Original Article

Interaction Between Working Memory and Long-Term Memory

A Study in Children With and Without Language Impairment

Klara Marton[1,2] and Naomi Eichorn[1]

[1]The Graduate School and University Center of the City University of New York, NY, USA,
[2]Barczi Gusztav College of Special Education of Eotvos Lorand University, Budapest, Hungary

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Abstract. Individual differences in working memory have been related to interactions between working memory and long-term memory (LTM). The present study examined this interaction in children with and without language impairment. We used two listening span tasks and two nonword repetition tasks. The results suggest a strong interaction among age, language status, and task complexity. Children with specific language impairment showed consistently poor performance across tasks and indicated a weakness in using long-term knowledge to support working memory performance. The findings show that these children do not benefit from various manipulations designed to enhance working memory performance via LTM support due to a combination of inefficiencies in maintaining and updating items in working memory and retrieving information from LTM, in part because of their poor resistance to interference.

Keywords: working memory, long-term memory, strategy use, children, specific language impairment

Findings on working memory (WM) functioning show significant individual differences in both children and adults, as well as in clinical populations (Coady & Evans, 2008; Unsworth, Brewer, & Spillers, 2013; Was & Woltz, 2007). One of the most significant contributing factors to these individual differences is the interaction between WM and long-term memory (LTM). Understanding the nature of this interaction in children is important from a theoretical point of view, as well as from an educational perspective. Although a number of studies (e.g., Marton & Schwartz, 2003; Montgomery, 2000) reported a deficit in WM and related executive functions in children with specific language impairment (SLI), underlying factors contributing to the WM capacity deficit in these children have not been identified.

Individual differences in WM capacity reflect both variations in the ability to actively maintain and update information over a limited period of time and the ability to retrieve task-relevant information from LTM (Unsworth & Engle, 2007). Efficient retrieval of task-relevant items is particularly difficult when interfering information is present (Oberauer, 2005). Individuals with low WM capacity tend to retrieve more irrelevant items because they are more susceptible to interference, which leads to less accurate recall of target items.

Interaction Between WM and LTM: A Developmental Perspective

Most cognitive models that describe the association between WM and LTM are based on data from adults (e.g., Oberauer, 2005; Unsworth & Engle, 2007). Developmental scientists often use these models in WM studies with typically developing children (e.g., Magimairaj, Montgomery, Marinellie, & MacCarthy, 2009) because there is no comprehensive developmental model to describe the interaction between WM and LTM and account for age-related changes. Most developmental studies examined performance in one of the memory systems. Only a few researchers investigated specific memory functions in both systems. Lloyd, Doydum, and Newcombe (2009) examined familiarity and recollection within a recognition paradigm and found that young children (age 4) produced an increased false alarm rate for recollection in LTM compared to WM tasks. The authors concluded that young children show a weakness in retrieval of relevant information from LTM. Studies on school-age children revealed that both WM and long-term knowledge predict learning outcomes but the interaction between the two systems was not examined in any of these studies (e.g., Alloway, 2009).

 © 2014 Hogrefe Publishing

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

An important phenomenon that has recently received considerable attention in the developmental literature is the ability to resist interference (Unsworth et al., 2013), which improves gradually during childhood and into adolescence (Schleepen & Jonkman, 2009). Interference may influence both memory systems, negatively affecting the maintenance and updating of relevant items in WM, as well as the retrieval of information from LTM. Effective resistance to interference is associated with larger WM capacity because relevant and irrelevant item representations compete for the same limited WM capacity.

The majority of developmental research examining WM and its relation to LTM used variations of two basic task types: verbal span tasks and nonword repetition (NWR) tasks. Although both tasks can be used to examine the relationship between WM and LTM, they target different aspects of WM. Verbal span tasks measure capacity limitations in processing complex linguistic information and require active maintenance and updating of items in WM, as well as systematic searches in LTM. Studies using verbal span tasks show that children's WM performance improves with age and that these developmental changes are associated with faster speed of processing (Kail, 2000), larger storage capacity (Magimairaj et al., 2009), and more efficient strategy use (Fatzer & Roebers, 2013).

Studies using NWR tasks focus on storage and processing of phonological information and its relationship to LTM. This interaction has been examined using nonwords with high and low phonotactic probability (frequency of certain phoneme combinations; Edwards, Beckman, & Munson, 2004) and nonwords with varying degrees of wordlikeness (Gathercole, 1995). Improved NWR skills in older compared to younger children have also been associated with the development of phonological sensitivity (Morra & Camba, 2009) and vocabulary (Edwards et al., 2004). Older children show more structured and stable word representations in LTM than younger children and these factors facilitate NWR.

WM Capacity in Children With Specific Language Impairment (SLI)

Children with SLI show significant limitations in WM capacity compared to their typically developing (TLD) peers (Marton & Schwartz, 2003; Montgomery, 2000) and perform poorly on verbal span tasks (Marton & Schwartz, 2003; Montgomery, 2000). Compared to TLD participants, children with SLI exhibit diminished primacy and recency effects in listening span tasks (Marton & Schwartz, 2003), their performance deteriorates to a greater extent when sentence complexity increases (Marton & Schwartz, 2003; Montgomery & Evans, 2009), and they show less efficient resistance to interference (Marton, Campanelli, Eichorn, Scheuer, & Yoon, in press, Spaulding, 2010). Most of these studies found that at a baseline level with simple sentences and minimal interference, children with SLI perform similarly to their TLD peers. With an increase in task complexity, children with SLI either perform similarly to younger children or show a distinct performance pattern. This WM deficit in children with SLI has been associated with difficulty with concurrent processing and storage (Marton & Schwartz, 2003; Montgomery, 2000); poor inhibition, particularly resistance to interference (Marton et al., in press); and slower processing speed (Miller, Kail, Leonard, & Tomblin, 2001). Based on these findings, we may conclude that WM shows a multifaceted deficit in children with SLI; however, none of these studies examined interactions between WM performance and LTM retrieval in children with SLI.

A second large set of studies focused on phonological WM in children with SLI. These studies used various versions of NWR tasks and consistently found that children with SLI performed more poorly than their peers (Marton & Schwartz, 2003; Montgomery & Windsor, 2007). An important aspect of NWR with regard to LTM is the wordlikeness effect. Repetition of nonwords with no meaningful syllables is a measure of phonological WM, whereas more wordlike nonwords reflect the support from LTM (Gathercole, 1995). Casalini and colleagues (2007) examined the wordlikeness effect using morphemes and compared nonwords with no meaning to nonwords with meaningful morphemes. Although children with SLI performed more poorly than TLD children, they showed a similar profile of recall, with highest recall for real words, followed by nonwords with real morphemes, followed by nonwords with no meaningful morphemes. These findings reflect facilitatory effects of LTM on phonological WM.

The Goal of the Current Study

We examined the interaction between WM and LTM in children with and without SLI in two experiments. We chose the Traditional Listening Span Task (TLST) and Traditional Nonword Repetition Task (NWR) as reference tasks because these measures are widely used in the SLI literature. We also developed modified versions of each task to more directly examine the contribution of LTM to WM. In Experiment 1, we administered a TLST and an Active Listening Span Task (ALST), in which we measured children's performance on items they had activated and retrieved prior to administration of the span task. In Experiment 2, we compared performance on traditional NWR to a modified NWR in which nonwords contained one monosyllabic meaningful word, with either high or low frequency of occurrence. Based on previous findings in TLD children, nonwords that contain meaningful syllables are easier to recall than less wordlike nonwords (Gathercole, 1995).

We tested the following hypotheses:

Hypothesis 1: In Experiment 1, all groups will show better performance on the ALST than on the TLST because in the former task, children repeat words that were previously activated and retrieved from LTM. Serial recall of items is highly influenced by LTM

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(Burgess & Hitch, 2005), thus recalling recently retrieved words from LTM is easier than recalling words that have not been previously activated.

Hypothesis 2: All children will perform better on syntactically simple compared to complex sentences in both listening span tasks; however, the benefit from LTM support in the ALST will be greater for complex sentences than for simple sentences. Processing complex sentence structures while maintaining sentence final words requires more efficient contribution from LTM than remembering sentence-final words of simple sentences (MacDonald, Just, & Carpenter, 1992).

Hypothesis 3: The WM deficit in children with SLI is highly influenced by their poor resistance to interference. These children will maintain more irrelevant items in WM and retrieve more irrelevant information from LTM than their peers, as reflected by a larger number of interference errors. If our task modifications facilitate WM performance, we should see a decrease in interference errors in the ALST compared to TLST.

Hypothesis 3.1: Interference errors in sentence-final word repetition reflect the maintenance of irrelevant items in WM.

Hypothesis 3.2: Interference errors in children's answers to content questions reflect retrieval problems from LTM. Because children could not anticipate specific details targeted by comprehension questions, this information required storage to and retrieval from LTM.

Hypothesis 4: In Experiment 2, NWR will be more accurate for stimuli in the modified compared to the traditional task because wordlike nonwords are repeated more accurately than nonwords with no meaningful syllables (Dollaghan, Biber, & Campbell, 1993). If children use LTM knowledge efficiently, then the following pattern is expected: nonwords with high frequency words > nonwords with low frequency words > nonwords with no meaningful syllables.

Methods

Participants

Three groups participated in this study (15 children with SLI, 15 age-matched peers, and 15 language-matched controls). All participants spoke English as their primary language and had average (> 85) nonverbal intelligence, based on the Test of Nonverbal Intelligence (TONI-3, Brown, Sherbenou, & Johnsen, 1997). Children with SLI (age range: 7;8–10;1 years; mean age: 8;11) were enrolled in speech-language therapy at the time of the study and scored at least 1.25 *SD* below the age appropriate level (range of standard scores: 81–64) on the Clinical Evaluation of Language Fundamentals-4 (CELF-4; Semel, Wiig, & Secord, 2003).

The other two groups consisted of typically developing children who all passed a language screening (CELF-3 Screening Test; Semel, Wiig, & Secord, 1995) and showed typical developmental patterns, as confirmed by parent and teacher interviews. The age-matched typically developing children (AM; age range: 7;9–10;3 years; mean age: 8;9) were each matched (within 3 months) to the children with SLI. The younger language-matched typically developing children (LM; age range: 6;8–9;11; mean age: 7;4) were each matched to one of the children with SLI within three raw scores on the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn & Dunn, 1997). The inclusion of the language-matched group allowed us to control for potential effects of vocabulary on WM performance.

Each group consisted of 2 females and 13 males. All participants passed a pure-tone audiometric screening at 20 dB HL (at 500; 1,000; 2,000; and 4,000 Hz) bilaterally. Apart from the language impairment in children with SLI, none of the participants demonstrated any other disorder. Exclusionary criteria included articulatory errors and motor, emotional, or physical disturbances. See Table 1 for a summary of participant characteristics.

Stimuli

Traditional and Active Listening Span Tasks

The task included two lists of sentences, each ranging in length between 2 and 5 sentences per set (6 sets of

Table 1. Participant characteristics. Mean (*SD*)

	SLI (*n* = 15)	AM (*n* = 15)	LM (*n* = 15)
Gender (male/female)	13/2	13/2	13/2
Age (years;months)	8;11 (10.3)	8;9 (10.2)	7;4 (16.3)
TONI-3 standard score	104.7 (16.8)	110.0 (16.6)	110.6 (16.7)
CELF standard/criterion score	72.53[a] (9.3)	25.71[b] (5.6)	20.93[b] (7.3)
PPVT-3 raw score	93.3 (18.1)	114.07 (14.2)	95.8 (17.4)

Notes. SLI = children with specific language impairment; AM = age-matched controls; LM = language-matched controls; TONI-3 = Test of Nonverbal Intelligence (3rd ed.); CELF = Clinical Evaluation of Language Fundamentals, PPVT = Peabody Picture Vocabulary Test (3rd ed.). [a]CELF-4 core language score, [b]CELF-3 Screening Test, criterion score.

 © 2014 Hogrefe Publishing

2 sentences per set, 6 sets of 3 sentences per set, etc., in each list; total of 84 sentences per list; total of 168 sentences for the task). Each set contained a question targeting sentence content. The first list was created as a TLST (e.g., "Kelly likes to play with her *doll*." "The conductor blows his *horn*."). The sentence-final words were high frequency words typically acquired during the preschool years (Hall, Nagy, & Linn, 1984) and were 1–2 syllables in length.

In the ALST, children were presented with incomplete sentences one at a time (e.g., "Fred reads poems to his ____.") and were asked to complete each sentence with a single word. There were no restrictions regarding the word type used, but it had to complete the sentence in a meaningful way. Once children completed all 84 sentences in this list, the sentences were presented to them as a listening span task (see section Procedures for additional details).

All sentences consisted of 10 or fewer syllables. Of the six sets at each set length in each task, three sets of sentences were syntactically simple (e.g., "Molly hates the big fat cat.") and three sets were syntactically complex (e.g., "Bill thinks that Sue likes the wise man.").

Nonword Repetition With and Without Meaningful Syllables

Two lists of nonwords were presented to each participant. The first list consisted of 16 nonwords with no meaningful syllables (/'giləsɛp/; /fo'kuʃəmɪk/; Marton & Schwartz, 2003). The second list consisted of 32 nonwords with one meaningful monosyllabic real word (e.g., /kaʊ'rimət/, /no'kuʃəmɪk/). Monosyllabic real words were either high or low frequency words (16-16 each, based on Hall et al., 1984). Nonwords in both lists were three or four syllables long. Previous data (Marton & Schwartz, 2003) showed that children with SLI perform similarly to age-matched peers in repeating two-syllable nonwords, but their performance deteriorates with three-syllable stimuli.

Procedures

Participants were tested individually in a single session. For the listening span tasks, participants were instructed to listen to a set at a time, repeat sentence-final words in the order of presentation, and answer the question that followed presentation of each set. Each set included a question to ensure that children processed the entire sentence, not only the sentence-final words. The correct answer to the question was never a sentence-final word. Sentences for both listening span tasks were read and recorded by the examiner. Sets with differing number of sentences (2-3-4-5) were presented in a random order for both lists to avoid proactive interference for only the longer sets. Practice trials with different stimuli were used with each participant.

Stimuli for the nonword repetition tasks were recorded by a female speaker and presented via headphones. Children's responses were audio-recorded. Nonwords from the two lists were presented in a random order. Testing and data analyses were performed by different investigators.

Results

Traditional Versus Active Listening Span

Accuracy data were entered into a mixed ANOVA comparing task type (traditional vs. active), syntactic complexity (simple vs. complex), and set length (2, 3, 4, and 5 sentences) for the three groups. For set length, Mauchly's test indicated violation of the sphericity assumption, $\chi^2(5) = 27.94$, $p < .001$; therefore, degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = .73$) for this factor. All reported post hoc pairwise comparisons were performed with Bonferroni adjustments (see descriptive data in Table 2).

Table 2. Listening span tasks. Descriptive statistics for recall accuracy by group, task type, set length, and sentence complexity. Mean (*SD*)

	SLI (*n* = 15)		AM (*n* = 15)		LM (*n* = 15)	
	Simple	Complex	Simple	Complex	Simple	Complex
TLST						
Set length 2	47.8 (28.1)	58.9 (24.3)	81.0 (17.1)	72.6 (22.3)	54.4 (29.2)	53.3 (29.7)
Set length 3	37.8 (22.1)	23.7 (16.7)	60.3 (18.9)	44.4 (24.7)	46.7 (27.8)	37.0 (25.1)
Set length 4	27.2 (17.1)	17.2 (20.8)	50.0 (25.3)	44.6 (26.1)	38.3 (22.4)	31.7 (23.6)
Set length 5	22.7 (12.0)	24.2 (18.8)	36.2 (25.0)	41.9 (21.6)	30.2 (17.4)	22.2 (25.4)
ALST						
Set length 2	60.0 (25.0)	60.0 (23.4)	79.8 (27.9)	84.5 (24.9)	67.8 (30.5)	65.6 (25.6)
Set length 3	31.8 (18.7)	32.6 (14.8)	53.2 (30.6)	59.5 (28.4)	51.1 (27.2)	52.6 (34.2)
Set length 4	28.9 (20.6)	27.8 (14.0)	55.4 (19.2)	51.8 (27.4)	38.9 (27.4)	52.8 (26.8)
Set length 5	25.8 (15.1)	25.8 (9.0)	44.8 (22.9)	38.1 (25.4)	40.0 (22.1)	30.7 (16.7)

Notes. SLI = children with specific language impairment; AM = age-matched controls; LM = language-matched controls; TLST = traditional listening span task; ALST = active listening span task.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ANOVA results showed a significant main effect of task type, $F(1, 41) = 11.57$, $p = <.01$, partial $\eta^2 = .22$, with better performance on the ALST ($M = 48.30$, $SE = 2.85$) compared to TLST ($M = 41.85$, $SE = 2.56$); nearly significant main effect for complexity, $F(1, 41) = 3.94$, $p = .054$, partial $\eta^2 = .09$, with higher scores on simple ($M = 46.25$, $SE = 2.59$) than complex sentences ($M = 43.9$, $SE = 2.62$); main effect for set length, $F(2, 90) = 107.63$, $p < .001$, partial $\eta^2 = .72$; and main effect for group, $F(2, 41) = 5.99$, $p = <.01$, partial $\eta^2 = .23$. Significant interactions were observed for Task Type × Complexity, $F(1, 41) = 6.69$, $p = <.05$, partial $\eta^2 = .14$ and Group × Set Length, $F(6, 80) = 2.33$, $p = <.05$, partial $\eta^2 = .10$.

Pairwise comparisons of groups indicated that overall performance of children with SLI differed significantly from AM children ($p < .01$) but not from LM children ($p = .33$). Comparisons of set lengths revealed that all set lengths differed significantly from each other ($p < .01$ across comparisons), with poorer performance on each increase in set length, as expected.

Examination of the Task Type × Complexity interaction revealed that task types differed significantly on complex ($p < .001$) but not simple sentences ($p = .09$). This finding was consistent with Hypothesis 2, which predicted greater benefit from LTM support specifically on complex sentences. The nonsignificant Group × Complexity interaction, $F(2, 41) = .14$, $p = .87$, further indicated that groups were affected similarly by changes in complexity. Based on these results, performance patterns for simple sentences are not discussed further in remaining analyses.

To test Hypothesis 1, which predicted better performance on ALST compared to TLST across groups, we conducted pairwise comparisons of task types within each group (complex sentences only). Results revealed that task types differed significantly for the LM group ($p < .001$) and approached significance for the AM group ($p = .056$), but did not differ for the SLI group ($p = .146$). Based on these results, Hypothesis 1 was partially supported for the control groups, but not for children with SLI.

Traditional Versus Active Listening Span: Error Analysis

To test Hypothesis 3, we first analyzed errors in repetition of sentence-final words, then examined errors in children's responses to content questions. We differentiated three error types in sentence-final word repetition but focused our analysis on the two more relevant types of errors: (1) interference errors, in which children incorrectly repeated a previously correct response or provided nontarget words from the sentence context; and (2) intrusion errors, in which children responded with words that had never been presented. We also compared the total number of different words provided by the children during the ALST, because the use of fewer different words in any group would have increased the likelihood of interference errors for these children. Descriptive data for these analyses are presented in Table 3.

A one-way ANOVA comparing total number of different words among groups revealed no between-group differences, $F(2, 42) = 1.73$, $p = .19$. The frequency of each error type was then entered into a mixed ANOVA comparing task type (traditional vs. active), error type (interference vs. intrusion), and group (SLI, AM, LM).

Results revealed a significant main effect of error type, $F(1, 41) = 64.74$, $p < .001$, partial $\eta^2 = .61$, with more interference ($M = 9.06$, $SE = 1.01$) compared to intrusion errors ($M = 1.70$, $SE = .23$); significant main effect of task type, $F(1, 41) = 4.57$, $p < .05$, partial $\eta^2 = .10$, with more errors on the TLST ($M = 6.03$, $SE = .59$) compared to ALST ($M = 4.74$, $SE = .70$); and significant main effect of group, $F(2, 41) = 7.54$, $p < .01$, partial $\eta^2 = .27$. Significant interactions were observed for Error Type × Task Type, $F(1, 41) = 5.09$, $p < .05$, partial $\eta^2 = .11$; and Error Type × Group, $F(2, 41) = 3.73$, $p < .05$, partial $\eta^2 = .15$.

Pairwise group comparisons indicated that the number of errors for children with SLI significantly exceeded errors for AM ($p < .01$) and LM children ($p < .01$) but that the TLD groups did not differ from each other ($p = 1.0$). Examination of the error Type × Task Type interaction

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3. Listening span error analysis. Descriptive statistics for total number of different words in ALST and total number of errors by error type, task, and group. Mean (*SD*)

	SLI (*n* = 15)	AM (*n* = 15)	LM (*n* = 15)
WM: TLST			
Interference	15.20 (8.17)	7.36 (6.99)	8.13 (3.83)
Intrusion	3.40 (3.09)	1.14 (1.35)	0.93 (0.88)
WM: ALST			
Interference	12.67 (11.02)	6.36 (6.89)	4.67 (4.62)
Intrusion	2.73 (4.73)	1.07 (1.21)	0.93 (1.22)
Different words	67.87 (5.08)	70.33 (5.21)	66.93 (5.22)
LTM: TLST/ALST			
Interference	11.27 (7.08)	7.67 (2.82)	8 (2.73)
Intrusion	4.87 (2.72)	2.8 (1.97)	2.93 (2.94)
Omissions	5.47 (4.96)	5.87 (4.85)	6.87 (7.85)

Notes. SLI = children with specific language impairment; AM = age-matched controls; LM = language-matched controls; TLST = traditional listening span task; ALST = active listening span task; LTM = long-term memory.

 © 2014 Hogrefe Publishing

further showed that the effect of task type was significant for interference ($p < .05$) but not for intrusion errors ($p = .68$). Consistent with accuracy data, within-group comparisons revealed that the effect of task type on the number of interference errors was significant for LM children ($p < .05$) but not for AM ($p = .53$) or SLI ($p = .11$) groups.

To analyze children's answers to content questions, we conducted four individual one-way ANOVAs to compare the total number of (1) correct responses, (2) interference errors, (3) intrusion errors, and (4) omission errors among groups. Results revealed significant between-group differences for correct responses, $F(2, 42) = 4.89$, $p < .05$, and for interference errors, $F(2, 42) = 7.78$, $p = .001$. Group effects on number of intrusion errors approached significance, $F(2, 42) = 3.02$, $p = .06$, and were not significant for omissions, $F(2,42) = .21$, $p = .8$.

Pairwise group comparisons indicated that overall comprehension performance of children with SLI was poorer than that of AM peers ($p < .05$) but comparable to that of LM children ($p = .07$); and that the SLI group demonstrated significantly more interference errors than both TLD groups, SLI-AM: $p < .01$, SLI-LM: $p < .01$.

Traditional Versus Modified Nonword Repetition

Preliminary comparison of overall accuracy (% words correct) for nonwords containing high- versus low-frequency meaningful syllables using a paired sample *t*-test showed no significant difference between the nonword sets, $t(41) = -1.55$, $p = .129$. High and low frequency stimuli were therefore combined for subsequent analyses. Comparisons of three- and four-syllable nonwords showed significant differences between lengths within each task type (traditional and modified NWR), $t(43) = 2.23$, $p < .05$; $t(42) = 4.14$, $p < .001$, respectively. The distinction between nonword stimuli based on length was therefore maintained in subsequent analyses.

Accuracy Data

The first analysis compared mean accuracy for groups, nonword lengths, and task type in a 3 × 2 × 2 mixed ANOVA. Pairwise comparisons were all performed with Bonferroni adjustments. Descriptive data are presented in Table 4.

Results revealed a significant main effect for task type, $F(1, 40) = 18.74$, $p < .001$, partial $\eta^2 = .32$, with higher accuracy on modified ($M = 54.09$, $SE = 2.85$) compared to traditional NWR ($M = 43.12$, $SE = 2.5$); main effect of length, $F(1, 40) = 13.0$, $p < .001$, partial $\eta^2 = .25$, with higher accuracy on three- ($M = 52.94$, $SE = 2.63$) versus four-syllable nonwords ($M = 44.27$, $SE = 2.64$); and main effect of group, $F(2, 40) = 23.67$, $p < .001$, partial $\eta^2 = .54$. Pairwise comparisons revealed that the SLI group differed from AM and LM groups ($p < .001$ for both comparisons), but AM and LM children did not differ from each other ($p = .44$).

There was a significant Group × Task Type interaction, $F(2, 40) = 3.64$, $p = .035$, partial $\eta^2 = .15$, indicating that the wordlikeness effect differed among groups. Additional comparisons of task types within groups indicated that task types differed significantly for the LM group ($p < .001$), with higher accuracy on modified ($M = 65.83$, $SE = 4.83$) compared to traditional NWR ($M = 45.42$, $SE = 4.17$). AM and SLI groups showed no change in performance ($p = .095$ and $p = .275$, respectively). These results partially support Hypothesis 4, which predicted better performance on modified compared to traditional NWR, but indicate that benefit derived on modified NWR was dependent on group.

Error Analysis

We differentiated between single errors, which contained a single phonemic error, and multiple errors, which contained two or more phonemic errors. Descriptive statistics are provided in Table 5. The total number of each error type was entered into a mixed ANOVA comparing task type (traditional vs. modified), nonword length (3- vs. 4-syllable), error type (single vs. multiple), and group (AM, LM, SLI). Results revealed significant main effects of task type, $F(1, 38) = 53.43$, $p < .001$, partial $\eta^2 = .58$, nonword length, $F(1, 38) = 15.97$, $p < .001$, partial $\eta^2 = .30$, group, $F(2, 38) = 23.24$, $p < .001$, partial $\eta^2 = .55$, and error type, $F(1, 38) = 5.84$, $p < .05$, partial $\eta^2 = .13$. Significant interactions were observed for Task Type × Group, $F(2, 38) = 10.09$, $p < .001$, partial $\eta^2 = .35$, Error

Table 4. Nonword repetition. Descriptive statistics for repetition accuracy (% correct) by group, list condition, and stimulus length (number of syllables). Mean (*SD*)

	SLI (*n* = 15)	AM (*n* = 15)	LM (*n* = 15)
Traditional NWR			
Three syllables	26.8 (16.9)	65.8 (21.9)	47.5 (24.2)
Four syllables	20.5 (10.5)	55.0 (14.8)	43.3 (22.6)
Modified NWR			
Three syllables	31.3 (20.7)	73.2 (19.7)	74.6 (18.5)
Four syllables	25.9 (20.0)	62.5 (21.5)	57.1 (23.2)

Notes. SLI = children with specific language impairment; AM = age-matched controls; LM = language-matched controls; NWR = nonword repetition.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 5. Nonword repetition error analysis. Descriptive statistics for total number and proportion of single and multiple errors by group and list type. Mean (*SD*)

	SLI (n = 15)		AM (n = 15)		LM (n = 15)	
	Number	Proportion	Number	Proportion	Number	Proportion
Traditional NWR						
Single errors	2.8 (1.1)	0.43 (0.18)	1.9 (0.7)	0.64 (0.22)	2.9 (1.2)	0.69 (0.19)
Multiple errors	3.4 (1.3)	0.57 (0.18)	1.2 (1.0)	0.36 (0.22)	1.4 (0.9)	0.31 (0.19)
Modified NWR						
Single errors	2.6 (0.9)	0.48 (0.18)	2.0 (1.1)	0.82 (0.18)	1.8 (0.9)	0.68 (0.15)
Multiple errors	3.1 (1.6)	0.52 (0.18)	0.5 (0.6)	0.18 (0.18)	1.0 (0.8)	0.32 (0.15)

Notes. SLI = children with specific language impairment; AM = age-matched controls; LM = language-matched controls; NWR = nonword repetition.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Type × Group, $F(2, 38) = 8.96$, $p < .001$, partial $\eta^2 = .32$, and Task Type × Nonword Length, $F(1, 38) = 8.88$, $p < .01$, partial $\eta^2 = .19$. The interaction between task type and error type approached significance, $F(1, 38) = 3.74$, $p = .06$.

Pairwise comparisons of error types within each group indicated that error types differed significantly from each other across groups ($p < .05$, $p < .01$, and $p < .01$ for SLI, AM, and LM groups, respectively); however, the predominant error type differed based on group. AM and LM children showed more single than multiple errors ($p < .01$ for both groups), whereas the SLI group showed more multiple than single errors ($p < .05$). Although the number of single errors was comparable for SLI and control groups, SLI-AM: $p = .132$; SLI-LM: $p = .682$, the number of multiple errors was significantly higher in children with SLI compared to AM and LM groups ($p < .001$, both comparisons).

Finally, the number of errors within each type varied within groups as a function of task type. For the SLI group, multiple errors exceeded single errors during traditional NWR ($p < .05$) but did not differ on modified NWR ($p = .175$). For AM children, error types did not differ on traditional NWR ($p = .225$) but single exceeded multiple errors on modified NWR ($p < .001$). For LM children, single exceeded multiple errors across task types (p = <.01 and $p < .05$, respectively).

Discussion

The goal of the present study was to examine the interaction between WM and LTM in children with and without SLI, based on evidence indicating that efficient retrieval from LTM facilitates WM performance (Casalini et al., 2007; Unsworth & Engle, 2007). Although it is well established that children with SLI show a deficit in WM across tasks and studies (Marton & Schwartz, 2003; Montgomery, 2000), it is unknown whether this deficit is limited to maintenance and updating of relevant items in WM or also involves difficulties in retrieving information from LTM. We compared children's performance on a traditional (TLST) and active listening span task (ALST). In the latter, target items consisted of words previously activated and retrieved from LTM. We also compared children's performance on a traditional NWR task and a modified task in which nonwords included one monosyllabic real word.

Findings on Listening Span Tasks

Overall results showed differences in performance accuracy between TLST and ALST (Hypothesis 1) and different performance profiles for the three groups. Performance on simple sentences was not significantly different in the two tasks, although there was a tendency for better performance in ALST. The benefit from modifications with complex sentences in ALST was clearly present (hypothesis 2).

All groups showed better performance on simple compared to syntactically complex sentences, which is consistent with developmental data and results from previous SLI studies (Marton & Schwartz, 2003; Montgomery & Evans, 2009). An increase in stimulus complexity requires more efficient allocation of attentional resources and efficient retrieval from LTM (Unsworth & Engle, 2007). The interaction between task type and sentence complexity suggests that as processing demands increased, children needed to support their WM more efficiently with long-term knowledge.

Overall group results showed that performance of children with SLI was similar to that of younger language-matched (LM) controls within tasks but differed from that of age-matched (AM) children, especially on the ALST. This finding is consistent with previous studies showing that children with SLI perform more poorly than age-matched peers on verbal span tasks, particularly as complexity increases (Marton & Schwartz, 2003; Montgomery, 2000). The comparison of ALST and TLST, however, further revealed that children with SLI did not show improvement on the ALST compared to TLST, indicating that these children did not support their WM performance with long-term knowledge as tasks became more complex.

One criticism toward using TLST with children with SLI is that children may not know some of the words presented as targets. One of our goals with the ALST

 © 2014 Hogrefe Publishing

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

was to overcome this problem by having the children finish incomplete sentences with their own words. These words existed in their vocabulary and had been retrieved at the time of sentence completion. If the problem in TLST was related to vocabulary knowledge, then performance in ALST should have improved. Our results demonstrate that even when children with SLI have the required knowledge, they have difficulty retrieving appropriate items according to task goals and using long-term knowledge to support WM performance. Our error analyses further suggest that one factor contributing to this deficit in children with SLI is poor resistance to interference.

The error analysis data for recall of sentence-final words show that children with SLI produced about twice as many interference errors as their typically developing peers (Hypothesis 3). Most of these interference errors consisted of items from previous trials. A potential explanation for the occurrence of this type of error during the ALST could have been that children with SLI produced a limited set of words when completing sentences. More frequent use of the same words would have increased the likelihood of interference errors. The number of different words provided by the children during the ALST, however, did not differ between groups. Thus, the larger number of interference errors in the SLI group reflects less efficient maintenance and updating of relevant items in WM. Children with SLI maintained items from previous trials that were not relevant in the current sets. A limitation of the present study is that the design did not allow us to differentiate proactive interference errors, in which children repeated previous responses, from distractor interference errors, in which non-target words from previous contexts are produced.

Our listening span tasks also included questions about sentence content, which measured efficiency of LTM retrieval. Children with SLI showed a greater number of interference errors in the answers to these questions, suggesting that these children's WM deficit is affected by difficulty retrieving relevant information from LTM. Results from the listening span tasks further indicate that interference negatively affected children's performance at both levels of the memory system (maintenance of information in WM and retrieval of relevant information from LTM) and that this negative effect was significantly larger in children with SLI than in either of the control groups.

In contrast to children with SLI, LM participants benefited most from modifications in ALST. This finding implies that although vocabulary may have contributed to task performance, the ability to resist interference, update WM contents, and retrieve relevant information from LTM played a more critical role in task performance, particularly in the ALST. An unexpected finding was that the difference in LM children's performance across span tasks was larger than that of the AM children. One explanation for this finding relates to strategy use and associated performance benefits. Compared to younger children, older children are more efficient in spontaneously using strategies to support their WM performance (Fatzer and Roebers, 2013). Children who use efficient strategies spontaneously do not benefit from external cues to the same extent as children who are less efficient in spontaneous strategy use (Kray, Kipp, & Karbach, 2009).

Findings on Nonword Repetition Tasks

Both the developmental and SLI literature suggest a relationship between vocabulary knowledge and NWR (Gathercole, Willis, Emslie, & Baddeley, 1992; Morra and Camba, 2009); however, most researchers have studied this relationship indirectly, by examining correlations between vocabulary size and NWR (Coady & Evans, 2008) or manipulating the frequency of phoneme combinations within nonwords (Munson, Kurtz, & Windsor, 2005). Casalini and colleagues (2007) provided the most direct examination by studying nonwords created from Italian morphemes. In the present study, we compared NWR on lists of words with no meaningful syllables to nonwords containing one monosyllabic real word. Our expectation was that children would repeat nonwords from the latter list more accurately because these stimuli facilitate the use of LTM to support NWR.

Overall results showed better performance for nonwords with meaningful syllables compared to nonwords with no meaningful syllables (Hypothesis 4). The accuracy results between lists with low- and high-frequency words did not differ, which was surprising. Recall of low- and high-frequency words typically differs due to the contribution of LTM to short-term memory (Roodenrys, Hulme, Alban, Ellis, & Brown, 1994). The lack of a difference in recall based on word frequency in our results may be related to several factors. First, all meaningful words were monosyllabic in length and were inserted into a nonword. Thus, meaning was embedded in a no-meaning context, which might have caused a masking effect. Word frequency also shows a correlation with phonotactic probability (Hulme, Stuart, Brown, & Morin, 2003). Phonotactic probability across the nonword lists in our study was similar; we only changed 1-2 phonemes in each word to manipulate meaning and frequency features.

The groups showed different profiles and there were several interactions among the independent variables. There was an overall group effect, with better performance in AM than LM children on the traditional NWR task but not on the modified task. Differences in the efficiency of spontaneous strategy use between these two groups may be responsible for this finding. In the traditional NWR task children had to rely on spontaneous strategy use, whereas the modified NWR task provided them with additional information. As in span tasks, LM children showed the most benefit from task modifications.

Children with SLI performed more poorly than the two control groups on both tasks. This finding is consistent with previous studies reporting poor NWR in this population (Marton & Schwartz, 2003; Montgomery & Windsor, 2007). The unexpected finding was that the children with SLI continued to show very low performance accuracy even when nonwords included meaningful words. These results further indicate poor interaction between LTM

© 2014 Hogrefe Publishing

and WM in these children. They either had difficulty recognizing the meaningful monosyllabic words within the nonwords or could not use long-term word knowledge as a scaffold to NWR.

The error analysis data (single vs. multiple) show a somewhat different pattern. Whereas children with SLI produced more multiple than single errors in the traditional NWR task, their data on the modified task yielded equal number of single and multiple errors. Thus, the proportion of single versus multiple errors changed with task modifications. This finding suggests that even if children with SLI did not repeat the entire nonword accurately, they made less complex errors in the modified than in the traditional NWR task. More refined examinations, including performance profiles and error pattern analyses, are needed in future studies to further elucidate these performance changes.

Error analysis data for the control groups showed that the number of single and multiple errors was equal for the AM group in the traditional NWR task but there were more single errors than multiple errors in the modified task. Thus, the benefit from LTM support resulted in simpler errors. The number of total errors in this group was significantly smaller than in the SLI group. The error patterns for the LM children did not change across tasks but their accuracy rate increased with task modifications.

Conclusions

The findings of this study show that the relationship between WM and LTM is complex in nature. Data from the two experiments revealed consistent patterns that indicate a strong interaction among age, language status, and stimulus complexity. Children with SLI showed consistently poor performance across tasks, indicating a weakness in using long-term knowledge to support WM performance. This study adds to the existing literature on the WM capacity deficit in children with SLI by showing that these children do not benefit from various manipulations designed to enhance WM performance via LTM support due to a combination of inefficiencies in maintaining and updating items in WM and retrieving information from LTM. These children showed a larger interference effect than the controls. The current findings suggest that children with SLI do not use efficient retrieval strategies to support their WM performance, in part because of their poor resistance to interference. These children may need direct strategy training to improve their resistance to interference at both levels of the memory system (LTM and WM). To develop such training methods, it is important that we further examine the nature of the deficit in resistance to interference by systematically investigating the factors that underlie proactive and distractor interference.

Task modifications had a smaller effect on the performance of older typically developing children compared to younger children, likely due to more efficient use of spontaneous strategies in the older children. Younger children, who are typically less efficient in using strategies spontaneously, showed the largest benefit from task manipulations that facilitated interaction between WM and LTM.

Acknowledgments

This study was supported by the National Institute on Deafness and Other Communication Disorders: "Working Memory Capacity in Children with SLI," R03DC41449; "The Impact of Inhibition Control on Working Memory in Children with SLI," 1R15DC009040-01; and by the European Union and the European Social Fund, TÁMOP: "Assessment of cognitive functions in children with different disabilities."

References

Alloway, T. P. (2009). Working memory, but not IQ predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Assessment, 25*, 92–98.
Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *Test of nonverbal intelligence*. Austin, TX: Pro-Ed.
Burgess, N., & Hitch, G. (2005). Computational models of working memory: Putting long-term memory into context. *Trends in Cognitive Sciences, 9*, 535–541.
Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, S., Pecini, C., & Burani, C. (2007). Non-word repetition in children with specific language impairment: A deficit in phonological working memory or in long-term verbal knowledge? *Cortex, 43*, 769–776.
Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*, 1–40.
Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech, Language, and Hearing Research, 36*, 1051–1054.
Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MN: American Guidance Service.
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.
Fatzer, S. T., & Roebers, C. M. (2013). Language and executive functioning: Children's benefit from induced verbal strategies in different tasks. *Journal of Educational and Developmental Psychology, 3*, 1–9.
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. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*, 887–898.
Hall, W. S., Nagy, W. E., & Linn, R. L. (1984). *Spoken words, effects of situation and social group on oral word usage and frequency*. Mahwah, NJ: Erlbaum.
Hulme, C., Stuart, G., Brown, G. D., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language, 49*, 500–518.
Kail, R. (2000). Speed of information processing: Developmental change and links to intelligence. *Journal of School Psychology, 38*, 51–61.
Kray, J., Kipp, K. H., & Karbach, J. (2009). The development of selective inhibitory control: The influence of verbal labeling. *Acta psychologica, 130*, 48–57.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

 © 2014 Hogrefe Publishing

Lloyd, M. E., Doydum, A. O., & Newcombe, N. S. (2009). Memory binding in early childhood: Evidence for a retrieval deficit. *Child Development, 80*, 1321–1328.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24*, 56–98.

Magimairaj, B., Montgomery, J., Marinellie, S., & MacCarthy, J. (2009). Relation of three mechanisms of working memory to children's complex span performance. *International Journal of Behavioral Development, 33*, 460–469.

Marton, K., Campanelli, L., Eichorn, N., Scheuer, J., & Yoon, J. (in press). Information processing and proactive interference in children with and without language impairment. *Journal of Speech, Language, and Hearing Research*.

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.

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

Montgomery, J. W. (2000). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 293–308.

Montgomery, J. W., & Evans, J. L. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*, 269–288.

Montgomery, J. W., & Windsor, J. (2007). Examining the language performances of children with and without specific language impairment: Contributions of phonological short-term memory and speed of processing. *Journal of Speech, Language, and Hearing Research, 50*, 778–797.

Morra, S., & Camba, R. (2009). Vocabulary learning in primary school children: Working memory and long-term memory components. *Journal of Experimental Child Psychology, 104*, 156–178.

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

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General, 134*, 368–387.

Roodenrys, S., Hulme, C., Alban, J., Ellis, A. W., & Brown, G. D. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition, 22*, 695–701.

Schleepen, T. M., & Jonkman, L. M. (2009). The development of non-spatial working memory capacity during childhood and adolescence and the role of interference control: An n-back task study. *Developmental Neuropsychology, 35*, 37–56.

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

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (4th ed.) (CELF-4). San Antonio, TX: Psychological Corporation.

Spaulding, T. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 725–738.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition, 41*, 242–254.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104–132.

Was, C. A., & Woltz, D. J. (2007). Reexamining the relationship between working memory and comprehension: The role of available long-term memory. *Journal of Memory and Language, 56*, 86–102.

Klara Marton

The Graduate Center of the City University of New York
365 Fifth Avenue, Rm. 7107
New York, NY 10016-4309
USA
Tel. +1 212 817-8802
Fax +1 212 817-1537
E-mail kmarton@gc.cuny.edu

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

© 2014 Hogrefe Publishing