JSLHR

Article

Nonword Repetition: The Relative Contributions of Phonological Short-Term Memory and Phonological Representations in Children With Language and Reading Impairment

Judith Rispens[a] and Anne Baker[a]

Purpose: This study investigates the relative contributions of phonological short-term memory and phonological representations to nonword repetition (NWR). This was evaluated in children with specific language impairment (SLI) and/or reading impairment (RI); it was also studied from a developmental perspective by comparing 2 groups of typically developing (TD) children who differed in age.
Method: NWR, digit span, vocabulary, and word and nonword discrimination were tested in 2 groups of TD children: one group matched on chronological age (CA TD group: *n* = 41, mean age = 7;8 [months;years]), and one language age–matched control group (LA TD group: *n* = 16, mean age = 5;8). Also, 10 children with SLI, 14 children with RI, and 23 children with SLI and RI (hereafter, SLI + RI) participated and were matched to the age of the CA TD group.
Results: For the TD children, NWR was predicted by discrimination, digit span, and age. The interaction between discrimination ability and age was also significant. Children with SLI + RI were significantly impaired on NWR compared with all other groups. A regression analysis, including the CA TD group and the children with SLI and/or RI, showed that digit span, discrimination ability, and group (SLI + RI) contributed significantly to NWR.
Conclusions: Phonological short-term memory and phonological representations both significantly contribute to NWR. The predictive strength of the quality of phonological representations changes during development.

Key Words: nonword repetition, specific language impairment (SLI), reading impairment, language development

Recent research has demonstrated that performance on a task of nonword repetition (NWR) is associated with vocabulary (Gathercole & Baddeley, 1989; see also Gathercole, 2006, for an overview) and literacy (Conti-Ramsden & Durkin, 2007). NWR has also been discussed extensively in the context of disordered language populations. Children with specific language impairment (SLI) often display poor NWR performance (Bishop, North, & Donlan, 1996; Ellis Weismer et al., 2000), as do children with reading impairment (RI; Bishop, McDonald, Bird, & Hayiou-Thomas, 2009; Catts, Adlof, Hogan, & Ellis Weismer, 2005; de Bree, Rispens, & Gerrits, 2007).

An important question here is: What skills are involved in repeating nonwords? NWR was initially viewed as a measure that was relatively free from lexical influences, as nonwords are not part of the lexicon. However, lexical and/or sublexical factors such as phonotactic frequency (e.g., Edwards, Beckman, & Munson, 2004; Vitevitch & Luce, 1999) and the similarity between an existing word and a nonword (known as the *wordlikeness factor;* Gathercole, 1995) have been shown to facilitate NWR. This suggests that established phonological/lexical representations are activated during encoding of new verbal stimuli (see also Archibald & Gathercole, 2006). Furthermore, multiple linguistic and cognitive skills are involved in NWR, including encoding, temporary storage, retrieval, and articulation of a phonological

[a]University of Amsterdam, Amsterdam Center for Language and Communication

Correspondence to Judith Rispens: J.E.Rispens@uva.nl

Editor: Janna Oetting
Associate Editor: Marc Joanisse

Received September 23, 2010
Revision received March 17, 2011
Accepted September 3, 2011
DOI: 10.1044/1092-4388(2011/10-0263)

sequence (Snowling, Chiat, & Hulme, 1991). Coupled with the assumption that lexical knowledge is also tapped to some extent during NWR, it is not surprising that different hypotheses exist as to the abilities underlying NWR.

An influential view is that NWR performance primarily reflects the capacity of phonological short-term memory (STM), even though lexical processes may play a role (Gathercole, 2006). It is assumed that phonological short-term memory is crucial for the development of vocabulary (Baddeley, Gathercole, & Papagno, 1998; Gathercole, Hitch, Service, & Martin, 1997), and it has been found that NWR predicts vocabulary growth in young children (e.g., Gathercole & Baddeley, 1989, 1990). This finding has led to the hypothesis that the ability to repeat a novel word—in a sense, the first step involved in learning a new word—greatly depends on the capacity of the phonological short-term memory. For SLI, which is a developmental language impairment that often implies delayed vocabulary development, it has been claimed that a reduced capacity in phonological short-term memory underlies poor NWR performance (see, e.g., Gathercole, 2006).

In contrast to the idea that NWR depends on phonological short-term memory, the *lexical restructuring hypothesis* claims that growth of vocabulary predicts NWR, with phonological awareness mediating this relation (Bowey, 1996; Metsala & Walley, 1998). With the expansion of vocabulary, lexical representations emerge in which all phonemes are specified. This promotes access to sublexical phonological units such as syllables and phonemes. According to Metsala (1999), fully specified lexical representations underlie NWR in two ways. First, existing phonological templates support the creation of novel phonological representations. Second, more detailed lexical representations lead to increased flexibility in rearranging individual phonemes in new patterns. In typically developing (TD) children, growth in lexical structure measured over the course of 1 year accounted for significant variance in growth of NWR independent of digit span, thus demonstrating the importance of long-term lexical knowledge, in the form of highly specified phonological representations for NWR (Metsala, Stavrinos, & Walley, 2009).

The proposed associations among vocabulary size, phonological representations, and NWR may also explain why children with SLI and children with RI often display poor NWR. Munson, Kurtz, and Windsor (2005) showed that children with SLI, with smaller vocabulary sizes than their age-matched TD peers, did not differ in NWR performance from younger TD children matched on vocabulary size. Maillart, Schelstraete, and Hupet (2004) demonstrated that children with SLI had more difficulty in detecting small phonemic changes within a word than did TD children. The authors suggested that children with SLI have underspecified phonological representations that may therefore interfere with NWR ability. Similarly, Boada and Pennington (2006) showed that children with RI have less mature phonological representations and, most important, demonstrated that phonological awareness correlated with NWR. These results thus suggest that vocabulary size and fine-grained phonological representations are associated with NWR and that a deficit in phonological specification may interfere with NWR performance.

In summary, the lexical restructuring hypothesis stresses the interaction of NWR with phonological representations and sublexical units within the mental lexicon, in contrast to the phonological short-term memory hypothesis. On the basis of the available evidence, it seems plausible that both hypotheses are correct to a certain extent and may, in fact, be complementary. The purpose of the present study was to determine the relative contributions of long-term verbal knowledge to NWR in the form of phonological representations on the one hand and phonological short-term memory on the other.

Furthermore, it is not necessarily the case that these relative contributions remain constant across development. For example, nonwords that are of low phonotactic frequency (i.e., in which the sequence of phonemes does not occur frequently in a given language) are relatively more difficult to repeat for children than for adults (Edwards et al., 2004; Munson, 2001). Edwards et al. (2004) concluded that individuals with larger vocabularies have increasingly flexible phonological representations from which they are able to make more phonological generalizations that benefit the repetition of nonwords.

Thus, it may be the case that the quality of phonological representations becomes less predictive for NWR performance over time. When children grow older, not only does their vocabulary size increase, but their phonological representations within their lexicon become more specified and robust. This increases the chance that phoneme sequences of these representations can be used flexibly for repeating or learning new words. At a young age, the interindividual differences in vocabulary size and in the quality of the already established phonological representations within the mental lexicon may be greater than at an older age. By comparing two TD groups, 7- and 5-year-old children, on precisely the same tasks, it is possible to gain some insight into such developmental changes.

Third, it is not yet clear whether the relative contributions of phonological representations and phonological short-term memory to NWR ability in TD children are the same in children with language and reading problems. Poor NWR performance is common in SLI, but recent studies have shown that an NWR deficit is more specifically associated with the presence of RI in children with SLI. In fact, it has been reported that children with SLI but without RI perform like their TD peers (Bishop et al., 2009; Catts et al., 2005; Rispens & Parigger, 2010). Even so, NWR ability in SLI may be dependent on whether children have already started to learn to read and write. Bishop et al. (2009) assessed children with SLI on NWR ability at the age of 9 years and found that children with only SLI (and no reading problems) performed normally. However, the results were different from those obtained at the age of 4 years, when there was no difference in NWR ability between SLI-only children and children with SLI plus reading problems (in retrospect). At that age, both groups scored more poorly than did the TD children. A study conducted by de Bree, Wijnen, and Gerrits (2010) also showed that children with SLI who were 4 years of age—and, therefore, preliterate—made significantly more errors on a task of NWR compared with their age-matched TD peers. These results suggest that the acquisition of orthography facilitates NWR (Bishop et al., 2009). Children with poor orthography cannot profit from the relatively strong phonological skills that usually result from a normal development of orthography. However, it is interesting to note that children with RI—and, thus, with poor orthography—do not necessarily perform poorly on NWR. Several studies have documented that children with RI perform similarly to their TD peers (Boada & Pennington, 2006; Larkin & Snowling, 2007), whereas other studies report poor NWR in children with RI (Bishop et al., 2009; Catts et al., 2005). Thus, there is mixed evidence on NWR ability in children with RI and in children with SLI. It is important to note that it has not yet been researched whether the contribution of variables such as phonological short-term memory and the quality of phonological representations to NWR is the same for children with SLI and/or RI compared with TD children.

Testing the independent contributions of phonological short-term memory and the quality of phonological representation to NWR is, of course, difficult, as both skills are related to one another. Testing phonological short-term memory using, for instance, a digit span task involves using linguistic (i.e., lexical) material, which draws on long-term memory resources. This implies that during a phonological short-term memory task, lexical representations are activated and, thus, that the results on a phonological short-term task are also related to phonological representations. In contrast, in order to develop and retain any phonological sequence for at least a minimal amount of time, phonological short-term memory is necessary, as this is part of the process of forming a phonological representation. Thus, the quality of phonological representations is related—to a certain extent—to phonological short-term memory. Furthermore, when testing both skills, abilities such as lower level auditory processing and articulation (in the case of a digit span task) are involved. The results of such measures thus do not reflect solely the capacity of phonological short-term memory, nor do they reflect the quality of phonological representations. Phonological short-term memory and phonological representations are, of course, not similar "constructs" despite the fact that both skills are usually tapped in such tasks.

In summary, in this study we compared the associations among phonological representations, phonological short-term memory, and NWR in children with SLI, children with RI, and children with SLI and RI (hereafter, SLI + RI). These groups were compared with two groups of TD children: one group matched on chronological age (CA) and one group matched on the vocabulary age of the children with SLI. This vocabulary age–matched TD group is approximately 2 years younger than the children with SLI and served as a language age–matched control group (LA TD group). This group also serves as a comparison group for the children with RI, as the children in this group were approximately 2 years behind in word decoding and were roughly at the same literacy level as the children in the LA TD group.

The main objective of this study was to evaluate whether the contributions of phonological representations and phonological short-term memory to NWR change with development and whether they are different across TD and SLI and/or RI populations. By comparing different types of children of different ages on the same tasks, we hoped to gain more insight than could be gained from studies in which that focus has been on only one aspect (e.g., phonotactic frequency in Munson et al., 2005) or on one type of population (e.g., TD children in Metsala et al., 2009).

Method

Participants

Five groups of participants took part in this experiment: children with SLI only, children with SLI + RI, children with RI only, TD children approximately 8 years of age, and TD children approximately 5 years of age.

Children With SLI and RI

Children with RI. Fourteen children with RI (see Table 1) participated. They all attended regular schools and were contacted via special needs teachers. They were selected on the criterion that they had poor reading and spelling skills but normal oral language development. Some children were diagnosed with developmental

Table 1. Overview of participant characteristics, including reading scores.

Group	*N*	Age Range	Age *M*	Age *SD*	RWT *M*	RWT *SD*	PWT *M*	PWT *SD*
CA TD	41	7;0–8:9	7;8	0;5	12.1	2.5	12.0	2.8
LA TD	16	5;4–6;4	5;8	0;3				
SLI + RI	23	7;0–8;10	8;0	0;2	2.8	1.9	4.3	2.2
SLI only	10	7;2–8;10	8;1	0;2	10.2	1.9	10.8	2.2
RI	14	7;1–9;7	8;3	0;7	4.9	2.7	4.5	1.5

Note. Ages are expressed in years;months. Blank cells indicate data not applicable. RWT = real-word test; PWT = pseudoword test; CA = chronological age; TD = typically developing; LA = language-matched control group; SLI = specific language impairment; RI = reading impairment.

dyslexia; others were in the process of diagnosis. To ascertain the presence of poor reading skills, we administered two tasks: a real-word reading task (RWT; Brus & Voeten, 1973) and a pseudoword reading task (PWT; van den Bos, Spelberg, Scheepstra, & de Vries, 1994). In the RWT, the child is required to read aloud a list of existing words as quickly and accurately as possible within 1 min. The PWT follows the same principle but uses nonwords and takes 2 min. In the Netherlands, it is common to use timed reading tests to detect reading difficulties, as speed is a better indicator of reading development than accuracy alone (De Jong & van der Leij, 2003). The raw score (number of items read correctly) can be converted to a standardized score with a mean (*M*) of 10 and an *SD* of 3. Reading problems are present when children score more than 1 *SD* below the mean on both the RWT and the PWT. The present experiment included only those children who scored at least 1 *SD* below the mean. None of the children with RI had been diagnosed with oral language problems, and none of the children had ever been referred to a speech-and-language therapist for concerns about their oral language development.

Children with SLI. In the first instance, 33 children diagnosed with SLI (M_{age} = 8;0 [years;months]) were selected from three special needs schools for children with developmental language disorders. SLI was diagnosed (a) when a child performed at least 1.5 *SD* below the mean in at least two language domains measured with Dutch standardized language tests or (b) when a child performed more than 2 *SD*s below the mean on a Dutch standardized general language test. The testing for the diagnosis was carried out by a multidisciplinary team of the special needs schools, and an SLI diagnosis was a requirement in order for a child to be accepted by the special needs schools. Children who had evidence (or a history) of speech output problems, such as dyspraxia, were excluded from participation in the present study, as such problems interfere with the reliability of NWR scoring. All children had normal hearing and normal or corrected-to-normal vision. Only those children who were raised in a family with at least one parent who was a native speaker of Dutch and where Dutch had been spoken from birth onward were accepted. The special needs schools were located in two different regions (north and middle) of the Netherlands.

In the second instance, the 33 children with SLI were screened for the presence of RI using the two reading tasks: PWT and RWT (see the *Children with RI* subsection above). From the total number of 33 children, 23 children scored at least 1 *SD* below the mean. They were characterized as *SLI + RI*. Ten children with SLI scored within normal limits on the RWT and the PWT tasks and were characterized as *SLI only* (see Table 1).

We were interested in dividing the group of children with SLI into two groups based on the presence/absence of RI because RI may have an impact on NWR performance, as described in the introduction. Our aim was to categorize the SLI-only children based on their reading performance but to keep the severity of the oral language impairment as comparable as possible. It was not possible to statistically compare the background language scores, as different diagnostic tests had been used with the individual children. However, the two groups of SLI children were able to be compared on two language measures: One is a standardized receptive vocabulary measure, the Dutch version of the Peabody Picture Vocabulary Test (PPVT; Schlichting, 2005), which is used in the present study as a variable (see Tasks section below), and the second is a past-tense production task that was administered to the same children for a different study (Rispens & de Bree, 2010). Problems with the production of the past tense are characteristic of SLI (Bishop, 1997; Leonard, 1998). Thus, it seems suitable to use results on past-tense marking seem as a background measure for evaluating language ability among the children. The past-tense task elicited past tenses of regular and irregular verbs and of nonsense verbs (similar to a Wug task; Berko, 1958). On both tasks, there were no statistical differences between the SLI children with RI and without RI (receptive vocabulary: raw score $p = .45$; percentiles $p = .67$; past tense $p = .22$), but all children performed significantly poorer than the TD children (Rispens & de Bree, 2010).

TD children. Two groups of TD children participated. One group of children matched with the SLI-only children, the SLI + RI children, and the RI-only children on CA; these children are referred to as the *CA TD children*. A second younger group of children matched with the SLI-only children and the SLI + RI children on their mean score on a receptive vocabulary task—the Dutch version of the Peabody Picture Vocabulary Test (PPVT;

Schlichting, 2005). These children are referred to as the *language age–matched (LA TD) children.* The results of the PPVT are presented in Table 3.

The CA TD children ($n = 41$, $M_{age} = 7;8$) were selected from four primary schools located in the north, middle, and southwest regions of the Netherlands. All children were in second grade, made normal progress in school, had normal hearing and normal or corrected-to-normal vision, and did not have any cognitive or emotional problems. All children were from families in which at least one parent was a native speaker of Dutch and in which Dutch was the language spoken at home from birth onward.

The LA TD children ($n = 16$, $M_{age} = 5;8$) were all attending the second year of kindergarten. As mentioned above, they were selected so that we could investigate any developmental change in the relative contributions of long-term phonological representations and phonological short-term memory to NWR. They also functioned as an LA TD control group for the SLI-only and SLI + RI groups. Because the group of RI children was approximately 2 years behind TD peers on their reading level, the LA TD group also functioned as a group roughly matched for reading age. Informed consent was obtained from the parents of all participants.

Tasks

Three standardized tasks were used, and three experimental tasks were constructed.

Standardized Task 1: Digit span. In order to evaluate the contribution of phonological short-term memory to NWR, we administered a standardized digit span task: the Digit Span Forward subtest of the Dutch version of the Wechsler Intelligence Scale for Children—Revised (Van Haasen et al., 1986). Children were presented with a series of numbers (auditorily) and were asked to repeat them in the correct order. For each digit length (three to eight digits), there were two separate items. If the child failed to repeat two items in the same category correctly, the test was broken off. For each item repeated correctly, a child was awarded 1 point. The raw scores were converted to standard scores according to the normative data provided by the test manual.

Standardized Task 2: Vocabulary. Vocabulary size was investigated because the lexical restructuring hypothesis predicts a correlation between vocabulary size and NWR performance. A receptive vocabulary task, rather than an expressive task, was used because the majority of studies have demonstrated that receptive measures correlate more strongly with NWR performance (Coady & Evans, 2008). The Dutch standardized version of the PPVT (Schlichting, 2005) was administered to the children.

Standardized Task 3: Raven. We used a measure of nonverbal ability, the Raven Standard Progressive Matrices (Dutch version; Raven, 2006), to examine the influence of general cognitive ability on NWR performance. The results were used as a measure to control for the influence that differences in nonverbal intelligence between the children may have on NWR performance. The raw scores were translated into the standardized percentile scores. The LA TD children were not assessed with this task because the Dutch norm scores are available only from age 6;6 forward.

Experimental Task 1: NWR. The NWR task was the central task and was designed for this study. It comprised 40 items divided equally into two-, three-, four-, and five-syllable items; see the Appendix for the stimuli used in this task. To control for the influence of phonotactic probability, we ensured that half the items were low in phonotactic probability and the other half were high in phonotactic probability, based on the Dutch Phonotactic Frequency database (Adriaans, 2006) derived from the corpus of spoken Dutch (Oostdijk, 2000). The items did not contain consonant clusters and followed the Dutch rules of prosody. The items, prerecorded by a female native speaker of Dutch, were played back using a laptop computer. The items were divided into two blocks of 20 items each that were presented to the children with a pause between the blocks. The order of the two blocks was pseudorandomized. All responses of the children were recorded and transcribed offline. For each item, the percentage of phonemes correctly repeated per word was calculated because this is a more sensitive scoring method than scoring the number of words repeated accurately (Graf Estes, Evans, & Else-Quest, 2007).

Experimental Task 2: Word discrimination (minimal pairs). A task of lexical discrimination was constructed as a measure of the specificity of phonological representations in the mental lexicon. The assumption that underlies this task is that the more detailed phonological representations are, the better children will discriminate lexical minimal pairs.

Forty-eight nouns were selected, and each noun was paired with a nonword that differed by one phoneme from that noun (e.g., *tijger–dijger* [tiger–diger]). In half the items, the target phoneme was in initial word position (*tijger–dijger* [tiger–diger]), and in the other half, the target phoneme was in medial word position (*toeter–toeser* [horn]). All items were highly imageable and had an early age of acquisition (below the age of 6 years; Schaerlaekens, Kohnstamm, & Lejaegere, 1999). Half of all items were low in frequency (log frequency between 0 and 0.5 based on the CELEX database; Baayen, Piepenbrock, & Gulikers, 1995); the other half were high in frequency (log frequency > 1.5). Six filler items were presented twice in the correct form. Children were shown a picture and then were asked the question, "*Is dit een … ?* [Is this a … ?]" followed by the target item. Thus, all pictures were presented twice in a pseudo-randomized order throughout the task. All minimal pair members were presented once in the correct form and once in a mispronounced form. Children had to press a yes/no button to indicate their response to the question. Each response was classified as a *hit, false alarm, miss,* or *correct rejection.* We used signal detection analysis to calculate the mean d' score (see Swets, 1996). In the present study, a perfect d' sensitivity score is 4.65.

Experimental Task 3: Nonword discrimination (minimal pairs). This task measures discrimination between minimal pairs of bisyllabic nonwords. For an accurate performance, children need to create new phonological representations with a detailed specification so that a minimal phonological deviation can be identified. This task is assumed to measure the degree of specificity of phonological representations. Twenty-four nonword pairs consisting of two syllables differing from one another by one phoneme (onset position of second syllable) were constructed. Half the items differed by two phonological features from the target phoneme, the other half by one phonological feature. Examples of two minimal pairs are *deesep–deetep* [de:sɛp–de:tɛp] and *voobur–voonur* [vo:bʏr–vo:nʏr]. Twenty identical nonword pairs served as filler items and as a check to see whether children were able to perform the task. The items were prerecorded and played back to the children from a laptop computer. The children were asked to indicate whether they heard the same or a different word, pressing a *yes* or *no* button on the computer keyboard. The mean percentage of correct responses was calculated.

Procedure and Data Analysis

All children were seen individually during two sessions at their schools, except for one child (RI-only subgroup) who was seen at her home. All measures were straightforward to score (either a yes/no response, a sequence of numbers, or pointing to a picture) except for the NWR task. All responses on the NWR task were transcribed and scored by at least two people. The people carrying out the transcriptions and the scoring included a linguist, a trained speech-and-language therapist, and two advanced master's students of linguistics who were trained in transcribing and scoring. All transcriptions and scores were compared between the two raters. In the majority of cases, there was high agreement between the scores. In case of disagreement or doubt, both raters would listen together to the recording and decide together how the response was scored.

To evaluate group differences, we used an analysis of variance followed by post hoc analyses (Games–Howell due to unequal variances and unequal sample sizes), and significance was set at $p < .05$. The word and nonword discrimination tasks were both designed to measure the quality of phonological representations, and the performances on these tasks correlated (Pearson's $r = .56$, $p < .001$). For the present study, we used a single-factor score to represent discrimination ability derived from a principal components analysis. This yielded a single factor explaining 78.9% of the variance and an eigenvalue of 1.56. The factor score is used in the statistical analyses instead of the two scores of the two separate tasks.

To investigate the associations between NWR and the other tasks, we carried out Pearson's product–moment correlation analyses, controlling for the effect of group. Two separate multiple regression analyses were carried out. The first included the two groups of TD children, which enabled us to examine the developmental perspective. In the second regression analysis, we included the CA TD children and the children with SLI and RI (SLI only, SLI + RI, and RI only) to examine what variables contribute to NWR and whether this is different in the presence of SLI and RI. We created dummy variables representing group status to test whether group status predicts NWR and whether group status interacts with the predictor variables. To do so, we calculated the product terms between the dummy variables and the predictor variables significantly explaining variance of NWR. In the regression analyses, we entered age and the Raven score (not for the analysis including the LA TD group, as the Raven could not be administered to them) as a first step to remove variance solely attributed to general developmental level.

Results

In Table 2, reliability data are reported for the three experimental tasks: NWR, word discrimination, and nonword discrimination. We carried out the reliability analyses for the two TD subgroups and for the whole sample. As can be observed, all measures were reliable (Cronbach's $\alpha > .7$).

Table 3 provides an overview of the *M*s and *SD*s of the children on all tasks. It is striking that the children with SLI + RI do far worse, in general, not only on the NWR but also on most of the other tasks. The SLI-only children and RI-only children did not differ significantly from

Table 2. Reliability figures for the three experimental tasks.

Group	NWR	Word discrimination	Nonword discrimination
CA TD	$\alpha = .84$	$\alpha = .79$	$\alpha = .71$
LA TD	$\alpha = .91$	$\alpha = .78$	$\alpha = .81$
Whole sample	$\alpha = .92$	$\alpha = .90$	$\alpha = .82$

Table 3. Mean performance, with *SD*s, and group comparison.

Task	Group									
	SLI + RI		SLI only		RI only		CA TD		LA TD	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
NWR	72.9[a,b,c**d*]	5.9	82.5	5.9	85.3	5.9	88.8	5.4	82.1[a*]	8.6
PPVT (%)	23.9[a,b,c**]	20.9	27.3[a**b,c*]	20.9	51.1	18.9	64.7	21.8	57.1	26.0
PPVT (raw score)	89.7[a,c**]	10.2	92.4[a,c**]	8.9	103.7	10.5	104.7	9.6	81.8[a,c**]	16.7
Digit span (standardized score)	4.2[a,b**]	1.9	4.8[a,b**]	1.4	4.6[a,b**]	1.7	7.0	2.3	7.5	1.9
Raven (%)	41.6[a*]	7.1	57.3	24.1	56.9	29.0	66.3	24.5		
Discrimination	−1.2b[*a,c,d**]	1.0	−.03[a*]	0.5	.29	0.4	.64	0.4	−.29[a**]	0.8

Note. NWR is expressed in percentage of phonemes correct; discrimination is represented by the factor score of the word and nonword discrimination task. NWR = nonword repetition; PPVT = Peabody Picture Vocabulary Test.

[a]Significantly lower than CA TD controls. [b]Significantly lower than LA TD controls. [c]Significantly lower than the RI group. [d]Significantly lower than the SLI-only group.

*$p < .05$. **$p < .01$.

the CA TD children on the NWR. The relationships among the scores are discussed in more detail below.

Predicting NWR: Typical Development

To gain insight into a possible change of variables contributing to NWR during development, we carried out a correlation analysis and a multiple regression analysis for the two groups of TD children.

A partial correlation analysis with the combined sample of LA TD and the CA TD groups, but controlling for the effect of group, demonstrated that NWR correlated significantly with digit span ($r = .52, p < .01$), vocabulary ($r = .29, p < .01$), and the discrimination factor score ($r = .49$, $p < .001$). We carried out a regression analysis and created a dummy variable (0 = *second grade,* 1 = *kindergarten*) to investigate whether there was a significant contribution of age and to test whether any of the predictor variables interacted significantly with age. In other words, using the dummy variable allowed us to examine the effect of being a member of the kindergarten group in relation to the predictor variables and relative to the children who were in second grade. First, we entered age and then the digit span and vocabulary scores together with the discrimination factor score, stepwise, in the second block. In the the third step, we entered the dummy variable, and in the final step, we entered the interactions between the dummy and predictor variables.

Table 4 shows the results of the multiple regression analysis. The final regression model was significant, $F(5, 58) = 13.5$, $p < .001$, and explained a total of 56% variance (R^2). The regression model showed that in the first three steps, the discrimination factor score (31%, $p < .001$) and digit span (10%, $p < .001$) accounted for a significant amount of the total sum of squares in the NWR data, just like the dummy variable (5%, $p = .01$). The final model—including the block with the Vocabulary × Dummy Variable interaction, the Digit Span × Dummy Variable interaction, and the Discrimination Factor Score × Dummy Variable interaction—accounted for a significant amount of the total sum of squares in the NWR data (4%, $p = .03$). The Discrimination Factor Score × Dummy Variable interaction was significant ($p = .03$). The Digit Span × Dummy Variable interaction ($p = .32$) and the Vocabulary × Dummy Variable interaction ($p = .23$) were not significant.

This regression analysis shows that group status significantly predicts NWR and that the strength of the predictors is different for the LA TD children, as the discrimination factor score interacted with group status. Two correlation analyses carried out separately for the two TD groups shows that the discrimination factor score only marginally correlated with NWR for the CA TD children ($r = .29$, $p = .058$), whereas the discrimination factor score strongly correlated with NWR for the LA TD children ($r = .73$, $p = .001$). This finding

Table 4. Multiple regression analysis predicting NWR in the groups of CA TD children and the LA TD children (represented by the dummy predictor).

Predictor	Final β	*p*
Age	−.22	.24
Digit span	.34	.003
Discrimination	.20	.18
Vocabulary	.07	.51
Dummy	−.42	.03
Dummy × Discrimination	.33	.03

suggests that NWR performance is more strongly related to discrimination ability in younger children than in older children. In addition, there was no significant correlation between vocabulary and NWR for the CA TD group ($r = .16$, $p = .32$), whereas NWR was significantly associated with vocabulary in the LA TD group ($r = .51$, $p < .05$).

In summary, both vocabulary and discrimination ability are significantly related to NWR in 5-year-old children, but these relationships were not significantly related to NWR (or, to a lesser extent, were not present) in 8-year-old children.

Predicting NWR: Children With SLI and RI

As a first step, we computed partial correlations between NWR and the other scores for the combined sample of CA TD children and for the children with SLI only, SLI + RI, and RI only, controlling for the effect of group. NWR correlated with all measures (digit span, $r = .57$, $p < .001$; Raven, $r = .23$, $p < .05$; Discrimination factor score, $r = .38$, $p < .001$) except for the receptive vocabulary score ($r = .11$). Because the PPVT score correlates significantly with NWR within the combined sample (not controlling for the effect of group, $r = .43$, $p < .01$), we decided to enter the vocabulary score as a predictor variable in the regression analysis.

We carried out a multiple regression analysis to determine the relative contributions of digit span, vocabulary, and the discrimination factor score to NWR and to evaluate these for the groups of children with SLI only, SLI + RI, and children with RI only relative to the CA TD children. We created three dummy variables: Dummy 1 representing the children with SLI + RI, Dummy 2 representing the children with SLI only, and Dummy 3 representing the children with RI only. The contributions of these dummy variables are relative to the reference category "no impairment and age-matched" (the CA TD children).

We carried out the multiple regression analysis with NWR as the outcome variable. Predictors were chronological age and the Raven score in a first separate block and, in a second block digit span, vocabulary and the discrimination factor score, which were entered stepwise. In the third block, we entered the three dummy variables (Dummy 1: SLI + RI; Dummy 2: SLI only; Dummy 3: RI only) and added, in the final block, the product terms of all the dummy variables with all the predictor variables.

The regression model was significant, $F(16, 87) = 11.5$, $p < .001$, and explained 71% of the variance in NWR (R^2) in total. We found that, after considering the effects of age and nonverbal IQ, digit span (32%) and the discrimination factor score (14%) were significantly related to NWR. When, in the next step, we added the dummy variables, we demonstrated that the Dummy 1 variable (representing the SLI + RI group) also accounted for a significant amount of the total sum of squares (10%, $p < .001$). The interactions between the dummies and the predictors did not make significant contributions to the regression model ($\Delta F = 3\%$, $p = .45$). Thus, even though group status (SLI + RI), relative to the CA TD children, significantly predicted NWR, the regression effects of the predictors are not significantly different for the three groups relative to the CA TD children. Table 5 presents the results of the significant regression model.

Discussion

We performed this research to investigate the relative contributions to NWR of phonological short-term memory and phonological representations. We wanted to evaluate how these aspects are associated with NWR performance, whether these associations change with development, and whether they are the same for children with LI and children with RI.

First, the regression analysis in which the younger group of TD children and the CA TD children were included showed that both digit span and discrimination ability significantly predicted variance in TD children. Second, group membership also significantly predicted NWR. It is important to note that in this sample, there was a significant Group Membership × Discrimination Ability interaction. This interaction indicates that the variable "discrimination" behaves differently as a predictor variable for the two groups. In this case, specifically, the interaction suggests that in children who are 5 years of age, the variable "discrimination" is more strongly associated with NWR than in children who are 8 years of age. Indeed, a correlation analysis demonstrated that for the younger group of children (the LA TD group),

Table 5. Results of the regression analysis predicting NWR in the combined sample of the CA TD group and the SLI + RI (Dummy 1), SLI-only (Dummy 2), and RI-only (Dummy 3) groups.

Predictor	Final β	*p*
Age	−.12	*ns*
Raven	.03	*ns*
Digit span	.28	< .001
Discrimination	.30	.04
Vocabulary	−.002	.79
Dummy 1	−.77	< .001
Dummy 2	−.10	.20
Dummy 3	−.03	.73

NWR is more strongly associated with discrimination ability than for the older group of children (the CA TD group). For the latter group, the significance of the correlation between NWR and discrimination is only marginal. Thus, phonological short-term memory plays a significant role in NWR performance for the group of 5-year-old children, but it appears that the role of phonological representations is a better predictor of NWR at age 5 years than at age 8 years. This finding confirms earlier research suggesting that lexical restructuring is relatively important at ages 5 and 6 years (Storkel, 2002). Previous investigations have, in fact, shown that lexical representations of children 5 and 6 years of age are less specified in comparison with 8-year-olds and adults (Walley, Metsala, & Garlock, 2003). Thus, the difference in the two TD groups may stem from the fact that at ages 5 and 6 years, lexical restructuring is still in progress so that more individual variance in discrimination ability will be related to NWR than at ages 7 and 8 years. By then, more lexical items will have been fully specified. There seems to be a developmental shift between the ages of 5 and 8 years whereby the quality of phonological representations becomes relatively less predictive for NWR ability. This may also explain the lack of correlation between NWR and vocabulary in the CA TD group. Of course, our results are based on a cross-sectional design and not on a longitudinal study, which is more suitable for investigating the developmental track of skills underlying NWR and possible developmental changes in the strength of these contributors. To consolidate our findings, a longitudinal study should be carried out. Furthermore, the strength of the predicting variables may also depend on the type of nonwords used. As discussed previously, our items were balanced for high and low phonotactic probability. It is possible that the LA TD children had particular problems with the items low in phonotactic probability compared with the CA TD children, as the LA TD children have less lexical resources at their disposal. We are presently investigating the interaction between growth of vocabulary size and the ability to repeat nonwords low in phonotactic probability in different age groups.

Turning to the results of the children with SLI and/or RI, we observed that not all children with SLI and RI had profound difficulty with NWR. Only the children with SLI + RI were significantly impaired on the NWR task, even more so than the LA TD children. In contrast, the SLI-only children and RI-only children performed similarly to the age-matched TD children.

Furthermore, a regression analysis showed that both digit span and discrimination ability were significantly associated with NWR in the combined sample of the CA TD children and in the three groups of children with SLI + RI, SLI only, and RI only. Digit span explained the greatest portion of variance (32% unique variance); the discrimination factor score also contributed significantly (14%), but less than digit span. Furthermore, the dummy variable representing the children with deficits in both language and reading (the SLI + RI group) was a significant predictor of NWR performance relative to the CA TD children. However, it is important to note that there were no significant interactions between the predictor variables and group status.

The results underline the finding that both short-term memory capacity (measured with a digit span task) and detailed phonological representations (represented by word and nonword discrimination ability) are independently important for NWR. At ages 7 and 8 years, it seems that phonological short-term memory is more predictive than phonological representations. Even though NWR was significantly predicted by group status (SLI + RI children compared with CA TD children), there were no significant Digit Span × Group Membership and Discrimination Ability × Group Membership interactions.

The SLI + RI children scored particularly poorly on NWR—even more so than the LA TD group—and also demonstrated poor scores on the digit span (< LA TD) and on discrimination ability (< CA TD, SLI only, and RI). The other two subgroups (SLI only and RI only) were as impaired as the SLI + RI group on the digit span task but scored significantly higher on discrimination ability. Thus, their better NWR performance may be facilitated by their relatively stronger phonological representations, which may have compensated for their limited phonological short-term memory capacity. The finding that the only significant group predictor of NWR was the one corresponding to an impairment in both language and reading fits the suggestion that a double deficit is responsible for poor NWR (see also Newbury, Bishop, & Monaco, 2005).

We did not find evidence in any of the groups that vocabulary size explained significant variance in NWR. This finding is not in line with the lexical restructuring hypothesis and some previous research (Gathercole, Willis, Emslie, & Baddeley, 1992; Metsala et al., 2009; Munson et al., 2005). However, vocabulary may have shared much of its variance with the discrimination tasks that measured specificity of phonological representations, which, according to the lexical restructuring hypothesis, is the result of vocabulary size (Metsala et al., 2009). This may have resulted in no significant independent variance of vocabulary explaining NWR.

Thus, our results demonstrate that phonological short-term memory and phonological representations significantly contribute to NWR but that the relative strength of these predictors depends on age. Furthermore, the same associations among NWR, phonological short-term memory, and phonological representations were found for children with SLI and RI relative to their CA TD peers. A deficit in phonological short-term memory combined with underspecified phonological representations has a severe impact on NWR. However, normal levels of the latter may facilitate NWR.

Acknowledgments

We thank all children and their parents for participating in this study, and we thank the staff at their schools for their help. Furthermore, we thank Wendy Boelhouwer, Christa Kerkhof, and Martine Jong for their assistance in collecting the data. Finally, we thank Elise de Bree and Marcel Giezen for their comments on earlier versions of this article and Titia Benders for her help with the statistical analysis.

References

Adriaans, F. (2006). PhonotacTools (Test version) [Computer program]. Utrecht, the Netherlands: Utrecht Institute of Linguistics OTS.

Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research, 49,* 970–983.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database Release 2, Dutch Version 3.1 [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

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

Berko, J. (1958). The child's learning of English morphology. *Word, 14,* 150–177.

Bishop, D. V. M. (1997). *Uncommon understanding*. London, United Kingdom: Psychology Press.

Bishop, D. V. M., McDonald, D., Bird, A., & Hayiou-Thomas, M. E. (2009). Children who read words accurately despite language impairment: Who are they and how do they do it? *Child Development, 80,* 593–605.

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.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95,* 153–193.

Bowey, J. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. *Journal of Experimental Child Psychology, 63,* 44–78.

Brus, B. Th., & Voeten, M. J. M. (1973). *Een-Minuut Test. Vorm A en B* [One-minute test. Form A and B]. Nijmegen, the Netherlands: Berkhout Testmateriaal.

Catts, H., Adlof, S. M., Hogan, T. P., & Ellis Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48,* 1378–1396.

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 and Communicative Disorders, 43,* 1–40.

Conti-Ramsden, G., & Durkin, K. (2007). Phonological short-term memory, language and literacy: Developmental relationships in early adolescence in young people with SLI. *Journal of Child Psychology and Psychiatry, 48,* 147–156.

de Bree, E., Rispens, J. E., & Gerrits, E. (2007). Non-word repetition in Dutch children with (a risk of) dyslexia and SLI. *Clinical Linguistics & Phonetics, 21,* 935–944.

de Bree, E., Wijnen, F., & Gerrits, E. (2010). Non-word repetition and literacy in Dutch children at-risk of dyslexia and children with SLI: Results of the follow-up study. *Dyslexia, 16,* 36–44.

De Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology, 95,* 22–40.

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.

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. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513–543.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28,* 200–213.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of children learning new names. *British Journal of Psychology, 81,* 439–454.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33,* 966–979.

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.

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50,* 177–195.

Larkin, R. F., & Snowling, M. J. (2007). Phonological skills and spelling abilities in children with reading and language impairments. *International Journal of Language and Communication Disorders, 43,* 111–124.

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

Maillart, C., Schelstraete, M., & Hupet, M. (2004). Phonological representations in children with SLI: A study of French. *Journal of Speech, Language, and Hearing Research, 47,* 187–198.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91,* 3–19.

Metsala, J. L., Stavrinos, D., & Walley, A. C. (2009). Children's spoken word recognition and contributions to phonological awareness and nonword repetition: A 1-year follow-up. *Applied Psycholinguistics, 30,* 101–122.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Munson, B. (2001). Phonological pattern frequency and speech production in children and adults. *Journal of Speech, Language, and Hearing Research, 44,* 778–792.

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 language impairments. *Journal of Speech, Language, and Hearing Research, 48,* 1033–1047.

Newbury, D. F., Bishop, D. V., & Monaco, A. P. (2005). Genetic influences on language impairment and phonological short-term memory. *Trends in Cognitive Sciences, 9,* 528–534.

Oostdijk, N. (2000, May–June). The Spoken Dutch Corpus. Outline and first evaluation. In M. Gravilidou, G. Caravannis, S. Markantonatou, S. Piperidis, & G. Stainhaouer (Eds.), *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC)* (Vol. 2, pp. 887–894). Athens, Greece.

Raven, J. (2006). *Raven Standard Progressive Matrices* [Dutch Version]. Enschede, the Netherlands: Harcourt Test Publishers.

Rispens, J. E., & de Bree, E. (2010). Past tense productivity in Dutch children with SLI: The role of phonology. In K. Franich, K. Iserman, & L. Keil (Eds.), *BUCLD (Boston University Conference on Language Development) Proceedings* (Vol. 34, pp. 327–338). Boston, MA: Boston University.

Rispens, J. E., & Parigger, E. M. (2010). Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. *British Journal of Developmental Psychology, 28,* 177–188.

Schaerlaekens, A., Kohnstamm, D., & Lejaegere, M. (1999). *Streeflijst woordenschat voor zesjarigen. Derde herziene versie gebaseerd op nieuw onderzoek in Nederland en België* [Vocabulary list of 6-year-olds. Third revised version based on new research in the Netherlands and Belgium]. Lisse, the Netherlands: Swets & Zeitlinger.

Schlichting, L. (2005). *Peabody Picture Vocabulary Test—III NL*. Amsterdam, the Netherlands: Harcourt Test Publishers.

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

Storkel, H. L. (2002). Restructuring similarity neighborhoods in the developing mental lexicon. *Journal of Child Language, 29,* 251–274.

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Mahwah, NJ: Erlbaum.

van den Bos, K. P., Spelberg, H. C. L., Scheepstra, A. J. M., & de Vries, J. R. (1994). *De Klepel. Een test voor de leesvaardigheid van pseudo-woorden* [The Klepel. A test for the ability to read pseudo-words]. Nijmegen, the Netherlands: Berkhout Testmateriaal.

Van Haasen, P. P., de Bruyn, E. E. J., Pijl, Y. J., Poortinga, Y. H., Lutje-Spelberg, H. C., VanderSteene, G., . . . Stinissen, J. (1986). *Wechsler Intelligence Scale for Children—Revised, Dutch version*. Lisse, the Netherlands: Swets & Zeitlinger.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374–408.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing: An Interdisciplinary Journal, 16,* 5–20.

Appendix. The items and their phonetic transcriptions of the nonword repetition (NWR) task.

Orthography	IPA
Two-syllable items	
1. weugof	wø:xɔf
2. kuimup	kœymʏp
3. luubuf	lybʏf
4. joefeum	jufø:m
5. feusut	føsʏt
6. hiewam	hiwam
7. raanom	ranɔm
8. geeres	xerɛs
9. woosel	wosɛl
10. daanes	danɛs
Three-syllable items	
1. woezuunim	wuzynɪm
2. muihuuguf	mœyhyxʏf
3. soeguipem	suxœypɛm
4. nuigeusup	nœyxø:sʏp
5. veujoetup	vø:jutʏp
6. kaaroodin	karodɪn
7. voopeeket	vopekɛt
8. loowaamas	lowamas
9. taanoolon	tanolɔn
10. deevoenos	devunɔs
Four-syllable items	
1. meufuusuinef	mø:fysœynɛf
2. suijiegoonif	søyjixonɪf
3. juuvuigoowuf	jyvœyxowʏf
4. guiweusoegir	xœywø:suxɪr
5. fuiseuwoesut	fœysø:wusʏt
6. liejootaanig	lijotanɪx
7. peewaatoopes	pewatopɛs
8. liekoovoepar	likovupar
9. saaviebeemer	savibemɛr
10. kooviewaalan	koviwalan
Five-syllable items	
1. baamerienooves	bamɛrinovɛs
2. geerutievaanot	xerʏtivanɔt
3. tieloniedaanag	tilɔnidanax
4. wookaloemoodon	wokalumodɔn
5. beemonievoekes	bemɔnivukɛs
6. fuugiwuinoefep	fyxɪwœynufɛp
7. soegonuifeusir	suxɔnœyfø:sɪr
8. geumuwoekuubir	xø:mʏwukybɪr
9. nuijigeufuusut	nœyjɪxø:fysʏt
10. jeunimeusuifir	jø:nɪmø:sœyfɪr

Note. IPA = International Phonetic Alphabet.