Poor performance on the retention of phonemes' serial order in short-term memory reflects young children's poor reading skills

Kirsten Schraeyen[1,2,4], Astrid Geudens[1], Pol Ghesquière[2], Wim Van der Elst[3] and Dominiek Sandra[4]
[1]Thomas More University College / [2]University of Leuven / [3]Hasselt University / [4]University of Antwerp

The present study aimed to identify crucial factors that underlie phonological representations in short-term memory (STM) of third-graders with different literacy skills. For this purpose, we used the Nonword Repetition Task (NRT) to disentangle the processing of phonemes' identity and their serial order. We found no evidence that children's literacy skills are linked to their capacity for retaining phonemes' identity. However, their literacy skills are linked to their capacity for retaining phonemes' serial order. The latter link can be interpreted in terms of a domain-general STM mechanism but is also compatible with the impact of literacy on children's knowledge of the phonotactic regularities in a language.

Keywords: phonological representation, nonword repetition task, phoneme identity, serial order, short-term memory

Research in normally developing children has shown that adequate phonological processing skills are crucial to the development of proficient reading (Melby-Lervåg & Lervåg, 2011). In addition, it is widely acknowledged that less skilled readers have difficulties in processing phonological information (Hulme & Snowling, 1992; Stanovich & Siegel, 1994). Evidence supporting this claim comes from studies focusing on three distinct but interconnected aspects of phonological processing (Wagner & Torgesen, 1987). At first, it has been shown that less skilled readers demonstrate poorer *phonological awareness* (Pratt & Brady, 1988; Lyon, Shaywitz, & Shaywitz, 2003). This ability to manipulate sounds has been identified as a primary factor underlying early reading achievement (Ehri et al., 2001). Second, there is a substantial body of evidence showing a significant relation

The Mental Lexicon 12:1 (2017), **129–158.** DOI 10.1075/ml.12.1.06sch
ISSN 1871-1340 / E-ISSN 1871-1375

between *rapid automatized naming* (RAN), i.e., the capacity to retrieve phonological codes for digits, objects etc., stored in long-term memory (LTM), and reading performance (e.g., Aarnoutse, van Leeuwe, & Verhoeven, 2005; Clarke, Hulme, & Snowling, 2005, but see de Jong, 2011). Third, a relationship between reading performance and performance on *verbal short-term memory* (VSTM) tasks has also been shown, better reading skills being associated with superior VSTM. This is, for instance, the case in memory span tasks such as a nonword repetition task (NRT), in which nonwords[1] have to be recalled immediately from VSTM (e.g., Beneventi, Tonessen, & Ersland, 2009). VSTM plays a crucial role in learning to read, as it is a temporary storage place before the sequence of phonemes can be read aloud and is also needed before a word's phonological information can be transmitted to and represented in LTM.

Note that all three phonological skills that are crucial for reading success involve the storage and retrieval of phonological representations, either in STM or in LTM. This observation has led to the widely accepted assumption that less skilled readers' phonological processing problems are all due to a basic problem at the level of these phonological representations, more particularly, the representation of individual phonemes (Boada & Pennington, 2006; Carroll & Snowling, 2004; Elbro, 1996; Elbro, Borstrøm, & Petersen, 1998; Ramus & Szenkovits, 2008). More particularly, these representations are believed to be of poorer quality. As a direct consequence, sequences of phonemes (real words in RAN, nonwords in the NRT) are also more poorly represented in these readers' LTM and STM. Hence, words and nonwords, i.e., phonological representations, will cause problems when they must be retrieved from LTM and STM, respectively, and when phoneme awareness tasks require that manipulations be performed on them.

In this view, fully developed representations of phonemes are at the core of reading development, as high-quality representations of this type are a prerequisite for the establishment of associations between phonemes and graphemes. A variety of labels has been used to describe the problem of these qualitatively poor representations, such as *poorly specified* (Elbro & Jensen, 2005), *indistinct* (Elbro, 1998), or *less mature* (Boada & Pennington, 2006) phonemic representations. In sum, many researchers adhere to the assumption that poor performance on several experimental tasks reveals the existence of poor phonological representations in STM and LTM, which betray the presence of poor representations of the phonemes in a language, which in turn are also the cause of reading problems.

1. In the literature the term 'nonwords' is used to refer to items in a NRT-task. However, note that all NRT-items in this paper were pseudowords, i.e., items respecting the phonological constraints of Dutch.

The central question of this paper is whether the quality of phoneme representations is the only crucial factor to ensure adequate encoding, storage, and/or retrieval of phonological representations, or whether there are other phoneme-related factors that affect literacy skills. We will address this question by examining a group of young children with a wide range of literacy skills and investigating their performance in encoding and recalling the phonological representations of nonwords in STM, i.e., when performing a NRT. The NRT is a measure that has frequently been used to probe the quality of phonological representations in STM. In this task, individuals need to immediately repeat auditorily presented stimuli, more particularly, novel phoneme sequences differing in length and complexity. Consequently, a new phonological representation, i.e., a *sequence of familiar phonemes in a specific, unfamiliar serial order*, has to be formed and must be retained for a very short period of time in the phonological loop of STM before active retrieval takes place and an articulatory response is made (Munson, Swenson, & Manthei, 2005; Rispens & Parigger, 2010). If the adequate formation, storage, and/or retrieval of phonological representations in STM is closely related to reading success and the NRT is a good measure for assessing the nature of these representations, performance on this task should correlate with children's reading performance. This is indeed what many experiments have shown. Most published NRT studies yield a consistent outcome, with lower performance in high-risk children for reading problems (e.g., Carroll & Snowling, 2004; Snowling, Gallagher, & Frith, 2003; de Bree, Rispens, & Gerrits, 2007) as well as in less skilled and dyslexic young readers (e.g., Baird, Slonims, Simonoff, & Dworzynski, 2011; Robertson & Joanisse, 2010). Most often, in these studies, the overall accuracy score, i.e., the number of correct nonword reproductions, is used to demonstrate an association between the ability to form, temporarily store, and/or retrieve new phonological representations in STM on the one hand and reading performance on the other hand.

Even though these accuracy scores are an important piece of evidence that skilled reading is related to the ability to retain phonological information in STM they do not reveal *why* less skilled readers find it more difficult to repeat an unfamiliar phoneme string. As a novel nonword consists of a sequence of sounds, both a good retention of the individual phonemes' *identity* and a good retention of their *serial order* are important. It follows that either one of these factors (or both) may be responsible for worse NRT performance in less skilled readers: their dependence on phoneme representations of a poorer quality and/or a problem with the correct storage, retention, and/or retrieval of the serial order of these phonemes in the nonword sound string. In the former case, the problem is situated at the level of item information in STM (i.e., phoneme identity), whereas in the latter case it is situated at the level of serial order information in STM (i.e., phoneme order).

Obviously, the underachievement of less skilled readers on the NRT may be due to a combination of these problems.

A dissociation between the processing of item information and the processing of serial order information in STM has already received some attention (for a recent review, see Majerus & Cowan, 2016). Many recent STM models assume that the temporary storage of item information directly depends on the activation and quality of LTM representations, whereas a specialized STM system, i.e., an ordering mechanism in STM, is dedicated to the representation of serial order information and thus, less sensitive to the influence of LTM representations (Gupta, 2003). Evidence supporting the idea of a dissociation between item and serial order processing in STM comes from both behavioral and neuroimaging studies. Nairne and Kelly (2004), for instance, observed that phonological and semantic knowledge (LTM knowledge) had a positive effect on item retention but not on serial order retention. In their study, they found an increased item STM performance (less errors in item recall) but decreased serial order STM performance in tasks involving the recall of phonologically similar/semantically related words vs. phonologically dissimilar/semantically unrelated words. They concluded that the recall of verbal item information is affected by linguistic properties while serial order information is not (Nairne & Kelly, 2004). Other studies revealed a strong relationship between immediate serial recall (ISR), NRT performance, and word learning (see, for instance, Page & Norris, 2009; Cumming, Page, & Norris, 2003). The relationship between those three abilities was attributed to the existence of a sequence memory in STM, which is thought to be responsible for temporarily maintaining the serial order of lexical and sublexical information, i.e., the order of the units in the list (ISR) or the phoneme units in the nonwords (NRT) (see Gupta, 2003; Gupta, Lipinski, Abbs, & Lin, 2005). Using neuroimaging techniques, Majerus et al. (2008, 2010) demonstrated activation of phonological and semantic processing areas in the bilateral temporal lobes when verbal item information had to be recalled during STM-tasks. In contrast, the recall of serial order information was associated with the activation of nonlinguistic brain areas in the right intraparietal sulcus (Majerus et al., 2008; Majerus et al., 2010). The conclusion that these two STM processes are, to a certain extent, distinct has prompted researchers to study both types of retention in more detail, in an attempt to better understand the nature of verbal STM problems in the occurrence of reading problems or dyslexia.

Recent work by Martinez Perez, Majerus, and Poncelet (2012a) and Martinez Perez, Majerus, Mahot, and Poncelet (2012b) directly examined the dissociation between item and serial order processing and their independent contributions to reading acquisition (Martinez Perez et al., 2012a) and reading impairment (Martinez Perez et al., 2012b). To achieve this goal the authors disentangled retention of the particular items to be remembered (item STM capacity) from retention of the serial order in which these items had been presented (serial order STM capacity). To this end, they developed two different tasks. In a first study, kindergarteners were presented with an item STM-task in which they had to recall monosyllabic CVC-nonwords after a filled delay, i.e., the order of the recalled items was irrelevant. Additionally, the authors presented an order STM-task, in which these children had to listen to a sequence of animal names with high lexical frequency and a low age of acquisition, such as 'cat' and 'dog', before ordering pictures of the auditorily presented animal names, i.e., children did not have to recall the items. They found that order STM capacity, but not item STM capacity, predicted independent variance in decoding abilities for reading (Martinez Perez et al., 2012a). In a second study, in which they used the same tasks to compare dyslexic children with reading-age matched and chronological-age matched children with respect to both retention processes, they observed that children with dyslexia had an inferior performance on the item STM task compared to chronological-age matched controls, but not compared to reading-level matched controls. In the serial order STM task they observed an inferior performance of dyslexic children compared to both chronological-age and reading-age matched controls (Martinez Perez et al., 2012b). Using verbal and visuospatial material Hachmann et al. (2014) reported the same dissociation between STM performance for item and serial order information. Dyslexics had impaired serial order STM abilities but no impaired item STM abilities. They concluded that dyslexia is characterized by a selective impairment in the ability to retain items' serial order in STM.

Staels and Van den Broeck (2014) could not confirm this dissociation when attempting to replicate Martinez et al.'s (2012b) study with reading impaired children. These authors first conducted the same analyses and found similar results: inferior performance for children with dyslexia on the item STM and order STM tasks. Next, they conducted the same experiment with some methodological adjustments, arguing that 'a specific problem in serial order retention can only be proven by directly comparing dyslexic and typical individuals who score equally on the item retention task' (p 3). Using state trace analyses, they were unable to find a specific deficit in serial order STM in dyslexic children when these children and their controls were statistically matched on item retention performance. However, when they were matched on their performance on the serial order task, the effect of group remained significant for the item STM task. The authors concluded that the item and serial order tasks that were used by Martinez Perez et al. do not measure entirely distinct processes and that both are related to phonological processing, i.e., that there is no independent (and more important) serial order component (Staels & Van den Broeck, 2014).

When assessing the relative importance of item and serial order processing in STM with respect to reading performance, all aforementioned studies used *different* *tasks* to dissociate the retention of item information and serial order information. In addition, they focused on the level of digits, words, or syllables, not on the segmental or phoneme level. Only the repetition of the entire item or the serial order of entire items mattered. More particularly, no study paid attention to the retention of the phonemes making up the items.[2] This is strange, as literacy skills are especially strongly correlated with high-quality representations of individual phonemes (see above). In the present study we will focus on the relation between literacy skills and young children's abilities to retain the identity and serial order of the phonemes in an auditorily presented nonword. Moreover, we will study both types of retention skills by comparing them within items. This is a novel way to find out whether the correct retrieval of phonological representations from STM is only dependent on the quality of phoneme representations or is also related to a serial order mechanism. In order to accomplish this goal we focused on the retention of phonemes and their relative order in a novel phoneme string, i.e., a nonword. In other words, we used the NRT as our experimental task.

Given our focus on two types of retention performance the data will have to be analyzed in a different way than is usually the case. Indeed, the typical focus on the number of correct retentions collapses across the retention of a nonword's constituent phonemes and the retention of their serial order. This is because a correct response by definition implies that all phonemes are recalled in the correct serial order. Similarly, all errors, whether they are due to the loss of a phoneme's identity or its serial position, are treated equally. Only a study in which different error types are distinguished makes it possible to disentangle participants' retention of the phonemes themselves (item STM processing) from their retention of the serial order (serial order STM processing). This is the explicit aim of the current study: we want to assess whether participants' literacy skills predict the probability of making the two types of retention errors described above.

Recently, Schraeyen, Geudens, Ghesquière and Sandra (under review) followed this rationale in a study where adults with and without dyslexia performed a NRT. They compared both groups with respect to their item and serial order processing capacities when repeating three- to five-syllable nonwords. They distinguished three error response categories: (1) Responses containing only one or more phoneme identity errors (item errors); (2) Responses containing only one or more serial order errors (serial order errors); (3) Responses containing a combination of both error types. This choice was based on the above characterization of the NRT as a task in which individuals need to retain two types of phoneme-related information in their STM: their identity and their serial position. Results indicated that adults with and

---

2. This is, of course, due to the use of words or items with a familiar name (e.g., digits). Errors at the phoneme level are not likely to occur with such items.

without dyslexia did not differ in the retention of phonemes' identity, whereas adults without dyslexia outperformed adults with dyslexia with respect to serial order retention. In other words, dyslexics performed worse than controls on serial order STM capacity but not on item STM capacity. From this study, it could be concluded that serial order processing rather than item processing on the phoneme level is key for being able to form, store, and/or retrieve phonological representations in STM.

The main questions addressed in the current study are (a) whether these findings for adults will also obtain for children who are still in an early stage of reading development and (b) whether the dissociation can also be observed when treating Literacy as a continuous factor rather than a dichotomous one, i.e., without making a distinction between dyslexic readers and controls. If so, dyslexics' problem with the retention of phonemes' serial order is not a deficit, in the sense that it only occurs in this group of readers, but a property that co-varies with reading skill and is at its worst in very poor readers, i.e., dyslexics. If serial order retention of phonemes is independent from phoneme identity retention, and good serial order retention is a key characteristic of successful phonological processing in STM (as suggested by previous studies, see above), we expect that the dissociation between item and serial order STM processing should already be traceable in beginning readers and, moreover, become more pronounced as their literacy skills decrease along the continuous Literacy variable. More particularly, we expect that children's literacy skills will not be linked to the retention of phoneme identity in STM when performing a NRT task. In contrast, we expect that their literacy skills will be linked to the retention of phonemes' serial order in STM, i.e., that there will be a positive correlation between the two.

Method

Participants

A group of 89 Dutch-speaking children took part in this study (mean age: 104 months, SD: 6 months). They were recruited from several schools in northern Belgium (Flanders). All children were native speakers of Southern Dutch (Flemish) and had no documented hearing or speech difficulties. All children had been exposed to reading instruction from first grade in primary school onwards, i.e., they had received 20 to 21 months of explicit reading instruction.[3] We did not make

3. We were not able to collect participants' intelligence scores, even though such data would have removed extra variance from the data and, hence, have made it possible to more reliably estimate the effect of Literacy on the two retention variables.

a distinction between two reader groups, i.e., poor/dyslexic readers and typical readers, but used the children's oral reading fluency as a continuous predictor variable (henceforth: Literacy) in the statistical analyses. Our choice for this design was inspired by methodological concerns. The young age of the children did not always make it possible to rely on a formal diagnosis of dyslexia. Whereas some of the children with low(er) oral reading fluency had already been officially labeled as dyslexics, other children with similar oral reading fluency scores and/or a history of familial heredity had not been diagnosed yet. By using a continuous variable instead of a discrete one, we avoided misclassification of the children. At the same time this correlational design also allowed us to find out whether young readers' performance on the NRT is due to STM retention capacities (for phonemes' identity and serial order) whose strength varies along a continuum and is a direct function of the child's reading skill, i.e., also in the non-dyslexic range, rather than to STM retention capacities that are intact in non-dyslexic readers but severely impaired in dyslexic ones.

To determine children's literacy level we administered a standardized oral word reading test and a nonword reading task. The One Minute Test (OMT) (Brus & Voeten, 1997) measures children's oral word reading fluency. This standardized test requires children to read as many words as possible in one minute. Secondly, we used a standardized nonword reading task, the Klepel test (Van den Bos, Spelberg, Scheepstra, & de Vries, 1998). This standardized test requires children to read as many nonwords as possible in two minutes. The mean raw score for the performance on the oral word reading test was 44 (SD = 12; range 19–79) and 42 on the nonword reading task (SD = 15; range 16–77). Based on the high correlation (r = .86) between raw scores on the OMT and the Klepel test (see also Van den Broeck, Geudens, & Van den Bos, 2010), we decided to calculate a standardized composite score as our measure for Literacy.

Materials and design

Materials consisted of 36 nonwords (Appendix A), all of which were taken from the Flemish NRT test (Boets, 2006). Items in this NRT test were manipulated only for a few phonological factors. All items were phonologically legal in Dutch. These items were organized in three categories, varying in length from three to five syllables. All nonwords observed the Dutch pattern of placing stress on the penultimate syllable. Note that these manipulations were not related to theoretical issues. Of these 36 items, 18 nonwords consisted of a CVC final syllable and CV non-final syllables (e.g., /*pi.po.ket*/), whereas the other 18 nonwords consisted of CVC syllables only (e.g., /*lem.ros.pag*/).

Procedure

The NRT was individually administered. All multisyllabic nonword items were digitally recorded by a female native speaker of Dutch. These nonwords were presented once to the children through headphones. The children were told that they would hear a funny non-existent word and were asked to repeat it back immediately. We started with three-syllable items followed by four-syllable items and then five-syllable items. All children's responses were recorded (with Audacity software, version 1.2.4) and phonetically transcribed and scored by the experimenter and a second judge, using SAMPA (Speech Assessment Methods Phonetic Alphabet). An inter-rater agreement of 94.6% was found, based on the ratio of the number of identically transcribed phonemes on the total number of transcribed phonemes. Three decisions had to be made in response scoring: (a) whether the item was correctly repeated or not (nonword level score), and, if not, (b) whether retention problems were due to problems with phoneme identity or (c) with the serial order of the correctly recalled phonemes (phoneme level scores). In the majority of cases, there was also high agreement between the attributed scores. In case of disagreement, the experimenter and second judge decided together how the response was scored. The following scoring scheme was used.

Nonword level score
A score reflecting the child's nonword retention performance. A response was scored as 1 if the entire nonword was correctly reproduced, i.e., no errors were made at the word stress level, the syllable, or the phoneme level. For example, for the target word */sir.peg.wót.nal/*, the response */sir.peg.wót.nal/*, would be scored as 1. Any other response would be scored as 0, i.e., an error response.

Phoneme level scores
To assess children's ability to retain the identity and serial order of the phonemes in the nonword stimulus each response was given a binary score on two phoneme-related variables. These two scores reflected different aspects of the child's phoneme retention performance. Thus, four response types could be distinguished. (1) A correct response had a binary score of 1 on both variables. (2) A response with only identity errors was given a binary score of 0 on the dependent variable Phoneme Identity and a score of 1 on the dependent variable Serial Order. (3) A response with only serial order errors was given a binary score of 1 on the Phoneme Identity variable and a binary score of 0 on the Serial Order variable. (4) Finally, a response with identity and serial order errors was given a binary score of 0 on both the Phoneme Identity and Serial Order variables. In the following paragraphs we

describe these different error types (and their subtypes) in the responses and their classification.

1. *No identity or serial order errors*
   Correct response: each target phoneme appeared in the same serial order in the response, for example, target: /*sir.peg.wót.nal*/, response: /*sir.peg.wót.nal*/. Note that correct responses were also scored at these two phoneme-related variables, as all responses were included in the analyses of identity retention and serial order retention (see below);
2. *Phoneme identity errors*
   (a) Phoneme addition: one or more phonemes were added in the response, so, did not appear in the stimulus. However, all target phonemes appeared in the correct serial order in the response. For example, target: /*sir.peg.wót.nal*/, response: /*sir.pegs.wót.nal*/. Phoneme addition did not influence phoneme serial order as it did not disturb the relative position of the correctly recalled target phonemes (for the importance of relative position vs absolute position, see below). *(b) Phoneme substitution*: one or more phonemes in the target were replaced by other phonemes in the response, e.g., target: /*sir.peg.wót.nal*/, response: /*sir.peg.vót.nal*/, the target phoneme /w/ was replaced by the phoneme /v/. *(c) Phoneme omission*: one or more phonemes in the target were omitted in the response, e.g., target: /*sir.peg.wót.nal*/, response: /*sir.peg.wó.nal*/, the target phoneme /t/ was omitted in the response. *(d) Phoneme substitution in combination with phoneme omission*: e.g., target: /*sir.peg.wót.nal*/, response: /*sir.pek.wó.nal*/, the target phoneme /g/ was replaced by the phoneme /k/ and the target phoneme /t/ was omitted.
3. *Serial order errors*
   We applied the McKelvie scoring method to decide whether the serial order of phonemes was respected or not (McKelvie, 1987). In this method, phoneme strings of a least 2 phonemes are considered correct irrespective of their absolute position when counting from the beginning or the end of the nonword item. In addition, it is checked whether any other single phoneme is in the correct absolute serial position when counting either from the beginning or the end of the sequence, but not both. For example, for the target word /*sir.peg.wót.nal*/ and the response /*sir.pek.nót.wal*/, the response would be checked as follows: 11 phonemes out of 12 are correctly identified. The phonemes /sir.pe/, /ot/ and /al/ were positioned correctly. However, the phonemes /n/ and /w/ were swopped (between positions 7 and 9), violating the rule that single phonemes have to be in the correct absolute serial order when counting either from the beginning or the end of the sequence. *Phoneme shifts* were the only error type involving a serial order error but no identity error: each target phoneme appeared in the

response, but some phonemes were recalled at another (i.e., incorrect) serial position, for example, target: /*sir.peg.wót.nal*/, response: /*sir.peg.nót.wal*/, the target phonemes /n/ and /w/ are shifted in the response.

4. *Combination of identity and serial order errors*
   (a) Phoneme shift in combination with phoneme substitution: for example: target: /*sir.peg.wót.nal*/, response: /*sir.pek.nót.wal*/, target phonemes /n/ and /w/ are swopped (giving rise to a serial order error) and target phoneme /g/ is replaced by /k/ (giving rise to an identity error). *(b) Phoneme shift in combination with phoneme omission:* for example: target: /*sir.peg.wót.nal*/, response: /*wir.pe.sót.nal*/, target phonemes /w/ and /s/ are swopped (giving rise to a serial order error) and target phoneme /g/ is omitted (giving rise to an identity error). *(c) Phoneme shift in combination with a phoneme substitution and a phoneme omission:* for example, target: /*sir.peg.wót.nal*/, response: /*wir.pek.só.nal*/, target phonemes /w/ and /s/ are swopped (giving rise to a serial order error), target phoneme /g/ is replaced by /k/ and target phoneme /t/ is omitted (both giving rise to identity errors). *(d) All previous errors in combination with a phoneme addition.*

Table 1 provides an overview of the four different response types and their subtypes, together with their binary values on the two phoneme-related dependent variables. The four types reflect all possible combinations of the binary values on these variables.

Table 1. Overview of response types, their associated binary values with respect to identity and serial order performance, and their percentage in the total set of responses

Response type	**Identity performance**	**Serial order performance**	**%**
Correct response	1	1	32.0
Phoneme addition	0	1	2.2
Phoneme omission	0	1	4.5
Phoneme substitution	0	1	22.2
Phoneme substitution and omission	0	1	3.6
Phoneme shift	1	0	4.4
Phoneme shift and phoneme substitution	0	0	15.9
Phoneme shift and phoneme omission	0	0	2.2
Phoneme shift and phoneme substitution and phoneme omission	0	0	12.9

Note. Binary values are used to classify each response type with respect to identity and/or serial order performance.
Note. Responses combining phoneme identity and serial order difficulties were incorporated in the statistical analyses of both phoneme identity performance and serial order performance (see text). Combined errors (final three rows) could occur in combination with or without phoneme additions.

It was a deliberate decision to define Phoneme Identity and Phoneme Serial Order as dichotomous variables. Previous researchers used a binary coding system to investigate whether literacy is related to the probability of repeating a nonword without any error. They scored whether the stimulus was repeated with perfect phoneme identity performance and perfect serial order performance or not. We were interested in two similar questions, but targeted the relationship between literacy and retention performance at the phoneme level rather than the nonword level. (1) Was the stimulus repeated without any phoneme identity error (irrespective of serial order errors) or not? (2) Was the stimulus repeated without any serial order error (irrespective of phoneme identity errors) or not? Binary scores are typical in NRT tasks and the majority of STM tasks (e.g., Martens & de Jong, 2006; Grainger, 1990). At the conceptual level, our binary scores reflect whether the mechanism responsible for the retention of phoneme identity (or serial position) functioned flawlessly when repeating the nonword.

Statistical analyses will be applied on the responses at the nonword level (i.e., correct repetition of the nonword or not) and on the two dependent variables at the phoneme level: (a) Phoneme Identity and (b) Serial Order. Note that in the latter two analyses all responses (correct and incorrect ones) will be included. Each analysis addresses the question whether participants' literacy level determines the probability of correctly repeating all phoneme identities in the nonword and the probability of repeating the correctly remembered phonemes in the same relative serial order as in the stimulus. These analyses will enable us to assess the relative importance of the two types of STM retention processes in children with different literacy skills.

Statistical analyses

As a first exploratory analysis, to gain more insight in the level of (linear) association between the different measures, Pearson correlation coefficients were computed between the literacy scores and the three dependent variables: nonword level performance, phoneme identity performance, and serial order performance. In this analysis, total sum scores were used, i.e., all scores reflected the total number of correct or incorrect responses per child.

Next, a more fine-grained analysis was conducted in which the effect of the covariates of interest on the individual item responses (rather than on the aggregated total scores) was examined. To this end, a generalized linear mixed-effects model (GLMM) was fitted to the data. Let us define Y_ij as the jth outcome (i.e., item) for subject i, with $i = 1,\ldots,$ N and $j = 1,\ldots,$n_i. Further, Y_i is the n_i-dimensional vector of all measurements that are available for subject i. In the current study, $N = 89$

and $n_i = 36$. The model assumes that, conditionally on the random effects b_i (i.e., the random child effects) and b_j (i.e., the random item effects), which are assumed to be drawn independently from a multivariate normal with mean vector zero and unstructured variance-covariance matrix D), the outcomes Y_{ij} are independent with densities: $f_i (y_{ij} | b_i,\beta,\phi) = \exp\{\phi^{-1} [y_{ij} \theta_{ij}-\psi(\theta_{ij})] + c(y_{ij},\phi)\}$, with $\eta(\mu_{ij}) = x_{ij}' \beta + z_{ij}' b_i$ for a known link function $\eta(.)$, x_{ij} and z_{ij} are the p-dimensional and q-dimensional vectors of known covariates, respectively, β = a p-dimensional vector of unknown fixed-effects, and ϕ = a scale parameter.

The models were fitted by maximization of the marginal likelihood (which is obtained by integrating out the random effects). Details can be found in Molenberghs and Verbeke (2005). The generalized linear mixed-effects models that were initially fitted in our study included the following fixed effects: 'Literacy' (a continuous variable), Syllable Length (a variable that can take values 3, 4 and 5; this variable was dummy coded with 2 dummies and syllable length = 3 the reference category), the Literacy x Syllable Length interaction, and Phoneme Identity performance or Serial Order performance (both binary variables, one of which was taken up as a predictor variable to remove its correlation with the one that was used as the dependent variable). Participant and items were included as random effects (see Baayen, 2008) to appropriately account for the clustering in the data (i.e., items are clustered within children).

Next, it was evaluated whether the fixed-effects structure of the 'full' model (i.e., the model that included all the fixed and random effects) could be simplified. A hierarchical approach was taken in which it was first evaluated whether the two-way Literacy x Syllable Length interaction term could be removed from the model. In case the interaction term could indeed be removed, it was subsequently evaluated whether the main effect of Syllable Length could be removed from the model too. These tests were conducted using likelihood-ratio tests, i.e., the log-likelihood of the more elaborated model was compared with the log-likelihood of the simpler (nested) model. In particular, the −2 log likelihood difference of both models was computed and evaluated using a $\chi 2$ distribution with the difference in fixed-effect parameters of both models as the degrees of freedom (df) and $\alpha = .05$. A non-significant difference between both models implies that the simpler model has a similar fit to the data as the more complex model (i.e., that the model with an extra predictor does not explain significantly more variance), and, hence, that the simpler model can be retained. Once the final model was established, the model was refitted using Restricted Maximum Likelihood (REML) estimation to allow for better estimates of the variance components of the model.

In all analyses Literacy was kept in the model, irrespective of its level of significance. This was done because the main aim of the current study was to test the null hypothesis that literacy skills are not linked to identity retention/serial order retention. Recall that the experimental hypothesis is that literacy is not linked to the retention of phoneme identity but is linked to serial order retention (see Introduction). Further, in the analysis of Phoneme Identity the measure of Serial Order retention was included as a covariate. In the analysis of Serial Order, the measure of Phoneme Identity retention was included as a covariate. These covariates were included in the analyses to account for their (potential) confounding effects on the outcomes of interest, i.e., to take their effect under 'statistical control'. It is common practice to use such an approach in many statistical techniques including ANOVA, regression analysis, and (generalized) linear mixed-effects models (Verbeke & Molenberghs, 2000). All models were fitted using R (packages lme4, Bates, Maechler, & Dai, 2008; language R, Baayen, 2008).

Results

Correlational analyses

As can be seen in Table 2, there were high positive correlations between the performance at the levels of whole nonword retention, phoneme identity retention, and serial order retention (i.e., all rs ≥ .69, all *p*s < .0001).

Table 2. Pearson's correlation and levels of significance (two-sided tests) among the main variables of the study

		1	2	3
1.	Nonword Level Performance	–		
2.	Phoneme Identity Performance	.93**	–	
3.	Serial Order Performance	.77**	.69**	–
4.	Literacy	.22*	.14	.23*

* $p < .05$
** $p < .0001$

This result basically indicates that children who have higher (or lower) levels of nonword level performance also tend to have better (or poorer) phoneme identity and better (or poorer) serial order skills. In other words, children who are better in recalling entire nonwords make more responses in which all phonemes are recalled correctly (with or without serial order errors) and more responses in which the phonemes that have been correctly recalled are all placed in their correct (relative) serial order (i.e., with or without identity errors). The correlations between literacy and the other three outcomes were substantially lower. Literacy was significantly positively correlated with nonword level performance ($r = .22$, $p < .05$) and serial order performance ($r = .23$, $p < .05$), but not with phoneme identity performance ($r = .14$, $p > .05$). These correlations, which represent a rough first test of our hypothesis, are in line with the formulated expectations. The following analyses will be GLMM's, in which the effect of Literacy on the three measures of NRT performance is analyzed: Nonword performance, Phoneme Identity performance, and Serial Order performance.

Analysis of nonword level performance

For the nonword level performance, the results of the GLMM are summarized in Table 3. It turned out that the two-way interaction Literacy x Syllable Length was not significant ($p > .05$). Therefore, in the final model, the interaction term was removed.

Table 3. Estimated fixed effects of the generalized linear mixed model predicting nonword level performance from literacy and syllable length

Parameter	β	SE(β)	Z	P(>|z|)
(Intercept)	.50	.39	1.28	.20
Literacy	.14	.06	2.22	.03
Syllable Length 4	−2.04	.53	−3.86	.0001
Syllable Length 5	−3.51	.55	−6.42	< .0001

Note. Reference levels for the generalized linear mixed model are Syllable Length 3 and mean literacy score.

We observed a main effect of Literacy ($p = .03$) indicating that the probability of a correct response increased with increasing literacy scores. Note that this is the result that is typically obtained in NRT studies, i.e., this result validates our materials. In addition, aside from one's literacy score (i.e., no interaction) the probability of a correct answer was higher in three-syllable nonwords compared to four-syllable nonwords ($p = .0001$) and five-syllable nonwords ($p < .0001$), and the probability of a correct answer was higher in four-syllable nonwords compared to five-syllable nonwords ($\beta = -1.47$ for Syllable Length 5, $p = .007$) (the latter β estimate is obtained by fitting the same model as shown in Table 3, but now using syllable length 4 as the reference level instead of syllable length 3; data not shown). In line with this result, the mean nonword level item scores for three-, four- and five-syllable nonwords equaled .59, .26, and .12, respectively.

Analysis of phoneme level performance

The second set of analyses examined phoneme level performance. To this end, we analyzed the retention of (1) phonemes' identity, and (2) serial order. Table 4 presents the means error rates for the different response types for three-, four-, and five-syllable items. Recall that all responses were included in these two GLMM analyses, as each response was given two scores with respect to the phoneme level: whether all phonemes were recalled correctly or not, and whether all correctly recalled phonemes had been placed in the correct serial order.

Table 4. Mean error percentages (no phoneme error, identity error, serial order error, combined error) for each syllable length

	3-syllable items	4-syllable items	5-syllable items
No phoneme error (%)	59	26	12
Phoneme identity error (%)	36	69	85
Serial order error (%)	12	37	58
Combined phoneme error (%)*	7	32	55

* Combined errors are not analyzed as a separate error type. As they count as both identity errors and serial order errors, they have also been included in the error rates for these two error types. As a result the sum of percentages exceeds 100 in all columns.

Analysis of phoneme identity performance

Table 5 summarizes the results of the final model for the analysis of phoneme identity performance. The interaction Literacy x Syllable Length was not significant and, hence, removed in the final model. The final GLMM revealed no main effect of Literacy ($p = .12$) after controlling for serial order performance. The effect of the latter predictor indicates a significant correlation between the two phoneme-related factors ($p < .0001$; see also Table 2). This outcome indicates that there was no influence of literacy score on the retention of phonemes' identity. However, there was a main effect of Syllable Length. Across participants (i.e., no interaction with Literacy) phoneme identity retention was higher on three-syllable nonwords compared to four-syllable nonwords ($p = .0004$) and five-syllable nonwords ($p < .0001$). Similarly, phoneme identity performance was better on four-syllable nonwords compared to five-syllable nonwords ($\beta = -1.31$ for Syllable Length 5, $p = .005$) (this β estimate was obtained by fitting the same model as shown in Table 5 but using syllable length 4 as the reference level; data not shown). In line with this result, the probability of making a response without phoneme identity errors for three-, four- and five-syllable nonwords equaled .64, .31, and .15, respectively.

Note that children's phoneme identity performance decreased as their serial order performance decreased ($p < .0001$). Recall that serial order was not included in the model from a theoretical perspective; instead, this variable was only included as a covariate, i.e., to make sure that the effect of Literacy on the retention of phoneme identity was not (partially) attributable to differences in the retention of serial order. The strong correlation between identity and serial order performance reflects the distribution of the different response types (see Table 4, averages over syllable types): correct responses (32.33%), pure identity errors[4] (32%), pure serial order errors (4.33%), and combined errors (31.33%). Hence, the probability of correctly predicting the binary value for phoneme identity, given the value on the serial order variable was about .69, i.e., (32.33/36.66 + 31.33/63.33)/2 = .69 (see Table 2). Note that this high correlation underscores the importance of removing all variance from the phoneme identity data that is due to the correlation with serial order performance, in order to estimate the 'uncontaminated' effect of Literacy on phoneme identity performance. Obviously, this high correlation made it more difficult to detect an effect of Literacy on phoneme identity performance, which may be the reason why we failed to reject the null hypothesis (see General Discussion).

Table 5. Estimated fixed effects of the generalized linear mixed model predicting phoneme identity performance from literacy and syllable length when statistically controlling for serial order performance

Parameter	β	SE(β)	Z	P(>|z|)
(Intercept)	.81	.33	2.5	.01
Literacy	.08	.05	1.54	.12
Syllable Length 4	−1.58	.45	−3.55	.0004
Syllable Length 5	−2.90	.47	−6.20	<.0001
Serial Order Performance	.09	.13	7.03	<.0001

Note. Reference levels for the generalized linear mixed model are Syllable Length 3 and mean literacy score.

Analysis of serial order performance

Table 6 summarizes the results of the final model for the analysis of serial order performance. The interaction Literacy x Syllable Length was not significant and, hence, removed in the final model. The final GLMM revealed a main effect of Literacy (p = .02), indicating that increased literacy scores coincide with better serial order performance. In addition, we observed a main effect of Syllable Length, indicating that serial order performance decreased with increasing syllable length.

---

4. Pure identity errors and pure serial order errors are the errors that remain after removing the combined errors from all identity and serial order errors, respectively, as combined errors contribute to the errors on both phoneme-related dependent variables.

Across participants (i.e., no interaction with Literacy) this performance was better for three-syllable nonwords compared to four-syllable nonwords ($p < .0001$) and five-syllable nonwords ($p < .0001$). Similarly, serial order performance was better in four-syllable nonwords compared to five-syllable nonwords ($\beta = -.97$ for Syllable Length 5, $p = .02$) (this β estimate is obtained by fitting the same model as shown in Table 6 but using syllable length 4 as the reference level; data not shown). In line with this result, the probability of making a response without serial order errors for three-, four- and five-syllable nonwords equaled .88, .63, and .42, respectively.

Note that children's serial order performance increased as their identity performance increased ($p < .0001$). Recall that phoneme identity was included in the analysis as a covariate, i.e., to make sure that the effect of Literacy on serial order performance is not (partially) attributable to differences in phoneme identity performance. As in the GLMM for identity performance, the strong effect of the predictor Phoneme Identity was not surprising. Considering the distribution of responses over the four response types (see Table 4) the probability of predicting performance on the serial order variable from the value on the phoneme identity variable was about .69, i.e., (32.33/64.33+31.33/35.66)/2 = .69 (see Table 2). Note that this high correlation underscores the importance of removing all variance that is due to the correlation with phoneme identity, in order to estimate the 'uncontaminated' effect of Literacy on serial order performance. We emphasize that the latter effect is significant despite this high correlation. Hence, even though this correlation made it more difficult to detect an effect of Literacy on serial order performance, its significant effect is all the more important (see General Discussion).

Table 6. Estimated fixed effects of the generalized linear mixed model predicting serial order performance from literacy and syllable length when statistically controlling for phoneme identity performance

Parameter	β	SE(β)	Z	P(>|z|)
(Intercept)	3.68	.35	10.59	<.0001
Literacy	.21	.09	2.40	.02
Syllable Length 4	−2.22	.45	−4.96	<.0001
Syllable Length 5	−3.19	.45	−7.08	<.0001
Phoneme Identity Performance	1.05	.13	8.24	<.0001

Note. Reference levels for the mixed model are Syllable Length 3 and mean literacy score.

General discussion

The present study was designed to further our understanding of the nature of phoneme-related skills in children varying along a continuum of literacy skills using a NRT task. It is well-known that in the majority of NRT studies more skilled readers perform better when having to repeat nonwords compared to less skilled readers. According to received wisdom this is likely the result of higher-quality representations of the phonemes in their language. The latter makes it easier to develop better mappings between the newly learnt graphemes and the already existing phonemes, which are crucial for good reading performance. In our experiment, we were interested in finding out whether the quality of phoneme representations (which is higher for good readers) is the crucial factor to ensure adequate encoding, storage, and/or retrieval of phonological representations, and, if so, whether it is the only factor at the phoneme level (e.g., Elbro & Jensen, 2005; Boada & Pennington, 2006). More particularly, we wondered whether the skill to retain phonemes' serial order also underlies good readers' superior performance in phonological tasks that require good skills at the phoneme level. As current STM models dissociate between the retention of item and serial order information in STM (e.g., Martinez Perez et al., 2012a, 2012b; Binamé & Poncelet, 2016; Hachmann et al., 2014) but never addressed this issue at the level of phonemes, we disentangled the ability to retain phonemes' identity (item information) and the ability to retain serial order (serial order information) when performing a NRT-task. By studying these two retention skills with the same set of nonwords, this also made it possible to contrast them within items unlike the study of these skills with different items (even different item types) in previous studies. As a result, in contrast to traditional NRT studies we not only investigated the effect of literacy on the retention probability of the entire nonword but also on the retention probability of all phonemes in this nonword (with or without serial order errors) and on the retention probability of all (relative) serial positions of the correctly recalled phonemes (with or without identity errors).

Retention of phonemes' identity vs. retention of serial order

We performed two types of analyses: (a) global correlations between the three dependent variables of interest, performed on participant totals, and (b) generalized linear mixed effects model analyses at the level of the individual responses. In the latter analyses, phoneme identity performance and serial order performance at the phoneme level were predicted from the theoretically relevant variables and from their correlation. Both types of analyses converged on the same conclusions.

First, we checked whether we obtained the pattern that is typically found in NRT studies, i.e., NRT performance being predicted by the participant's literacy level. We indeed found a positive correlation between literacy and the number of correct nonword repetitions: as literacy scores increased the probability of making a correct NRT responses increased ($r = .22$, $p < .05$). We found the same effect when analyzing the data at the level of individual responses by means of a generalized linear mixed model (allowing the inclusion of both participants and items as random effects): the probability of a correct nonword retention increased as the child's literacy level increased. This outcome is in line with studies indicating that less skilled and dyslexic readers show lower overall nonword level performance on a NRT-task compared to typical readers (Ramus et al., 2003; Schraeyen et al., under review). Given this validation of our experimental items, we could address the question that motivated this study, i.e., whether the probability of making no phoneme identity errors (phoneme identity performance) and the probability of making no serial order errors for the correctly retained phonemes (serial order performance) is predicted by children's literacy level.

To this end, our crucial analyses pertained to the relationship between participants' literacy scores and their recall of phonemes' identities and phonemes' serial order. The correlation between participants' overall scores on these measures indicated a significant correlation between literacy scores and serial order performance ($r = .23$, $p < .05$), better performance being associated with higher literacy scores. In contrast, there was no significant correlation between literacy scores and phoneme identity performance ($r = .14$, $p > .05$). Apparently, the ability to retain phonemes' identity is not dependent on participants' literacy level. However, there was a strong positive correlation between participants' phoneme identity performance and their serial order performance, which makes it difficult to rely on these correlations. This high correlation was the result of the way in which responses were distributed across the four response categories that are defined by the two phoneme-related binary variables. In order to find out whether literacy really determines serial order performance and/or phoneme identity performance, we decided to perform GLMM analyses in which one type of phoneme performance was included as one of the fixed factors for predicting the scores on the other phoneme variable. In these analyses, we predicted one type of phoneme performance (phoneme identity, serial order) from participants' literacy scores and the number of syllables in the nonword while statistically controlling for the other type of phoneme variable (see also Staels & Van den Broeck, 2014).

When analyzing performance at the phoneme level, we found that phoneme identity performance, i.e., the processing of item information (phonemes), was not influenced by readers' literacy scores, as there was no main effect of Literacy on Phoneme Identity performance ($p > .05$) after statistically controlling for Serial

Order performance, i.e., when including Serial Order performance as a covariate. In contrast, the analysis of Serial Order performance did show a main effect of Literacy (p = .02), after statistically controlling for Phoneme Identity performance. This made us conclude that the retention of phonemes' serial order in the NRT is driven by children's literacy scores, higher serial order performance being associated with better literacy scores, or to put it in other words, the probability of making a response without serial order errors increased with increasing literacy scores.

Overall, these findings are in line with other recent studies that focused on serial order STM (for a review, see Majerus & Cowan, 2016) and confirm the results of a similar study by Schraeyen et al. (under review), in which adults with and without dyslexia were compared on their NRT performance. As in many previous studies, controls had a higher probability of making a correct nonword repetition. Performance at the two dimensions of phoneme performance was also studied: the retention of phoneme identity and of phonemes' serial order. There was no difference between the two groups in the retention of phonemes' identity. In contrast, there was an underperformance of adults with dyslexia on the retention of serial order. Even though the two studies investigated different age groups and differed in whether they investigated dyslexics vs. controls or individuals ranging in reading skills they converge on the same conclusion. The signature of poor reading skills in the NRT appears to be situated at the level of serial order retention, not at the level of phoneme identity retention. The fact that the current study corroborated Schraeyen et al.'s findings with dyslexic participants suggests that serial order problems are not exclusively associated with dyslexic readers but generalize to poor readers in general.

It may come as a surprise that children's literacy level did not affect how well they can retain phonemes' identities, given the strong emphasis in the literature on less skilled readers' low-quality representations of phonemes. This is plausibly due to the strong correlation between phoneme identity performance and serial order performance (.69). As mentioned in our mixed model analyses we included serial order performance as a predictor of phoneme identity performance (and vice versa) to remove the variance resulting from the strong correlation between these two factors – and, thus, to enable the model to calculate an uncontaminated estimate of the effect of Literacy on the dependent variable. Obviously, taking this correlation into account made it much more difficult to detect an effect of Literacy on the dependent variable, especially because it was so high. Hence, our failure to find an effect of Literacy on the retention of Phoneme Identity should be interpreted with caution. It is reasonable to believe that, if the correlation had been smaller, an effect of Literacy on Phoneme Identity performance had emerged. The fact that this effect was associated with a relatively small p-value (.12) in the current study makes this even quite plausible.

At the same time, our finding that Literacy had a significant effect on Serial Order performance, despite this strong correlation, highlights the important relationship between literacy skills and the ability to retain the serial order of phonemes in STM. This result strongly supports the conclusion that good retention skills for phonemes' serial order constitute a hallmark of good readers.

Item retention vs. serial order retention in STM

The most important conclusion of our study is that the strong relationship between children's literacy level and their retention capacity for serial order explains (at least in part) why more skilled readers typically obtain higher NRT (correct) scores than their less skilled peers: they are better able to keep the order of the phonemes from an unfamiliar phoneme string in memory.

As far as we know, our finding that children's literacy level is related to their retention ability for phonemes' serial order has not been demonstrated in previous research, even though some recent experiments have focused on the relationship between reading skill and serial order retention performance in STM tasks (for a review, see Majerus & Cowan, 2016). However, none of these studies focused on the phoneme level, i.e., they never involved a task that required the recall of the order of individual phonemes. As highlighted in the Introduction, different types of studies, using a variety of methods, item types, and participant groups, have addressed and confirmed the crucial role of serial order in STM performance (e.g. Gupta, 2003; Page & Norris, 2009; Cumming, Page, & Norris, 2003). In addition, other studies offer evidence for impaired serial order capacity in dyslexic populations. Martinez Perez et al. (2012b), for instance, reported an impaired STM for serial order information in children with dyslexia compared to chronological age-matched controls (CA) and reading age-matched controls (RA), although the dyslexic and RA groups did not differ on the item information STM task. Similar results were reported by Hachmann et al. (2014). These authors performed a study with adults and found that dyslexics performed worse on serial order processing in STM, both when using verbal and nonverbal material. Taken together, these studies suggest a domain-general serial order processing problem in dyslexics (but see Staels & van den Broeck, 2015).

However, the present results with children with a variety of literacy skills are different in at least one major aspect. In our study we disentangled item and serial order retention with the same items (nonwords) and within the same task (NRT), whereas others used different item sets. Martinez Perez et al. (2012b) used two different recall tasks to capture item and serial order retention. In their item task, participants had to recall one-syllable nonwords after a filled delay. In their serial order task, participants first had to listen to a series of auditorily presented high-frequency animal names. Then they received pictures of these animals, which they had to order, i.e., they did not have to recall the animal names, only their order. Hachmann et al. (2014) used a recognition task to capture item and serial order retention, based on the idea that this type of task directly taps into the storage function of STM without interference from working memory's executive functions (Smith-Spark & Fisk, 2007). After having seen a list of pictures, participants heard a word and had to decide whether the corresponding picture had been in the list or not (i.e., verbal item task). Then participants saw two consecutive lists of digits and had to decide whether the order of digits in both lists was the same or not (i.e., verbal order task). Martinez Perez et al. (2012b) suggested that it is reasonable that sequential memory also plays a crucial role in orthographic representations. Indeed, graphemes should be stored in a correct serial order to ensure adequate reading and writing.

A serial order mechanism in STM vs. phonotactic knowledge in LTM

To our knowledge our study is the first to directly distinguish item and serial order information at the phoneme level in an attempt to find out how literacy level affects both types of retention in an NRT, and, hence, to determine what causes poor readers' worse NRT performance. We addressed this question in a group of young readers varying along a continuum of literacy skills. The main question is how to relate our finding of serial order effects at the phoneme level to the widely accepted notion that poor readers perform worse on the NRT because the quality of their phonological representations in STM is lower than in good readers.

A first possibility is that the observed serial order effect at the phoneme level supports the existence of a domain-general sequence memory mechanism in STM, i.e., the type of mechanism that has been proposed by the studies discussed in the previous paragraph. A domain-general mechanism would be involved in maintaining the order of any series of units, i.e., also phoneme units in nonwords (Gupta, 2003; Majerus et al., 2008; Martinez Perez et al., 2012b). In the latter interpretation the worse NRT performance of children with poor literacy skill are attributed to a basic difference in working memory. This has been the suggestion throughout the paper so far. If true, our results would extend the earlier findings on the relationship between literacy and a serial order STM mechanism, more particularly, by linking this serial order mechanism to the smallest constituents of words, their phonemes. This would suggest that good literacy skills do not only depend on phonemic awareness but also on the ability to retain their serial order in STM. There is some evidence that dyslexics indeed have difficulties with processing temporal information, even in tasks that differ from the type of STM tasks described above (with children, see for instance Romani, Tsouknida, & Olson, 2015, who used visual sequences; with adults, see for instance Laasonen et al., 2012, who used audio-tactile sequences). Serial order STM memory impairments have also emerged in the study of other developmental problems, such as dyscalculia (Attout & Majerus, 2015), which suggests that learning difficulties in multiple cognitive areas might suffer from serial order STM problems (Jaroslawska, Gathercole, Logie, & Holmes, 2016).

However, even though it is intriguing to entertain this idea, a second possibility should also be considered. Good readers might experience fewer serial order problems with phonemes in the NRT because they have a better knowledge of the phonotactic system in their language, i.e., the set of permissible phoneme sequences and their occurrence frequency. According to this account phonotactic knowledge offers an extra encoding scheme in the NRT beyond the level of individual phonemes, i.e., familiar phoneme sequences can be used as 'chunks' in the STM representation. Obviously, the more participants can rely on such phoneme chunks the smaller will be their risk of making a serial order error.

Note that sensitivity to serial order must play an important role in the development of children's phonological knowledge. After all, the retention of serial order is required for the establishment of a word's phonological representation. In line with this, even very young children have been found to be sensitive to phonotactic segments in their native language (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Luce, & Charles-Luce, 1994). In addition, it has been shown that nonwords with a high phonotactic probability are easier to remember (Majerus, Poncelet, Van der Linden, & Weekes, 2008), and that learning new words with high-frequency phonological sequences is easier (Storkel, 2001).

In short, we are faced with two possible explanations of our finding that literacy scores negatively correlate with performance in the retention of serial order: the existence of a serial order mechanism in STM that is not specifically dedicated to phonemes or the importance of good phonotactic knowledge in long-term memory. Whatever be the correct account, our results do indicate that inferior reading ability is reflected in the poor retention of serial order in unknown phoneme sequences, i.e., nonwords in a NRT.

Cause or effect?

One may wonder whether the capacity to retain serial order is the *cause* the *consequence* of children's reading level. Given the correlational design of our study, it is impossible to determine the direction of the association. It is indeed reasonable to entertain the reverse hypothesis than the one we have suggested, i.e., that a strong sensitivity to the order of phonemes is the consequence of learning to read rather than a cause. Indeed, learning to read is likely to influence a child's sensitivity to phoneme sequences, an ability that is required for adequate nonword repetition and that we showed to be stronger in skilled readers than in less skilled ones. Given the requirement to attend to the order of phonemes during reading instruction, both explicit reading instruction and reading practice may nurture sequential processing and the ability to retain the order of phonemes. Nation and Hulme (2011), for instance, showed that reading at the age of 6 predicted growth in NRT performance between 6 and 7 years, independently of the effects of oral language skills and the autoregressive NRT-effect at the age of 6. However, NRT performance did not turn out to be a longitudinal predictor of the growth in reading (Hulme & Nation, 2011).These findings show that learning to read has a powerful effect on children's phonological processing and may cause better retention for phonemes' serial order in the NRT.

If our finding that skilled readers' stronger sensitivity to the serial order of phonemes is a consequence of reading practice rather than a prerequisite for becoming a good reader, this would have implications for the theoretical repercussions of the present study. However, also note that if our finding indexes a consequence of reading, researchers will have to reconsider the theoretical value of all NRT results. In such a scenario the superior NRT performance of skilled readers would be the result of having had more practice in the retention of serial order rather than reflecting the cause of poor reading, as for instance, poor phonemic or phoneme-related representations (e.g., phoneme sequences) or poor phoneme-related skills (e.g., an ill-functioning general-purpose serial order mechanism). Note that this alternative interpretation is also possible for earlier studies on STM order skills in less skilled readers or dyslexic readers (see for instance Martinez Perez, 2012b). Reading practice may have the side-effect that children become increasingly sensitive to the order of items (phonemes, syllables, words), a sensitivity that they may transfer to other cognitive domains (e.g., recalling the order of names in a list).

Despite the above two possible interpretations of our main result, the importance of a strong relationship between serial order performance at the phoneme level and literacy scores itself cannot be denied and calls for further research. More particularly, having established the relationship between reading capacity and a sensitivity to phonemes' serial order, the obvious next step is to set up a longitudinal study to find out whether preliterate children who are likely to become less skilled readers are those who have difficulties with the retention of serial order in an unfamiliar phoneme string. This finding would prove that inferior sensitivity to phonemes' serial order is causally involved in learning to read. For the time being, however, our finding offers an interesting challenge to delve deeper into the theoretical significance of NRT data.

Conclusion

Previous studies consistently revealed a relationship between reading performance and NRT performance. Until recently, it has remained unclear which factors are at the basis of this association. We reasoned that the study of phoneme-related errors in nonword repetition, rather than focusing on correct responses, makes it possible to reach a more precise understanding of NRT performance by using only a single set of items (nonwords). More particularly, it allowed us to follow the rationale that has been used in recent studies: making a distinction between participants' skill in the retention of items' *identity* and items' *serial order* in STM and their reading skill. In the NRT participants have to recall both the identity and the serial order of the phonemes in the nonword stimulus. We found that young children's literacy level predicted their performance in the retention of phonemes' *serial order* but not performance in the retention of phoneme identity), better literacy scores being associated with a higher probability of repeating the nonword without making serial order errors. Thus considered, our results could be interpreted along the lines of recent results in studies with adults, i.e., that better skills in temporarily storing and retaining the serial order among individual items correlate with better literacy skills. Hence, more skilled readers' superior retention of phonemes' serial order might be the result of a better developed domain-independent STM mechanism. However, as serially ordered phonemes form higher-order phonological units, governed by language-specific phonotactic constraints, knowledge of these phonotactically defined units, stored in LTM, might offer an alternative account of the data. Finally, it remains to be seen whether this better developed serial order skill is the consequence of learning to read or its cause.

References

Aarnoutse, C., van Leeuwe, J., & Verhoeven, L. (2005). Early Literacy from a Longitudinal Perspective. *Educational Research and Evaluation*, 11, 253–275. doi:10.1080/08993400500101054

Attout, L., & Majerus, S. (2015). Working memory deficits in developmental dyscalculia: the importance of serial order. *Child Neuropsychology*, 21, 432–450. doi:10.1080/09297049.2014.922170

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511801686

Baird, G., Slonims, V., Simonoff, E., & Dworzynski, K. (2011). Impairment in non-word repetition: a marker for language impairment or reading impairment? *Developmental Medicine and Child Neurology*, 53, 711–716. doi:10.1111/j.1469-8749.2011.03936

Bates, D. M., Maechler, M., & Dai, B. (2008). lme4: Linear mixed effects models using S4 classes (R Package Version 0.999375-25). [Software]. Retrieved from http://cran.r-project.org/web/packages/lme4/index.html.

Beneventi, H., Tonnessen, F. E., & Ersland, L. (2009). Dyslexic children show short-term memory deficits in phonological storage and serial rehearsal: an fMRI study. *International Journal of Neuroscience*, 119, 2017–2043. doi:10.1080/00207450903139671
Biname, F., & Poncelet, M. (2016). Order short-term memory capacity predicts nonword reading and spelling in first and second grade. *Reading and Writing*, 29, 1–20. doi:10.1007/s11145-015-9577-9
Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology*, 95, 153–193. doi:10.1016/j.jecp.2006.04.003
Boets, B. (2006). *Early literacy development in children at risk for dyslexia. A longitudinal study of the general magnocellular theory* (Unpublished doctoral dissertation). University of Leuven, Belgium.
de Bree, E. H., Rispens, J., & Gerrits, E. (2007). Non-word repetition in Dutch children with (a risk of) dyslexia and SLI. *Clinical Linguistics and Phonetics*, 21, 935–944. doi:10.1080/02699200701576892
Brus, B.Th., & Voeten, M. J. M. (1997). *Een Minuut Test*. Harcourt: Amsterdam.
Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high-risk of reading difficulties. *Journal of Child Psychology and Psychiatry*, 45, 631–640. doi:10.1111/j.1469-7610.2004.00252
Clarke, P., Hulme, C., & Snowling, M. (2005). Individual differences in RAN and reading: a response timing analysis. *Journal of Research in Reading*, 28, 73–86. doi:10.1111/j.1467-9817.2005.00255
Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory*, 11, 43–63. doi:10.1080/741938175
De Jong, P. F. (2011). What Discrete and Serial Rapid Automatized Naming Can Reveal About Reading. *Scientific Studies of Reading*, 15, 314–337. doi:10.1080/10888438.2010.485624
Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36, 250–287. doi:10.1598/RRQ.36.3.2
Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing*, 8, 453–485. doi:10.1007/BF00577023
Elbro, C. (1998). When reading is "readn" or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scandinavian Journal of Psychology*, 39, 149–153. doi:10.1111/1467-9450.393070
Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten. The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, 33, 36–60. doi:10.1598/RRQ.33.1.3
Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, 46, 375–384. doi:10.1111/j.1467-9450.2005.00468
Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology*, 56A, 1213–1236. doi:10.1080/02724980343000071
Gupta, P., Lipinski, J., Abbs, B., & Lin, P-H. (2005). Serial Position Effects in Nonword Repetition. *Journal of Memory and Language*, 53, 141–162. doi:10.1016/j.jml.2004.12.002
Hachmann, W., Bogaerts, L., Szmalec A., Woumans, E., Duyck, W., & Job, R. (2014). Short-term memory for order but not for item information is impaired in developmental dyslexia. *Annals of dyslexia*, 64, 121–136. doi:10.1007/s11881-013-0089-5

Hulme, C., & Snowling, M. (1992). Phonological deficits in dyslexia: a reappraisal of the verbal deficit hypothesis. In N. Singh & I. Beale (Eds.), *590 Current perspectives in learning disabilities* (pp. 270–331). New York: Springer-Verlag. doi:10.1007/978-1-4613-9133-3_9

Jaroslawska, A. J., Gathercole, S. E., Logie, M., & Holmes, J. (2016). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, 44, 580–589. doi:10.3758/s13421-015-0579-2

Jusczyk, P., Friederici, A., Wessels, J., Svenkerud, V., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402–420. doi:10.1006/jmla.1993.1022

Jusczyk, P., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630–645. doi:10.1006/jmla.1994.1030

Laasonen, M., Virsu, V., Oinonen, S., Sandbacka, M., Salakari, A., & Service, E. (2012). Phonological and sensory short-term memory are correlates and both affected in developmental dyslexia. *Reading and Writing*, 25, 2247–2273. doi:10.1007/s11145-011-9356-1

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53, 1–15. doi:10.1007/s11881-003-0001-9

Majerus, S., & Cowan, N. (2016). The nature of verbal short-term impairment in dyslexia: The importance of serial order. *Frontiers in Psychology*, 7, 1–8. doi:10.3389/fpsyg.2016.01522

Majerus, S., Poncelet, M., Van der Linden, M., & Weekes, B. (2008). Lexical learning in bilingual adults: the relative importance of short-term memory for serial order and phonological knowledge. *Cognition*, 107, 395–419. doi:10.1016/j.cognition.2007.10.003

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A. L., Martinez, T., Schmidt, C., *et al.* (2008). Neural networks for short-term memory for order differentiate high and low proficiency bilinguals. *Neuroimage*, 42, 1698–1713. doi:10.1016/j.neuroimage.2008.06.003

Majerus S., D'Argembeau A., Martinez T., Belayachi S., Van der Linden M., Collette F., *et al.* (2010). The commonality of neural networks for verbal and visual short-term memory. *Journal of Cognitive Neuroscience*, 22, 2570–2593. doi:10.1162/jocn.2009.21378

Martinez Perez, T., Majerus, S., & Poncelet, M. (2012a). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, 111, 708–723. doi:10.1016/j.jecp.2011.11.007

Martinez Perez, T., Majerus, S., Mahot, A., & Poncelet, M. (2012b). Evidence for a specific impairment of serial order short-term memory in dyslexic children. *Dyslexia*, 18, 94–109. doi:10.1002/dys.1438

McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *Journal of General Psychology*, 114, 75–88. doi:10.1080/00221309.1987.9711057

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension. *Journal of Research in Reading*, 1, 114–135. doi:10.1111/j.1467-9817.2010.01477.x

Molenberghs, G., & Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. New York: Springer.

Munson, B., Swenson, C. L., & Manthei, S. C. (2005). Lexical and phonological organization in children: evidence from repetition tasks. *Journal of Speech, Language, and Hearing Research*, 48, 108–123. doi:10.1044/1092-4388(2005/009)

Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science*, 14, 649–659. doi:10.1111/j.1467-7687.2010

Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language*, 50, 113–133. doi:10.1016/j.jml.2003.09.005

Page, M., & Norris, D. (2009). Is there a common mechanism underlying word-form learning and the Hebb repetition effect? Experimental data and a modelling framework. In A. Thorn & M. Page (Eds.). *Interactions between short-term and long-term memory in the verbal domain* (pp. 136–156), Hove, UK: Psychology Press.

Pratt, A., & Brady, S. (1988). Relation of Phonological Awareness to Reading Disability in Children and Adults. *Journal of Educational Psychology*, 80, 319–323. doi:10.1037/0022-0663.80.3.319.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, 126, 841–865. doi:10.1093/brain/awg076

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology*, 61, 129–141. doi:10.1080/17470210701508822

Rispens, J., & Parigger, E. (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. doi:10.1348/026151009X482633

Robertson, E. K., & Joanisse, M. F. (2010). Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics*, 31, 141–165. doi:10.1017/S0142716409990208

Romani, C., Tsouknida, E., & Olson, A. (2015). Encoding order and developmental dyslexia: a family of skills predicting different orthographic components. *Quarterly Journal of Experimental Psychology*, 68, 99–128. doi:10.1080/17470218.2014.938666

Schraeyen, K., Geudens, A., Ghesquière, P., & Sandra, D. (under review). *Serial order effects in young adults with dyslexia: new perspectives from nonword repetition errors*. Manuscript under review.

Smith-Spark, J., & Fisk, J. (2007). Working memory functioning in developmental dyslexia. *Memory*, 15, 34–56. doi:10.1080/09658210601043384

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development*, 74, 358–373. doi:10.1111/1467-8624.7402003

Staels, E., & Van den Broeck, W. (2014). Order short-term memory is not impaired in dyslexia and does not affect orthographic learning. *Frontiers in Human Neuroscience*, 8, 1–16. doi:10.3389/fnhum.2014.00732

Staels, E., & Van den Broeck, W. (2015). No solid empirical evidence for the SOLID (Serial order learning impairment) hypothesis of dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 650–669. doi:10.1037/xlm0000054

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, 86, 24–53. doi:10.1037//0022-0663.86.1.24

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, 44, 1321–1337. doi:10.1044/1092-4388(2001/103)

Van den Bos, K. P., Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. C. (1998). *De klepel pseudowoordentest*. Harcourt: Amsterdam.

Van den Broeck, W., Geudens, A., & Van den Bos, K. P. (2010). The nonword-reading deficit of disabled readers: A developmental interpretation. *Developmental Psychology*, 46, 717–734. doi:10.1037/a0019038

Verbeke, G., & Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer-Verlag.

Wagner, R. K. & Torgesen, J. K., (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192–212. doi:10.1037/0033-2909.101.2.192

Appendix A. Target items for the Nonword Repetition Task (NRT)

Three syllable items	Four syllable items	Five syllable items
/pi.pó.ket/	/sir.peg.wót.nal/	/bi.ko.ta.dú.kep/
/ves.róf.sif/	/kep.dar.bít.puk/	/gam.bis.kef.lún.jor/
/lem.rós.pag/	/lu.mo.gá.pes/	/wog.tum.dis.jáf.kel/
/pir.dáp.ket/	/ni.ne.lú.mar/	/bog.nup.sar.líf.tek/
/ku.mí.gar/	/jaf.ker.túm.sil/	/ga.go.li.sú.gef/
/po.sá.lin/	/be.fo.tí.ral/	/na.mo.ni.fú.nem/
/ni.mú.naf/	/zos.gef.zíl.vas/	/vug.zas.gof.líf.sef/
/nom.lún.fam/	/hi.so.ré.fum/	/pip.taf.bip.két.duk/
/pig.dúl.mek/	/no.si.gé.fas/	/so.ra.mi.ké.tul/
/har.lón.wig/	/keg.jol.bíf.mas/	/lu.re.fa.ní.pos/
/gi.ká.lom/	/pe.da.tí.tup/	/hu.sa.li.mó.gep/
/sa.vé.fus/	/mif.nem.lúm.zan/	/mun.fom.lin.zém.bam/

Corresponding address

Kirsten Schraeyen
Department of Speech Language Therapy
and Audiology
8, Molenstraat
2018 Antwerp
Belgium
Kirsten.schraeyen@thomasmore.be