each nonword repetition attempt is scored online as either correct or incorrect. In contrast, the NRT (Dollaghan & Campbell, 1998) consists of 16 nonwords ranging in length from one to four syllables. All stimuli contain single consonants only drawn from a set without late-acquired phonemes that are acoustically salient and do not include any constituent syllables corresponding to lexical items. The nonwords are spoken with equal stress on each syllable, facilitated by the inclusion of tense vowels only (e.g., /teɪvɑk/). Repetition accuracy is scored from transcriptions as the percentage of phonemes correctly repeated in appropriate positions (see also Gray, 2003; Sahlen et al., 1999).

Why should these differences matter? The reason is that the two nonword sets differ in many of the factors that have been suggested to play a role in the nonword repetition deficit in SLI. An account of this deficit advanced by Gathercole and Baddeley (1990) is that it reflects an underlying impairment of verbal short-term memory. The evidence in support of this claim is as follows. First, the children with SLI in this study also performed poorly on conventional measures of short-term memory such as digit span and word recall (see also, Archibald & Gathercole, in press; Montgomery, 1995), consistent with abundant evidence from other developmental and neuropsychological studies that nonword repetition and digit span are highly correlated with one another (see Baddeley, Gathercole, & Papagno, 1998; Gathercole et al., 1994, for reviews). Second, this group showed the greatest repetition decrement for the lengthiest nonwords, which were four syllables long. Decreased recall accuracy for memory sequences that have lengthy articulatory durations is a hallmark of verbal short-term memory and is typically attributed to temporal decay of the phonological representations in a short-term store (Baddeley, Thomson, & Buchanan, 1975; Cowan, Saults, Winterowd, & Sherk, 1991). By this account, the greater repetition decrement for lengthier nonwords in children with SLI could arise either from accelerated rates of decay before output or from inadequate encoding in the short-term store. Third, it was argued that the unfamiliarity of the phonological structure of nonwords forces participants to rely heavily on temporary phonological representations to support their repetition attempt, preventing the reliance on activated lexical representations that arises in memory tasks using familiar verbal stimuli (e.g., Hulme, Maughan, & Brown, 1991).

Note that according to the original Gathercole and Baddeley (1990) short-term memory account of nonword repetition, children with SLI should be disadvantaged in repeating any lengthy nonwords due to the lack of availability of compensatory lexical support. However, more recent research has established that even memory for nonwords can benefit from some support from knowledge of the lexical and phonotactic composition of the language (Roodenrys & Hinton, 2002; Vitevitch & Luce, 2005). Mechanisms for such support may include lexical and sublexical processing by which internal phonological representations of sound sequences are activated to encode the stimuli (Gupta & MacWhinney, 1997; Martin & Gupta, 2004), or the process of redintegration, by which incomplete memory representations are filled in at the time of retrieval (Gathercole, Frankish, Pickering, & Peaker, 1999; Gathercole, Willis, & Baddeley, 1991). Such lexical and redintegrative processes may be expected to compensate to some degree for the short-term memory deficit in SLI. If this is the case, the nonword repetition impairment of the SLI group in the present study may be expected to be greater on the NRT than the CNRep, due to the lesser opportunity for knowledge-based support in the former than the latter test stimuli.

Several alternative accounts of the nonword repetition deficit in SLI have since been advanced. Children with less extensive vocabulary knowledge may be at a disadvantage in nonword repetition because they have fewer opportunities to supplement temporary representations in short-term memory due to their more impoverished repertoire of lexical and sublexical knowledge (Snowling, Chiat, and Hulme, 1991) or less robustly abstracted representations of individual phonemes (Edwards, Beckman, & Munson, 2004). It may be, also, that children with SLI have less efficient mechanisms either for using lexical knowledge to support short-term memory or for creating representations within long-term memory leading to difficulties in repeating uncommon phoneme sequences even relative to children with similar vocabulary skills. Perhaps, then, the source of the nonword repetition deficit in children with SLI is their more poorly differentiated representational system arising from less efficient lexical mediation. On this basis, children with SLI in the present study should be more disadvantaged on the CNRep than the NRT, as the stimuli used incorporate many more lexical and morphological elements that could potentially benefit children with more extensive vocabulary knowledge or more efficient lexical mediation processes.

Nonword repetition accuracy is also undoubtedly influenced by the quality of speech output processes (Wells, 1995), and this is an area in which children with SLI are known to have problems. Individuals with SLI have more difficulty producing well-organized and stable rhythmic speech motor movements than typically developing children of the same age (Goffman, 1999, 2004); this may provide one possible cause of the nonword repetition deficit. Sahlen et al. (1999) found that maturity of phonological output processes was strongly associated with nonword repetition scores in a sample of young children with language impairment. Also, children with SLI have been reported to be differentially impaired in repeating nonwords containing consonant clusters, which place greater demands on speech output processes due to the need to coordinate a variety of articulatory gestures within a syllable (Bishop et al., 1996). Although the present sample of children with SLI excluded individuals with marked articulatory or phonological impairments, it is possible that the children had more subtle problems with speech motor output that could jeopardize the accuracy of their nonword repetition attempts. In the present study, such difficulties may be reflected as greater SLI decrements in the repetition of the clustered consonants and later developing phonemes of the CNRep nonwords than the single consonants and earlier developing phonemes in the NRT stimuli.

One interpretational problem that commonly arises in investigations of the theoretical underpinnings of nonword repetition and of SLI more generally concerns the fact that language status at the time of testing may place limitations on performance that make it difficult to determine whether poor performance is due to language impairment or some other underlying ability. To address this, many studies include a control group matched with the SLI participants on language abilities. Findings of deficits in children with SLI relative to a language control group are extremely valuable in identifying areas of disproportionate deficit and so in providing clues as to the etiology of the disorder. However, as children matched for language ability are necessarily younger than children with SLI (by nearly 4 years on average in the present study), their cognitive development is less mature, making null effects difficult to interpret (Edwards & Lahey, 1998; Ellis Weismer & Hesketh, 1996; Plante, Swisher, Kiernan, & Restrepo, 1993). For example, accuracy on the CNRep in typically developing children improves by about 20% between the ages of 6 and 9 years (Gathercole et al., 1994). In the present study, general cognitive development is taken into account by adjusting repetition scores using a measure of nonverbal ability as a means of providing a fair comparison between different age groups. Findings that the SLI group performed more poorly relative to both typically developing children matched for age and for language level when scores were adjusted for nonverbal skills would indicate that the nonword repetition deficit in SLI was disproportionate both to language and nonverbal cognitive abilities.

In this study, two tests of nonword repetition, the CNRep (Gathercole & Baddeley, 1996) and the NRT (Dollaghan & Campbell, 1998), were completed by three groups of children: school-age children with SLI and two groups of typically developing children, one matched for age and one matched for language abilities. The study aimed to examine the pattern of group differences in performance across both measures to assess factors that may contribute to the nonword repetition deficit in SLI. Performance was generally expected to be superior on the CNRep than the NRT for all groups due to the inclusion of more wordlike nonwords on the CNRep that provide opportunities for support by existing lexical knowledge. Poorer repetition on both tests was predicted for the SLI group, at least relative to the control group matched for age.

In addition, it was hypothesized that the pattern of SLI decrements across the tests may reflect the relative contribution to the poor nonword repetition in SLI of three factors: short-term memory, lexical mediation, and speech output. Specific predictions for the present study corresponding to each of these cognitive processes are as follows: (a) Disproportionate impairments on the NRT would be consistent with a verbal short-term memory account of the nonword repetition impairment in SLI, as would greater SLI deficits on the lengthier nonwords of both tests. (b) Greater SLI deficits on the CNRep, particularly on the highly wordlike stimuli, would point to poor lexical mediation processes as an important factor. (c) Poorer performance by the SLI group on the CNRep than the NRT may also point to deficits in motor speech output as a factor, as may findings of greater SLI deficits on the nonwords of the CNRep with a high proportion of consonant clusters.

Method

Participants

Thirty-six children participated in three groups in the present study: 12 children with SLI, 12 chronological-age-matched control children (age matched), and 12 language-age-matched control children (language matched). Each group comprised 8 boys and 4 girls. The mean age of the groups was as follows: SLI, 9;8 (years;months; SD = 1.70, range = 7;3–12;5), age matched, 9;9 (SD = 1.64, range = 7;0–12;5), and language matched, 6;1 (SD = 1.61, range = 4;4–10;4). All participants achieved a standard score of 85 or greater on a test of nonverbal reasoning (Raven's Coloured Progressive Matrices; Raven, Court, & Raven, 1986), and a test of articulation (Goldman Fristoe Test of Articulation—Second Edition (GFTA–2; Goldman & Fristoe, 2000). All of the children were native English speakers. None of the children had a diagnosis of attention-deficit disorder or attention-deficit/hyperactivity disorder, autism spectrum disorder, or hearing impairment. Descriptive statistics for the criterion measures completed by the children are summarized in Table 1.

SLI group. The children in the SLI group were recruited from language units (n = 10) and special schools (n = 2) over a 3-month period. The children met the identification criteria for SLI described by Records and Tomblin (1994). They performed at least 1.25 standard deviations below the mean on two of four language measures including one receptive measure. The receptive measures were the British Picture Vocabulary Scale, Second Edition (BPVS–II; Dunn, Dunn, Whetton, & Burley, 1997) and the Test for Reception of Grammar (TROG; Bishop, 1982). The expressive measures were the Expressive Vocabulary Test (EVT; Williams, 1997) and the Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals—UK3 (CELF–UK3; Semel, Wiig, & Secord, 1995). Group means for the BPVS–II, TROG, and EVT were approximately 1.25 standard deviations below the standardized mean and that for the Recalling Sentences subtest was more than 2 standard deviations below the mean. For each test, the proportion of children who scored below the 1.25-standard-deviations cutoff was as follows: BPVS–II, .75; TROG, .67; EVT, .58; Recalling Sentences, 1.0. Individual language profiles conformed closely to the group pattern, with 9 participants scoring more than 1.25 standard deviations below the mean on three or four of the language measures.

Control groups. The children participating in the control groups were recruited over a 1-month period from a school with a similar socioeconomic intake to the schools attended by the SLI group 2 months after recruitment of the SLI group. None of the children had any history of speech or language problems or any type of exceptional educational needs, according to school records. All of the children scored within 1 standard deviation of the mean for their age on three of the language measures described above (BPVS-II, TROG, EVT) and either the Recalling Sentences subtest or an expressive grammatical subtest (see below). The age-matched group was matched to the SLI group on sex and age, and the language-matched group on sex and BPVS–II raw score (mean difference in raw score= 2.5, SD = 2.07).

In one-way analyses of variance (ANOVAs), group differences were found for raw scores on all measures at $p < .001$, with the exception of the Raven score (p = .018). The age-matched group had significantly higher raw scores

Table 1. Descriptive statistics for standardized criterion measures for all groups.

		Participant group					
		SLI		Age matched		Language matched	
Measure	Score	*M*	*SD*	*M*	*SD*	*M*	*SD*
Raven's Matrices	RS	25.33	5.05	29.67_a	3.31	20.58_b	4.10
	SS	102.83	10.37	114.00	9.09	109.75	7.90
GFTA–2	RS	5.50_a	2.28	0.58_b	1.73	5.92_a	5.33
	SS	92.25	3.11	103.17	5.75	103.50	6.20
BPVS–II	RS	66.00_a	13.71	103.50_b	17.64	65.50_a	15.06
	SS	77.08	10.00	107.17	12.34	104.83	11.10
TROG	RS	11.58_a	2.39	17.75_b	1.55	12.50_a	3.03
	SS	77.08	7.82	108.58	11.03	101.25	11.19
EVT	RS	68.58_a	10.64	93.08_b	19.14	59.58_a	10.18
	SS	81.67	11.26	101.83	13.47	100.42	11.81
Recalling Sentences	RS	17.08_a	3.37	43.50_b	12.15	20.08_a	8.22
	ScS	3.25	0.62	9.00	2.45	6.33	2.64

Note. SLI = specific language impairment; Raven's Matrices = Raven's Coloured Progressive Matrices; GFTA–II = Goldman Fristoe Test of Articulation—Second Edition; BPVS = British Picture Vocabulary Scale, Second Edition; TROG = Test for Reception of Grammar; EVT = Expressive Vocabulary Test; RS = raw score; SS = standard score ($M = 100$, $SD = 15$); ScS = scaled score ($M = 10$, $SD = 3$). Means in the same row with different subscripts differ at $p < .01$ in the Tukey honestly significant difference comparison.

on all measures at $p < .01$ except the Raven's ($p = .043$), whereas the SLI and language-matched groups did not differ significantly on raw score on any of the screening measures except the Raven's ($p = .025$). On the basis of this group difference, Raven raw score was used as a covariate in all parametric analyses of group differences.

The language measures used in the present study were selected on the basis of their widespread use for the identification of children with SLI (e.g., Bishop, Bright, et al., 2000; Bishop, North, & Donlan, 1996; Botting & Conti-Ramsden, 2001). It is clear, though, that sentence imitation tasks such as the Recalling Sentences subtest used here not only test language ability but also are linked to short-term memory (Blake, Austin, Cannon, Lisus, & Vaughan, 1994; Willis & Gathercole, 2001). To ensure that participant selection decisions were based on language rather than memory skills, additional expressive language tests were administered in two situations: when children with SLI scored below the cutoff score on only two language tests and one of them was Recalling Sentences ($n = 5$) and when control children scored below age level on Recalling Sentences (age matched: $n = 2$; language matched: $n = 2$). The children in the SLI group completed the two additional expressive subtests of the CELF–UK3 required to compute the test's Expressive Language score for their age (Sentence Assembly and Formulating Sentences). In all cases, the Expressive Language score was greater than 1.25 standard deviations below average, establishing the presence of an expressive language deficit. Those in the control groups completed one test of expressive grammatical skills, the Word Structure subtest of the CELF–UK3, and scored either above 7 for their age or at ceiling, establishing intact expressive grammatical skills.

Procedure

The measures reported in the present study were completed in four individual half-hour sessions in a quiet room in the child's school. In addition to the language screening measures listed above, each child completed the CNRep (Gathercole & Baddeley, 1996) and the NRT (Dollaghan & Campbell, 1998). The language screening measures were completed in two sessions at the time of recruitment. Once recruitment was complete, the nonword repetition tasks were administered in a single session. The nonword repetition tasks were completed in a session in which the participants also completed four visuospatial memory tasks not reported here. The order of tasks for this session was as follows: Completion of two visuospatial tasks, the CNRep, two visuospatial tasks, and the NRT. Responses for all measures were recorded on a digital minidisk player.

Nonword repetition tasks. For each of the nonword repetition tasks, children were told that they would hear some made-up words and be asked to repeat each one exactly as they had heard it. Each word was played once followed by a 3-s pause during which the child responded.

The CNRep (Gathercole & Baddeley, 1996) involves the presentation of 40 nonwords, divided equally into two-, three-, four- and five-syllable items, that the child is required to repeat. Half of the nonwords contain consonant clusters (e.g., /'blɒntə'steɪpɪŋ/), and the remainder have single consonants only (e.g., /wʊɡə'læmɪk/). The nonwords are presented in a fixed random order by audiotape recording. Typical English stress patterns are used in the presentation.

The NRT (Dollaghan & Campbell, 1998) consists of 16 nonwords; 4 stimuli contain one, two, three, and four syllables, respectively. The nonwords are constructed from a limited set of phonemes (11 consonants, 9 vowels) excluding late developing sounds. The nonwords follow an alternating CV structure, and none of the syllables correspond to English lexical items. Only tense vowels are used, and therefore the stress patterns of the nonwords are unlike typical English words in that they have no weak syllables. A detailed description of the criteria guiding the development of the NRT is provided by Dollaghan and Campbell (1998). A fixed random order of nonwords was adopted for this study to ensure consistency across both tests. The nonwords were presented by digital audio recording of a native British adult female speaker following the phonetic transcription and pronunciation guidelines described by Dollaghan and Campbell (1998).

Scoring and reliability. All responses to the nonword repetition tests were scored at the phoneme level to provide comparable data across both tests. Each phoneme was scored as correct or incorrect in relation to the target phoneme. Phoneme substitutions and omissions were scored as incorrect; correctly articulated phonemes with slight distortions were scored as correct. As several research groups in the field do not score phoneme additions as errors (e.g., Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000), phoneme additions were not counted as errors in the present study. In cases of syllable omissions, an anchoring procedure based on vowel alignment was used to align the syllable sequences as closely as possible to the target syllable before individual phoneme scoring. For each nonword repetition task and each nonword length, the number of phonemes correct was divided by the total number of phonemes in the set, resulting in a percentage phonemes correct at each nonword length. These values were used to compute a mean percentage phonemes correct for the entire set for each task. This method of calculating a total score avoids disproportionate contribution by the longest nonwords to the total score (Kane et al., 2004).

Scoring was completed by the first author, a trained speech-language pathologist. A second listener trained in phonetic transcription with no knowledge of the participants' language status and group transcribed 14% of the samples independently (five audiotaped responses for each test). Phoneme percentages of interrater agreement ranged from 79% to 90%, with an average of 86% for the NRT, and from 84% to 96%, with an average of 90% for the CNRep. In addition, four audiotaped responses from participants completing the NRT for each participant group (33%) were transcribed a second time independently by the first author, 4 months after the initial transcriptions. Phoneme percentages of intrarater agreement ranged from 90% to 100%, with an average of 98%.

Results

Descriptive statistics in percentage phonemes correct are presented for both nonword repetition tests and all participant groups in Table 2. As the two tasks differed in numbers of stimuli, scores were expressed as percentage values for the purposes of comparison across tests and syllable lengths. For all participant groups, mean scores were higher for the CNRep than the NRT at equivalent syllable lengths. Accuracy decreased with increase in nonword length for all groups, except for the longest length of the CNRep (see also Gathercole et al., 1994). The decline in repetition accuracy with increasing length was most marked on the NRT. Means across participant groups were similar for shorter nonword lengths, but the age-matched group scored more highly at the longer lengths. The SLI group means were lower than both control groups on the CNRep but lower than the age-matched group only on the NRT.

A rationalized arcsine transform function was used to convert all scores into interval-level data for statistical analysis (Studebaker, 1985). Group differences were evaluated in a series of analyses of covariance (ANCOVA) with Raven's raw score entered as covariate. Before each ANCOVA, a test for homogeneity of regression slopes was completed (Wildt & Ahtola, 1978). In all cases, the interaction between group and Raven score was not significant ($p > .05$), indicating that there were no group differences in the Raven slope function.

Group Comparisons: Nonword Repetition

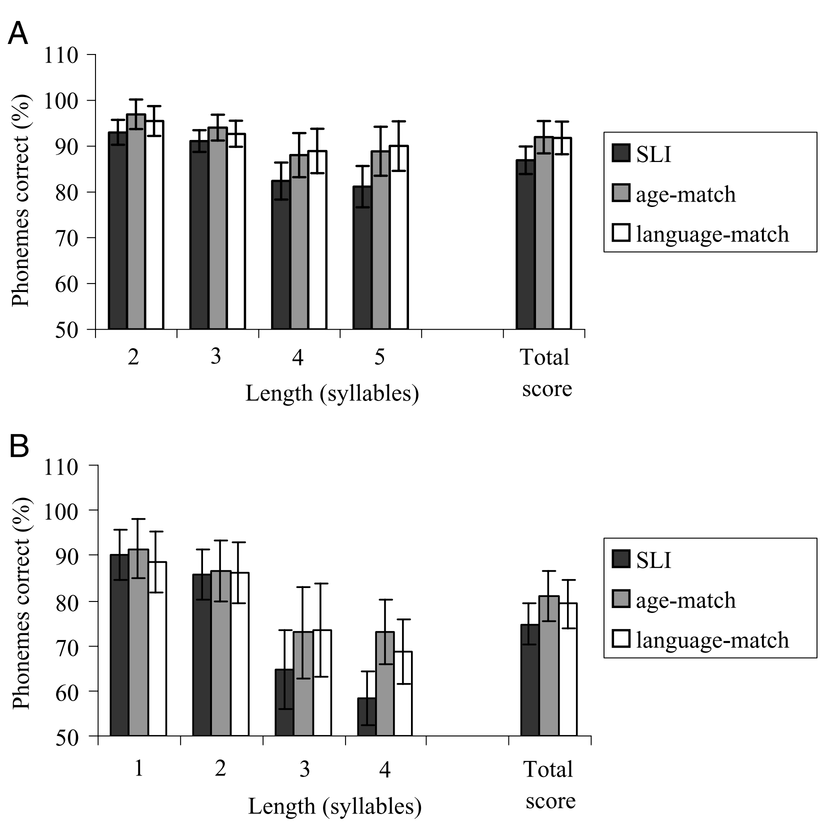
Mean phoneme accuracy scores on both nonword repetition measures adjusted for Raven score for all participant groups are summarized in Figure 1. Separate univariate ANCOVAs were used to assess group differences in the total score of each test for each of the three groups with Raven score as covariate. For the CNRep, the main effect of group was significant, $F(2, 32) = 4.212$, $p = .024$, $\eta_p^2 = .21$, and Raven score was a significant covariate, $F(1, 32) = 7.637$, $p = .009$, $\eta_p^2 = .19$. Planned contrasts revealed that the SLI group performed more poorly than

Table 2. Mean percentage of phonemes correctly repeated at each nonword length for each task and participant group.

		Participant group					
		SLI		Age matched		Language matched	
Measure	**Length**	***M***	***SD***	***M***	***SD***	***M***	***SD***
CNRep	2 syllables	92.99	6.56	96.54	2.25	95.91	3.76
	3 syllables	91.17	5.06	95.78	2.48	90.91	4.83
	4 syllables	82.50	9.18	91.83	4.79	85.05	8.30
	5 syllables	81.30	10.86	93.12	4.71	85.65	8.90
	M	86.99	6.97	94.32	2.68	89.38	5.79
NRT	1 syllable	90.28	7.81	93.75	7.22	86.11	12.48
	2 syllables	85.83	10.41	92.08	6.20	80.42	14.21
	3 syllables	65.18	18.84	85.71	12.56	60.42	23.47
	4 syllables	58.56	15.96	81.48	6.84	59.95	13.48
	M	74.96	10.17	88.26	4.77	71.72	13.96

Note. SLI = specific language impairment; CNRep = Children's Test of Nonword Repetition; NRT = Nonword Repetition Test.

Figure 1. Mean percentage of phonemes correctly repeated, with 95% confidence interval bars for each participant group for each stimulus length and total score. A: Children's Test of Nonword Repetition. B: Nonword Repetition Test. SLI = specific language impairment.



both of the control groups (age match: $p = .024$, language match: $p = .035$). For the NRT, Raven score was a significant covariate, $F(1, 32) = 24.949$, $p < .001$, $\eta_p^2 = .44$, but the main effect of group failed to reach significance, $F(2, 32) = 2.525$, $p = .096$, $\eta_p^2 = .14$. Cognitive development had a significant effect on performance on both tests; however, the effect size of the Raven covariate was considerably larger for the NRT (0.44) than for the CNRep (0.19). Thus, when scores were adjusted for individual differences in cognitive ability, the SLI group performed more poorly than both of the control groups on the CNRep but at equivalent levels to the control groups on the NRT.

In the next set of analyses, the performance of the three groups across syllable lengths was compared for each nonword repetition test separately in mixed-model ANCOVAs as a function of group and syllable length, with Raven score as covariate. For the CNRep scores, there were significant main effects of length, $F(3, 96) = 12.806$, $p < .001$, $\eta_p^2 = .29$, and group, $F(2, 32) = 3.456$, $p = .044$, $\eta_p^2 = .18$. Raven score was a significant covariate, $F(1, 32) = 5.94$, $p = .021$, $\eta_p^2 = .16$, and interacted significantly with length, $F(3, 96) = 6.684$, $p < .001$, $\eta_p^2 = .17$. The interaction between length and group was not significant, $F(6, 96) = 1.246$, $p = .290$, $\eta_p^2 = .07$. The main effect of length reflects a significant decrease in repetition accuracy with increasing length, and the interaction of length with Raven score reflects a significant developmental trend in nonword repetition. Planned contrasts between groups revealed that the SLI group performed more poorly than the age-matched group ($p = .044$). In comparison with the analysis on the CNRep total score, the contrast with the language-matched group in this analysis marginally failed to reach significance ($p = .057$), which likely reflects the reduced power associated with this Group × Length analysis.

In the corresponding analysis on the NRT phoneme accuracy scores across syllable lengths, the main effect of length was significant, $F(3, 63) = 8.337$, $p < .001$, $\eta_p^2 = .21$, due to the decrease in repetition accuracy with increasing stimulus length. Raven score was a significant covariate, $F(1, 32) = 22.429$, $p < .001$, $\eta_p^2 = .41$, and interacted significantly with length, $F(3, 96) = 6.145$, $p = .006$, $\eta_p^2 = .16$. The interaction between length and group was not significant, $F(6, 96) = 1.095$, $p = .371$, $\eta_p^2 = .06$. The main effect of group did not reach significance, $F(2, 32) = 2.103$, $p = .139$, $\eta_p^2 = .17$, although the effect size was similar to that found for the main effect of group in the corresponding CNRep analysis (.18).

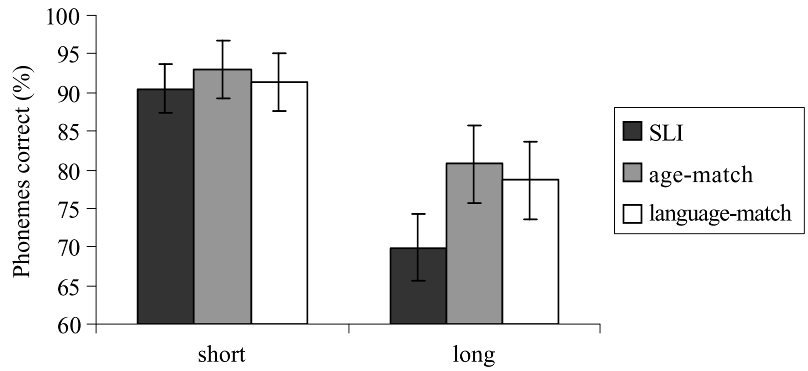
The failure to find a group effect in the NRT total score or syllable length analyses appears to contradict previous findings (e.g., Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000); however, the statistical comparisons involved in these studies differed from the current work in important ways: They compared larger groups, of similar ages, and without covarying nonverbal abilities. To provide directly comparable results, the NRT scores of the SLI and age-matched groups in the present study were compared in an analysis of variance as a function of group and syllable length. All terms were significant: group, $F(1, 22) = 16.080$, $p = .001$, $\eta_p^2 = .42$; length, $F(1, 22) = 19.728$, $p < .001$, $\eta_p^2 = .47$; length and group, $F(1, 22) = 3.154$, $p = .031$, $\eta_p^2 = .13$. Note that when the analysis was repeated with Raven score as covariate, the main effects (both group and length) remained significant. In agreement with previous studies, these results indicate that the SLI group did perform more poorly than the age-matched group with similar nonverbal abilities. Level of cognitive development as indexed by Raven score, however, did have a greater effect on NRT than CNRep performance. It is possible that there was insufficient power to detect a group difference when all three groups (age range = 4–12 years) were included in the analysis due to the small sample size.

To summarize the findings for the group comparisons of total scores and syllable lengths, the SLI group performed more poorly than the age-matched group on the CNRep and NRT but was impaired relative to the language-matched group only on the CNRep once scores were adjusted for nonverbal abilities.

Group Comparisons: Specific Features

Nonword length. Although the performance decrement of the SLI group was greater for the longest nonwords in the present study, the preceding analyses did not establish a significant increase in the SLI deficit with lengthier nonwords as has been found in previous studies (e.g., Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990). One possibility is that the analyses reported above lacked sufficient power to detect an interaction given the presence of four levels in the length variable together with a small sample size. To increase the power for detecting such an interaction, we compared group performance on the shortest (one or two syllables) and longest (four or five syllables) item lengths in a single ANCOVA as a function of group, length, and test, with Raven score as covariate. In the case of the longest items of the CNRep (4 and 5 syllable lengths) and shortest items of the NRT (1 and 2 syllable lengths), where two syllable lengths in the test comprised either long or short nonwords respectively, an accuracy score was computed by averaging across the two respective lengths. As in the previous analyses, there was a main effect of length, $F(1, 32) = 38.363$, $p < .001$, $\eta_p^2 = .55$, and group, $F(2, 32) = 3.563$, $p = .040$, $\eta_p^2 = .18$. There was also a main effect of test, $F(1, 32) = 29.315$, $p < .001$, $\eta_p^2 = .48$, in favor of the CNRep. Raven score was a significant covariate, $F(1, 32) = 10.053$, $p = .003$, $\eta_p^2 = .24$, and interacted significantly with length, $F(1, 32) = 12.458$, $p = .001$, $\eta_p^2 = .28$, and test, $F(1, 32) = 10.194$, $p = .003$,

Figure 2. Mean percentage of phonemes correctly repeated, with 95% confidence interval bars for short and long nonword items for each participant group. SLI = specific language impairment.



η_p^2 = .24, but the three-way interaction between Raven score, test, and length was not significant, $F(1, 32) = 0.050$, $p = .852$, $\eta_p^2 = .002$. Neither the interaction between test and length, $F(1, 32) = 0.033$, $p = .858$, $\eta_p^2 = .001$; test and group, $F(2, 32) = 0.340$, $p = .714$, $\eta_p^2 = .021$; nor test, length, and group, $F(2, 32) = 2.619$, $p = .088$, $\eta_p^2 = .14$, was significant. Importantly, there was a significant interaction between length and group, $F(2, 32) = 4.794$, $p = .015$, $\eta_p^2 = .23$. Analysis of simple effects revealed that group differences occurred on the long nonwords, $F(2, 32) = 7.398$, $p = .002$, $\eta_p^2 = .32$, but not the short nonwords, $F(2, 32) = 0.491$, $p = .617$, $\eta_p^2 = .03$. In the case of the long nonwords, the SLI group performed more poorly than both the age-matched ($p = .002$) and language-matched ($p = .013$) control groups. Figure 2 displays adjusted means for this interaction and establishes that the decline in performance on the long versus short items was more marked for the SLI as compared with the other two groups.

Wordlikeness. The possible contribution of poor lexical knowledge to the nonword repetition deficit in SLI was examined in a post hoc analysis of the CNRep. One way of investigating the degree of lexical mediation in nonword repetition is by comparing repetition accuracy on stimuli rated as low and high in wordlikeness (Gathercole, 1995; Gathercole, Willis, Emslie, & Baddeley, 1991). The wordlikeness ratings obtained by Gathercole, Willis, Emslie, and Baddeley (1991) on a 5-point scale ranging from 1 (*very unlike a real word*) to 5 (*very like a real word*) were used to create two subsets of 14 CNRep items matched for number of phonemes, syllables, and consonant clusters (see Appendix A). One set contained nonwords of low-rated wordlikeness, whereas the other contained high-wordlikeness items. Mean percentage phoneme accuracy and standard deviations for each participant group on the high- and low-wordlikeness sets are summarized in Table 3. The pattern of slightly higher repetition accuracy for the high- than the low-wordlikeness set was consistent for all groups, although the difference was minimal for the age control group. In the ANCOVA performed on the scores for the wordlikeness sets between groups with Raven score as covariate, the main effect of wordlikeness failed to reach significance, $F(1, 32) = 3.653$, $p = .065$, $\eta_p^2 = .10$. There was a significant main effect of group, $F(2, 32) = 4.441$, $p = .020$, $\eta_p^2 = .22$, but not a significant Wordlikeness × Group interaction, $F(2, 32) = 0.163$, $p = .850$, $\eta_p^2 = .01$. Raven score was a significant covariate, $F(1, 32) = 8.314$, $p = .007$, $\eta_p^2 = .21$, but did not interact significantly with

Table 3. Mean percentage of phonemes correctly repeated for subsets of CNRep items for all participant groups.

	Participant group					
	SLI		Age-matched		Language-matched	
Subset type	*M*	*SD*	*M*	*SD*	*M*	*SD*
Wordlikeness						
High wordlikeness	88.85	5.87	94.43	2.71	91.11	3.60
Low wordlikeness	83.38	10.58	92.74	4.57	86.26	7.76
Articulatory complexity						
Single consonants	89.68	6.09	93.41	3.36	89.98	4.81
Consonant clusters	82.36	15.76	94.68	1.87	90.12	5.29

wordlikeness, $F(1, 32) = 0.980$, $p = .330$, $\eta_p^2 = .03$. The SLI group had lower scores than both the age- ($p = .20$) and language-matched groups ($p = .033$) in planned contrasts. In the present data, then, there was no difference in sensitivity to wordlikeness across the SLI and control groups.

Articulatory complexity. A second post hoc analysis of the CNRep data explored the possible effects of speech motor output processes on nonword repetition, as indexed by the presence of consonant clusters. As clusters are more complex to produce than single consonants, a finding of increased SLI deficits with stimuli containing these structures may point to motoric differences as a source of the nonword repetition deficits. In the present study, two subsets of 15 CNRep items were created matched for syllable length (see Appendix B). Items in the single-consonant group had an alternating CV structure only, with no adjacent consonants even across syllable boundaries, whereas items in the consonant clusters group had clusters in at least half of their syllables. The subsets included 5 items at the two-syllable length, 5 at the three-syllable length, and 5 at the four-syllable length. The five-syllable length was not included because there were insufficient tokens of five-syllable items without consonant clusters for comparison. Mean percentage phoneme accuracy and standard deviations are shown in Table 3 for each group on the lists containing words with either single consonants or consonant clusters. Mean scores were lower on the consonant clusters than single-consonant list for all groups, but the reduction in performance across lists was greatest for the SLI group. Results of the ANCOVA comparing articulatory complexity, and participant groups on phoneme accuracy with Raven score as covariate revealed one significant effect, a significant interaction between articulatory complexity and group, $F(2, 32) = 4.321$, $p = .022$, $\eta^2_p = .21$. Raven score was a significant covariate, $F(1, 32) = 6.960$, $p = .533$, $\eta^2_p = .13$, but did not interact significantly with articulatory complexity, $F(1, 32) = 0.398$, $p = .533$, $\eta^2_p = .01$. Neither the main effects of articulatory complexity, $F(1, 32) = 0.224$, $p = .640$, $\eta^2_p = .01$, or group, $F(2, 32) = 2.823$, $p = .074$, $\eta^2_p = .15$, reached significance. Analysis of simple effects indicated that the groups differed on the consonant cluster, $F(2, 32) = 6.268$, $p = .005$, $\eta^2_p = .28$, but not the single-consonant nonword list, $F(2, 32) = 0.529$, $p = .594$, $\eta^2_p = .03$. Planned contrasts confirmed that the SLI group performed at a lower level on the consonant cluster word list than either the age-matched ($p = .003$) or the language-matched ($p = .014$) groups.

Discussion

Nonword repetition deficits were found in the present study in a group of children with SLI when compared across two measures of nonword repetition: the CNRep (Gathercole & Baddeley, 1996) and the NRT (Dollaghan & Campbell, 1998). The performance of the SLI group was impaired relative also to younger, typically developing children with similar language abilities once scores were adjusted for differences in general cognitive ability on the CNRep only. The magnitude of the impairment of the SLI group increased for lengthier and for more articulatory complex nonwords, relative to both typically developing groups. Of the two tests, the NRT was influenced to a greater degree by differences in developmental cognitive ability.

In line with previous findings, the results indicate that children with SLI have a disproportionate difficulty in repeating novel phonological sequences. Why is this? One proposal has been that an impairment of short-term memory may underlie the nonword repetition deficit in SLI (Gathercole & Baddeley, 1990). In keeping with this suggestion, the nonword repetition deficit of the SLI group relative to the control groups was greatest for the longest nonwords (see also Botting & Conti-Ramsden, 2001; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Ellis Weismer et al., 2000; Kamhi & Catts, 1986; Kamhi et al., 1988; Marton & Schwartz, 2003; Montgomery, 1995, 2004). According to Baddeley et al. (1998), short-term memory plays a key role in learning new words by generating a phonological representation of brief and novel speech events that mediates the creation of a phonological entry within the long-term lexical store. Children with SLI therefore may have more difficulty learning new words because their short-term memory representations are inadequate.

The present findings cannot, however, be readily accommodated solely by a verbal short-term memory deficit in SLI. One crucial difference between the two repetition tests is that nonwords in the NRT do not incorporate any lexical or sublexical units, whereas items in the CNRep include sublexical and morphological components. On this basis, it would be expected that the NRT would have a greater dependence on short-term memory, due to the reduced opportunities for support via lexical and phonotactic redintegration processes. In line with this view, repetition accuracy was greater on the CNRep than the NRT even though the order of test administration (NRT second) may have been expected to benefit the NRT (Gray, 2003). If a short-term memory deficit alone underpins the nonword repetition deficit in SLI, children with SLI should be more disadvantaged on the NRT than the CNRep as a consequence of lack of opportunity for successful redintegration. Contrary to this prediction, the SLI group deficit was found to be greater on the CNRep. The children with SLI obtained lower scores on the CNRep than both groups of control children, whereas the typically developing groups did not differ from one another. In contrast, performance of the SLI and language groups was equivalent on the NRT even when adjusted for nonverbal ability.

One explanation for this finding is that children with SLI have less efficient lexical mediation processes to support short-term memory in the course of nonword repetition. The finding of a greater SLI deficit on the CNRep is consistent with this view. In addition, though, it would be expected that the SLI group should show a reduced advantage to nonwords high in wordlikeness, whereas in fact no group deficits in sensitivity to wordlikeness were observed. It should, however, be acknowledged that performance approached ceiling levels for the age-matched control children in part of this analysis, which could potentially have masked a greater benefit of high wordlikeness in this group. Further systematic investigation is required to resolve this issue.

One factor that does appear to have contributed a differential effect on the SLI group is the articulatory complexity of the nonword stimuli. It may be that the greater SLI deficit on the CNRep arises in part from the inclusion of consonant clusters in contrast to the NRT. In keeping with this notion, only the SLI group showed a marked decline in repetition accuracy on the nonwords containing consonant clusters, in line with Bishop et al.'s (1996) findings. Thus, although in both studies, the children with language impairments had no gross motor speech deficits, they were further disadvantaged in nonword repetition when the articulatory demands of the stimuli were particularly complex. There are at least two possible interpretations of this finding: The children with SLI may have less robust phonological representations for these relatively uncommon phoneme combinations, although recent evidence from children with phonological impairments has not supported this suggestion (Munson, Edwards, & Beckman, 2005). Alternatively, the children with SLI may have difficulty forming the novel phonological sequences required in nonword repetition. In line with this view, Goffman (2004) reported that children with SLI have difficulty producing well-organized and stable rhythmic speech motor movements, which may affect their ability to repeat nonwords. It is possible, then, that the (poor) speech motor output skills of children with SLI contribute in part at least to their difficulties in nonword repetition.

The two nonword repetition tasks also differed in the strength of their associations with the nonverbal reasoning measure that is widely interpreted as tapping general cognitive maturity, Raven's Matrices. Scores on the NRT were much more highly associated with performance on the Raven test than CNRep scores, suggesting that the former nonword repetition task may be more closely linked to general cognitive development than the CNRep. One limitation of the present study was the small sample size together with the large age range of participants across groups (4–12 years), which may account for the failure to detect a group difference when all three groups were compared on the test more sensitive to cognitive abilities, the NRT. One further possibility is that the SLI group may have benefited from a differential practice effect, resulting in greater improvements on the second repetition test administered (NRT) relative to the control groups (Gray, 2003). SLI deficits on the CNRep, on the other hand, were demonstrable even for the small sample sizes involved in the present study, as they were for the NRT when the children with SLI were compared with their age peers.

In summary, this study confirms previous findings of poor nonword repetition in children with SLI for two tests of nonword repetition: the CNRep (e.g., Bishop et al., 1996; Botting & Conti-Ramsden, 2001; Gathercole & Baddeley, 1990) and the NRT (e.g., Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000). In line with a verbal short-term memory account of the deficit, the SLI group had more difficulty holding novel phonological forms in mind as reflected by the increased magnitude of their repetition impairment for longer nonwords. The test with the greater ability to identify SLI deficits was the CNRep, in which items incorporate sublexical units, grammatical morphemes, and consonant clusters. The children with SLI had more difficulty repeating words with increased articulatory complexity but benefited from lexical similarity of nonwords to the same extent as typically developing children. These results suggest that verbal short-term memory alone may not provide a full explanation of the nonword deficit in SLI. It is possible that there are multiple origins to the deficit in nonword repetition: verbal short-term memory, lexical knowledge, output processes, and others. The CNRep, therefore, may better reflect the nonword repetition deficit in SLI because it incorporates stimuli that tap several of these components.

Acknowledgments

This research was supported by a Bamford-Lahey Children's Foundation Scholarship awarded to Lisa M. D. Archibald. We thank the children and staff of all of the participating schools for their cheerful cooperation.

References

Archibald, L. M. D., & Gathercole, S. E. (in press). Short-term and working memory in SLI. *International Journal of Language and Communication Disorders*.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105,* 158–173.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589.

Bishop, D. V. M. (1982). *Test for Reception of Grammar*. Oxford, United Kingdom: Medical Research Council/Chapel Press.

Bishop, D. V. M., Bishop, S. J., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language, and Hearing Research, 42,* 155–168.

Bishop, D. V. M., Bright, P., James, C., Bishop, S. J., & van der Lely, H. (2000). Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics, 21,* 159–181.

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37,* 391–403.

Blake, J., Austin, W., Cannon, M., Lisus, A., & Vaughan, A. (1994). The relationship between memory span and measures of imitative and spontaneous language complexity in preschool children. *International Journal of Behavioral Development, 17,* 91–107.

Botting, N., & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *International Journal of Language and Communication Disorders, 36,* 421–432.

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

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

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

Cowan, N., Saults, J. S., Winterowd, C., & Sherk, M. (1991). Enhancement of 4-year-old children's memory span for phonologically similar and dissimilar word lists. *Journal of Experimental Child Psychology, 51,* 30–52.

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

Dunn, L. M., Dunn, L. M., Whetton, C. W., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd ed.). Windsor, UK: NFER-Nelson.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47,* 421–436.

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

Ellis Weismer, S., & Hesketh, L. (1996). Lexical learning by children with specific language impairment: Effects of linguistic input presented at varying speaking rates. *Journal of Speech and Hearing Resarch, 39,* 177–190.

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

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition, 23,* 83–94.

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

Gathercole, S. E., & Baddeley, A. D. (1996). *The Children's Test of Nonword Repetition*. London: Psychological Corporation.

Gathercole, S. E., Frankish, C., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 84–95.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British Journal of Psychology, 82,* 387–406.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12,* 349–367.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1994). The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory, 2,* 103–127.

Goffman, L. (1999). Prosodic influences on speech production in children with specific language impairment and speech deficits. *Journal of Speech, Language, and Hearing Research, 42,* 1499–1517.

Goffman, L. (2004). Kinematic differentiation of prosodic categories in normal and disordered language development. *Journal of Speech, Language, and Hearing Research, 47,* 1088–1102.

Goldman, R., & Fristoe, M. (2000). *Goldman Fristoe 2: Test of Articulation*. Circle Pines, MN: AGS.

Gray, S. (2003). 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, 36,* 129–151.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59,* 267–333.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30,* 685–701.

Kamhi, A., & Catts, H. (1986). Toward an understanding of developmental language and reading disorders. *Journal of Speech and Hearing Disorders, 51,* 337–347.

Kamhi, A. G., Catts, H. W., Maurer, D., Apel, K., & Gentry, B. F. (1988). Phonological and spatial processing abilities in language- and reading-impaired children. *Journal of Speech and Hearing Disorders, 53,* 316–327.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189–217.

Martin, N., & Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology, 21,* 213–228.

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech Language and Hearing Research, 46,* 1138–1153.

Montgomery, J. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research, 38,* 187–199.

Montgomery, J. (2004). Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. *International Journal of Language and Communication Disorders, 39,* 115–134.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research, 48,* 61–78.

Newcomer, P., & Hammill, D. (1988). *Test of Language Development—Primary, Second Edition*. Austin, TX: Pro-Ed.

Norbury, C. F., Bishop, D. V. M., & Briscoe, J. (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Linguistics, 23,* 247–268.

Plante, E., Swisher, L., Kiernan, B., & Restrepo, M. A. (1993). Language matches: Illuminating or confounding? *Journal of Speech and Hearing Research, 36,* 661–667.

Raven, J. C., Court, J. H., & Raven, J. (1986). *Raven's Coloured Matrices*. London: H. K. Lewis.

Records, N., & Tomblin, J. B. (1994). Clinical decision making: Describing the decision rules of practicing speech-language pathologists. *Journal of Speech and Hearing Research, 37,* 144–156.

Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on the serial recall of nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 29–33.

Sahlen, B., Reuterskiold-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Non-word repetition in children with language impairment–Pitfalls and possibilities. *International Journal of Language and Communication Disorders, 34,* 337–352.

Semel, E., Wiig, E., & Secord, W. (1995). *Clinical Evaluation of Language Fundamentals—UK 3*. London: The Psychological Corporation.

SLI Consortium. (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *American Journal of Human Genetics, 70,* 384–398.

SLI Consortium. (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *American Journal of Human Genetics, 74,* 1225–1238.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords and phonological processes: Some comments on Gathercole, Willis, Emslie and Baddeley. *Applied Psycholinguistics, 12,* 399–373.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41,* 407–418.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28,* 455–462.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language, 52,* 193–204.

Wells, B. (1995). Phonological considerations in repetition tasks. *Cognitive Neuropsychology, 12,* 847–855.

Wildt, A. R., & Ahtola, O. T. (1978). *Quantitative applications in the social sciences: Analysis of covariance*. Beverly Hills, CA: Sage.

Williams, K. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: AGS.

Willis, C. S., & Gathercole, S. E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory, 9,* 349–363.

Received May 12, 2005

Revision received September 8, 2005

Accepted January 13, 2006

DOI: 10.1044/1092-4388(2006/070)

Contact author: Lisa M. D. Archibald, Department of Psychology, University of Western Ontario, London, Canada N5X 3V6. E-mail: larchiba@uwo.ca

Susan E. Gathercole is now at University of York, York, United Kingdom.

Appendix A. Nonwords from CNRep in each wordlikeness set with wordlikeness ratings.

Low-wordlike set		High-wordlike set	
Nonword	**Rating**	**Nonword**	**Rating**
lɒdə'eɪpɪʃ	1.1	'hampənt	2.7
wʊgə'læmɪk	1.6	'sladɪŋ	2.7
'skɪtɪkʊlt	1.7	vɜsə'treɪʃanɪst	2.8
'tæfləst	1.8	'bərəzən	2.9
ditrəta'pɪlɪk	2.0	'glɪstərɪŋ	3.0
'brastərə	2.1	fɛnə'raɪze	3.0
empli'fɔvənt	2.2	'kəm'isəteɪt	3.2
'dɪlə	2.3	sɛprə'tɛnɪʊl	3.3
'glɪstəʊ	2.3	kən'frantʃulɪ	3.4
prɪstər'akʃənḷ	2.3	'pɛnḷ	3.4
'baləp	2.4	'trʊmpətin	3.4
pən'ɛrɪfḷ	2.4	stɒpə'gratɪk	3.5
riʌtə'peɪʃn	2.4	'rubɪd	3.8
altʃu'peɪtəri	2.5	difɜmɪ'keɪʃn	3.9

Note. CNRep = Children's Test of Nonword Repetition.

Appendix B. Subsets of CNRep items in each articulatory complexity group.

Length	High complexity	Low complexity
2 syllables	'hampənt	'pɛnḷ
	'glɪstəʊ	'baləp
	'sladɪŋ	'rubɪd
	'tæfləst	'dɪlə
	'prɪndḷ	'banəʊ
3 syllables	'glɪsterɪŋ	'dɒpəleɪt
	'frɛskəvɛnt	'banɪfə
	'trʊmpətin	'barəzən
	'skɪtɪkʊlt	'kɒmərin
	'brastərə	'θɪkərɪ
4 syllables	'kəntrampənɪst	wʊgə'læmɪk
	pɜ'plɪstəroɲk	fɛnə'raɪzə
	'blɒntə'steɪpɪŋ	'kəm'isəteɪt
	stɒpə'gratɪk	lɒdə'eɪpɪʃ
	empli'fɔvənt	pən'ɛrɪfḷ

Note. CNRep = Children's Test of Nonword Repetition.

Copyright of Journal of Speech, Language & Hearing Research is the property of American Speech-Language-Hearing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.