KU LEUVEN

FACULTY OF PSYCHOLOGY AND
EDUCATIONAL SCIENCES

Universiteit
Antwerpen

FACULTY OF ARTS

Parenting and Special Education

Linguistics

Item and serial order retention in phonological short-term memory representations in relation to reading performance

Kirsten Schraeyen

Doctoral thesis offered to obtain the degree of
Doctor of Educational Sciences (PhD) and the degree of
Doctor of Linguistics (PhD) within the framework of
a joint doctorate at
KU Leuven and Universiteit Antwerpen

Supervisors: Prof. Dr. Pol Ghesquière, Prof. Dr. Dominiek Sandra
Co-supervisor: Dr. Astrid Geudens

Kirsten Schraeyen, Item and serial order retention in phonological short-term memory representations in relation to reading performance. Doctoral thesis offered to obtain the joint degree of Doctor of Educational Sciences (KU Leuven) and Doctor of Linguistics (Universiteit Antwerpen), 2018.

Supervision: Prof. Dr. Pol Ghesquière, Prof. Dr. Dominiek Sandra, Dr. Astrid Geudens

This thesis focuses on item and serial order retention in verbal short-term memory (STM) of monolingual good and poor readers and typically developing bilingual children learning to read in a second language. One of the leading cognitive theories of dyslexia, i.e., the phonological deficit theory, postulates that individuals with dyslexia and poor readers have specific problems in the representation, storage, and/or retrieval of (a sequence of) speech sounds. It has been suggested that these poorly represented, stored, or retrieved phonological representations are at the core of their reading problems, as these representations are crucial for the establishment of associations between the orthographic and phonological representations of words, more particularly, by connections between the constituent phonemes and the corresponding graphemes. Support for this theory comes from evidence that individuals with dyslexia and poor readers perform particularly poorly in three phonological domains (phonological awareness, verbal STM, and lexical retrieval), which all involve the storage and retrieval of phonological information. Given that a phonological representation is a sequence of speech sounds, it involves several phonemic representations. Hence, both the identity of the individual speech sounds (item information) and their serial order (serial order information) are potentially important when processing phonological information. The strong research focus on the quality/accessibility of phonemic representations within the phonological deficit theory reflects an almost exclusive interest in the representation/retrieval of phonemes' identity. However, an important but so far unanswered research question is whether phonemes' serial order also plays a role in the adequate representation/retrieval of phonological representations. To address this question, we examined item and serial order retention in verbal STM processing by means of a nonword repetition task (NRT). We analyzed the retention of phonemes and their relative serial order in a novel (but phonotactically well-formed) phoneme string. In contrast to traditional NRT studies, we applied a unique scoring technique at the phoneme level, to study the effect of literacy on both the retention probability of phonemes' identity and phonemes' serial order. This technique was first applied in a sample of 46 normal readers and 45 adults with dyslexia (study 1), and later in a sample of 89 third-grade children representing a continuum of literacy skills (study 2). In both studies, individuals with dyslexia and poor readers demonstrated a decreased retention of serial order, but had an unaffected retention of phoneme identity. These results confirm that serial order retention is a crucial factor for describing differences in overall NRT performance (hence, for understanding the processing of phonological information in STM) between poor and better readers.

According to recent STM models, item retention depends on lexical knowledge stored in long-term memory (LTM), whereas serial order retention is regulated by a specific STM device that holds serial order information. To assess this idea, a sample of 20 typically developing Dutch monolingual children and 18 typically developing Turkish-Dutch bilingual children was studied (study 3). We examined the retention probability of phonemes' identity and phonemes' serial order, using NRT items that are in line with the phonotactic rules of the monolinguals' native and the bilinguals' non-native language. The results confirmed that the retention of phoneme identity was influenced by lexical knowledge, whereas the retention of serial order was not. These results provide more detailed information on the determinants of typically developing bilinguals' STM capacity for the encoding and retrieval of non-native phonological information. In addition, they provide evidence for distinct phoneme identity and serial order processing in STM.

We conclude that the distinction between the STM retention of phonemes' identity and serial order contributes to a more precise understanding of NRT performance on the one hand and the relation between these two task components and reading and language proficiency on the other hand. Hence, our research adds to a growing body of research regarding the involvement of serial order memory for reading success, although the potential role of phonotactic knowledge cannot be excluded.

Kirsten Schraeyen, Item and seriële orde retentie in fonologische kortetermijngeheugen representaties in relatie tot leesvaardigheid. Proefschrift aangeboden tot het verkrijgen van de gezamenlijke graad van Doctor in de Pedagogische Wetenschappen (KU Leuven) en Doctor in de Taalkunde (Universiteit Antwerpen), 2018.
Promotoren: Prof. Dr. Pol Ghesquière, Prof. Dr. Dominiek Sandra, Dr. Astrid Geudens

Dit proefschrift onderzoekt het bijhouden van item en volgorde informatie in het verbaal kortetermijngeheugen (KTG) bij eentalige goede en zwakke lezers en normaal ontwikkelende tweetalige kinderen die leren lezen in hun tweede taal. De fonologische-deficittheorie, één van de meest toonaangevende cognitieve theorieën over dyslexie, stelt dat zwakke lezers en personen met dyslexie specifieke problemen vertonen in de representatie, het opslaan en/of het ophalen van (een opeenvolging) van spraakklanken. Er wordt gesuggereerd dat deze gebrekkig gerepresenteerde, opgeslagen of opgehaalde fonologische representaties aan de basis liggen van hun leesproblemen aangezien ze cruciaal zijn voor het vormen van associaties tussen orthografische en fonologische representaties van woorden, door connecties tussen fonemen en de overeenstemmende grafemen. Ondersteuning voor deze theorie wordt geleverd door evidentie dat personen met dyslexie en zwakke lezers een specifieke zwakte vertonen binnen drie fonologische domeinen (fonologisch bewustzijn, verbaal KTG en ophalen van lexicale informatie), die allemaal de opslag en ophaling van fonologische informatie vereisen. Omdat een fonologische representatie een opeenvolging van spraakklanken betreft, gaat het om een aaneenschakeling van fonemische representaties. We kunnen dus stellen dat zowel de identiteit van individuele spraakklanken (iteminformatie) als hun volgorde (seriële-orde-informatie) potentieel belangrijk zijn bij de verwerking van fonologische informatie. De sterke onderzoeksfocus op de kwaliteit van/toegang tot fonemische representaties binnen de fonologische-deficittheorie weerspiegelt een bijna uitsluitende interesse in de representatie/ophaling van de foneemidentiteit. Echter, een belangrijke maar tot dusver onbeantwoorde onderzoeksvraag is of de foneemvolgorde ook bepalend is voor de representatie/ ophaling van fonologische representaties in het verbaal KTG. Om deze vraag te beantwoorden onderzochten we de verwerking van item- en volgorde-informatie in het verbaal KTG aan de hand van een nonwoord repetitietaak (NRT). We analyseerden de retentie van fonemen en hun relatieve volgorde in onbestaande (maar fonotactisch welgevormde) foneemsequenties. In tegenstelling tot traditionele NRT-studies, pasten we een nieuwe scoringtechniek toe, nl. op het foneemniveau, waardoor we het effect konden nagaan van leesvaardigheid op het vasthouden van de foneemidentiteit en -volgorde. Deze techniek werd eerst gebruikt bij 46 normaallezende en 45 volwassenen met dyslexie (studie 1) en daarna bij 89 kinderen uit het derde leerjaar met een variatie in leesvaardigheid (studie 2). Zowel volwassenen met dyslexie als jonge zwakke lezers vertoonden meer problemen met het vasthouden van de foneemvolgorde terwijl dat niet het geval was met het vasthouden van de foneemidentiteit. Deze resultaten bevestigen dat het vasthouden van foneemvolgorde een cruciale factor is voor de beschrijving van verschillen in NRT-prestatie (en dus de verwerking van fonologische informatie) tussen zwakke en betere lezers.

Volgens recente KTG-modellen is het vasthouden van iteminformatie afhankelijk van lexicale kennis in het langetermijngeheugen (LTG), terwijl het vasthouden van seriële volgorde gereguleerd wordt door een specifiek KTG-mechanisme dat volgorde-informatie bijhoudt. Om die hypothese te onderzoeken werden 20 normaal ontwikkelende Nederlands sprekende eentalige kinderen en 18 normaal ontwikkelende Turks-Nederlands sprekende tweetalige kinderen onderzocht (studie 3). We bestudeerden de retentie van de foneemidentiteit en -volgorde aan de hand van NRT-items die opgebouwd waren volgens de fonotaxis van het Nederlands en het Turks. De resultaten bevestigden dat het vasthouden van de foneemidentiteit beïnvloed werd door LTG-kennis, terwijl dit niet gold voor het vasthouden van volgorde-informatie. Deze resultaten geven meer inzicht in hoe normaal ontwikkelende tweetalige kinderen niet-taaleigen fonologische informatie verwerken. Bovendien leveren ze bewijs dat foneemidentiteit en -volgorde op een verschillende manier verwerkt worden.

We besluiten dat het onderscheid tussen twee types KTG-retentie bijdraagt tot een beter begrip van NRT-prestaties enerzijds en tot een inzicht in het verband tussen deze twee taakcomponenten en lees- en taalvaardigheid anderzijds. Deze resultaten ondersteunen het toenemend aantal bevindingen over de rol van een geheugen voor seriële volgorde in relatie tot leessucces, hoewel de potentiële rol van fonotactische kennis niet genegeerd kan worden.

DANKWOORD

It always seems impossible until it's done
Nelson Mandela

Nu de laatste hand gelegd is aan mijn proefschrift, wil ik graag even de tijd nemen om de vele mensen te danken die de afgelopen jaren bewust of onbewust hebben bijgedragen tot de realisatie ervan.

Allereerst wil ik mijn promotoren prof. Pol Ghesquière, prof. Dominiek Sandra en dr. Astrid Geudens bedanken voor het aanvaarden van het promotorschap en de unieke kans die ze me boden om dit onderzoek uit te voeren in een interuniversitair kader. Pol, ik wil jou bedanken voor je vakkundige begeleiding. Jouw deur stond altijd open, ook al was ik misschien wel één van de meest atypische doctoraatsstudenten die je ooit onder jouw vleugels nam. Je hebt me steeds de ruimte gegeven om mijn traject uit te stippelen op mijn tempo. Jouw nuchtere kijk, waardevolle raad en aanmoedigingen onderweg, maar ook je constructieve feedback op mijn onderzoek, gaven me steeds het nodige zelfvertrouwen om door te gaan. Dominiek, ik wil jou bedanken voor de bevlogenheid waarmee je je schouders zette onder onze samenwerking. Ik heb veel geleerd van onze boeiende discussies en je scherpzinnig inzicht in de materie. Jouw deskundige commentaar bij verschillende onderzoeksobstakels motiveerden me telkens om nog dieper in de data te duiken. Met je enthousiasme gaf je de onderzoeksmicrobe door. Pol en Dominiek, ik bewaar mooie herinneringen aan onze gezamenlijke overlegmomenten. Naast jullie inhoudelijk advies gaven jullie me ook telkens de moed om vol te houden en het onderste uit de kan te blijven halen. Astrid, woorden schieten te kort om jou te bedanken. Ik herinner me nog levendig onze eerste kennismaking op de trein richting onze respectievelijke thuishavens. We bleken veel gemeenschappelijk te hebben. Je hebt me met open armen verwelkomd binnen de expertiseteams leerstoornissen en daarna meertaligheid, onderdeel van het MDCL (later CODE) dat je jarenlang met hart en ziel geleid hebt. Jij hebt me bovendien het vertrouwen gegeven om dit doctoraatsproject te starten. De kiem van het onderzoek stelden we samen voor op SSSR in Berlijn (editie 2010). Ik wil je bedanken voor het delen van je ongelooflijke vakkennis, je hulp en deskundig advies onderweg, je woorden van steun, je oppeppende berichtjes dag én nacht en zoveel meer… maar bovenal voor jouw geloof in mij!

Een bijzonder woord van dank aan dr. Wim Van der Elst. Wim, verschillende cursussen R-statistiek waren mijn deel doorheen de jaren, maar pas na onze kennismaking gingen de *Mixed Effects Models* leven voor mij. Waarschijnlijk had je bij ons eerste overleg nooit gedacht dat je nog moeilijk van me af zou geraken. Jouw statistische kennis en geduld om mij hierin wegwijs te maken, je kritische blik op de uit te voeren statistiek zijn van onschatbare waarde gebleken voor dit doctoraat.

Furthermore, I would like to thank the members of my doctoral guidance committee, prof. Bert De Smedt, prof. Gert Storms, prof. Helen Grech and prof. Walter Daelemans for generously offering their time and their useful suggestions. I am also grateful that prof. Bert De Smedt, prof. Helen Grech, prof. Steven Gillis and prof. Lieven Verschaffel (chair exam committee) accepted to be members of the jury. I am not only indebted to them for their critical reading of the manuscript and their valuable recommendations, but also gratefully acknowledge their interest in my work and for being present at my defense.

Daarnaast een woord van dank aan het management van Thomas More om me de vrijheid te geven een deel van mijn werktijd te kunnen besteden aan dit onderzoek. Ook mijn collega's binnen de opleiding Logopedie en Audiologie, Thomas More, wil ik bedanken. Lieve collega's, dank voor jullie interesse, de fijne gesprekken en jullie bemoedigende woorden op tijd en stond. Een speciaal dankwoordje voor An. Wij hebben doorheen de jaren een bijzondere band opgebouwd. Ik ken je als iemand waarop je altijd kan vertrouwen, een fantastische collega die staat voor wat ze zegt en doet wat ze belooft. Je bent een belangrijk rolmodel voor studenten. Ik wil je bedanken omdat je er altijd was voor mij, ook als het even wat minder ging. Verder wil ik twee oud-collega's vermelden die mijn professionele loopbaan mee gevormd hebben. Kris, ik startte mijn onderzoek toen jij nog aan het hoofd van de opleiding Logopedie en Audiologie stond. Jij hebt me vanaf mijn aanstelling in 2004 binnen Thomas More (voorheen Lessius) het vertrouwen gegeven om mijn grenzen te verleggen. Dit doctoraatstraject was daar slechts één van. Ook na jouw afscheid als opleidingshoofd bleef je geïnteresseerd in mijn (doctoraats)werk. Lieve Hilde, jij hebt me vanaf dag één letterlijk in je armen gesloten. Je hebt me geïntroduceerd binnen het internationaal Speech Language Pathology netwerk dat jij uitbouwde. Samen trokken we de wereld rond en ik ben nog elke dag dankbaar voor onze mooie momenten en gezamenlijke realisaties. Vandaag zet ik met de nodige fierheid het internationaal SLP-netwerk verder. I am honoured to coordinate this worldwide SLP-IN Network and thankful for the great opportunity to work together with so many colleagues from 19 partner universities within this network. Dear international colleagues and friends, thanks for your support and interest in my PhD research throughout the years.

Zij die me kennen, weten dat ik de laatste jaren veel tijd heb doorgebracht in Suriname in het kader van een aantal onderwijs- en onderzoeksprojecten, aansluitend bij mijn onderzoeksinteresse rond (meer)talige ontwikkeling in relatie tot lees- en bij uitbreiding schoolsucces. Ik bedank dan ook graag prof. Wim Van Petegem, Lieven Faes, prof. Henry Ori, Ranoe Mangal-Jhari, dr. Renata de Bies, Naomi Lieuw-Tjoen-Sioe, Fiona Pawirodikromo en Gracia Nelson voor de boeiende samenwerking.

Ook een oprecht woord van dank aan de volwassenen en kinderen die hun medewerking verleenden aan dit onderzoek. Zonder hun bereidheid zou dit niet mogelijk geweest zijn. Verder bedank ik graag de studenten Isabeau Beerts, Monika De Cremer en Mukaddes Cevik om me te ondersteunen in een deel van de dataverzameling.

Tot slot wil ik me richten tot die mensen die misschien niet inhoudelijk betrokken waren bij mijn doctoraat, maar wel van onschatbare waarde zijn geweest voor het slagen in dit traject.

Lieve mams en paps, Jochen en Ruth, ontzettend bedankt voor jullie rotsvast vertrouwen in mij. Jullie hebben mij altijd gesteund bij elke beslissing die ik genomen heb. Bedankt voor het luisterend oor, de aanmoedigingen om mijn dromen te durven nastreven en het opvangen van zoveel praktische zaken. Lieve Mia en Jos, Pieter en Nena, mijn unieke schoonfamilie, jullie waren meer dan eens mijn vangnet en hebben op de achtergrond steeds gezorgd dat de trein bleef sporen. Aan familie en vrienden, bedankt voor jullie steun en gezellige afleiding. Bedankt ook voor jullie begrip dat ik er geregeld niet bij kon zijn omdat mijn onderzoekswerk prioriteit kreeg.

Lieve Conrad, wat is mama trots op jou! Van kleins af heb je me vaak moeten missen, de ene keer omdat ik in het buitenland zat, de andere keer omdat ik werkte aan mijn doctoraatsonderzoek. Gelukkig werd je steeds fantastisch opgevangen door papa, moeke en vake, oma en opa en zoveel anderen. Ik kijk al uit naar de *zee van tijd* in de komende jaren om samen de wereld verder te ontdekken, maar ook gewoon samen te genieten van een partijtje tennis of jouw kunsten in het skatepark!

De laatste in de rij maar met stip de belangrijkste persoon die ik wil bedanken is Toon, mijn rots in de branding en steun en toeverlaat sinds meer dan 20 jaar. Schattie, er zijn geen woorden om te beschrijven hoe dankbaar ik ben dat jij al die tijd aan mijn zijde gestaan hebt. Jij hebt me steeds de ruimte gegeven om mijn dromen te realiseren, ondanks jouw drukke carrière. Jouw woorden van steun, bemoedigende berichtjes, verplichte ontspanningsmomenten, maar ook jouw droge humor waren goud waard. Eerlijk is eerlijk, nog meer dan ik, heb jij uitgekeken naar dit moment, terug wat meer *quality time* met je vrouwtje … Ik hoop dat we al onze toekomstdromen verder vorm kunnen geven. Zoals jouw motto luidt: '*Samen kunnen wij alles aan*'. Dit werk is daar nog maar eens het bewijs van.

Kirsten Schraeyen
Mei 2018

TABLE OF CONTENTS

CHAPTER 1 ..... **1**

GENERAL INTRODUCTION **1**

THE IMPORTANCE OF PHONOLOGICAL SKILLS IN MODELS OF READING **2**

EMPIRICAL FINDINGS 2
THE PHONOLOGICAL DEFICIT HYPOTHESIS 5
Defective phonological representations? *5*
Defective access to phonological representations? *6*

PHONOLOGICAL REPRESENTATIONS **7**

THE ROLE OF PHONEMES IN THE DEVELOPMENT OF PHONOLOGICAL REPRESENTATIONS IN CHILDREN'S FIRST LANGUAGE (L1) 8
THE ROLE OF PHONEMES IN THE DEVELOPMENT OF PHONOLOGICAL REPRESENTATIONS IN CHILDREN'S SECOND LANGUAGE (L2) 10

THE GOAL OF THIS DISSERTATION: TOWARDS A NOVEL USE OF THE NRT **11**

PROPERTIES AND PAST USE OF THE NRT 12
TOWARDS AN IN-DEPTH ANALYSIS OF ERROR PATTERNS IN THE NRT 14

THE RETENTION OF ITEMS' IDENTITY AND SERIAL ORDER IN STM **15**

SEQUENCE (HEBBIAN) LEARNING MEDIATES BETWEEN STM AND LTM 16
EVIDENCE FOR A DISTINCT SERIAL ORDER MECHANISM IN STM 17
THE RELATIONSHIP BETWEEN SERIAL ORDER STM PROCESSING AND READING PERFORMANCE 18

RATIONALE OF THE CONDUCTED STUDIES **20**

REFERENCE LIST **25**

CHAPTER 2 ..... **37**

SHORT-TERM MEMORY PROBLEMS FOR PHONEMES' SERIAL ORDER IN ADULTS WITH DYSLEXIA: EVIDENCE FROM A DIFFERENT ANALYSIS OF THE NONWORD REPETITION TASK **37**

ABSTRACT **37**

INTRODUCTION **38**

METHOD **42**

PARTICIPANTS 42
MATERIALS AND DESIGN 42
PROCEDURE 42
SCORING 43
STATISTICAL ANALYSES 47

RESULTS **49**

CORRELATIONAL ANALYSES 49
NONWORD LEVEL PERFORMANCE 51
PHONEME LEVEL PERFORMANCE 52
PAIRWISE COMPARISONS BETWEEN RESPONSE TYPES 56

DISCUSSION **60**

SERIAL ORDER PROBLEMS IN A PHONOLOGICAL REPRESENTATION IN STM 63
SERIAL ORDER PROBLEMS IN AN ORTHOGRAPHIC REPRESENTATION IN STM 65
SERIAL ORDER PROBLEMS IN A DOMAIN-GENERAL SEQUENCING MECHANISM IN STM 66
SERIAL ORDER PROBLEMS DURING RETRIEVAL FROM STM 67

CONCLUSION **69**

REFERENCE LIST **70**

CHAPTER 3 ..... **75**

POOR PERFORMANCE ON THE RETENTION OF PHONEMES' SERIAL ORDER IN SHORT-TERM MEMORY REFLECTS YOUNG CHILDREN'S POOR READING SKILLS **75**

ABSTRACT **75**

INTRODUCTION **76**

METHOD **82**

Participants 82
Materials and design 83
Procedure 83
Nonword level score *84*
Phoneme level scores *84*
Statistical analyses 87

RESULTS **89**

Correlational analyses 89
Analysis of nonword level performance 90
Analysis of phoneme level performance 91
Analysis of phoneme identity performance 92
Analysis of serial order performance 94

DISCUSSION **96**

Retention of phonemes' identity vs. retention of serial order 96
Item retention vs. serial order retention in STM 99
A serial order mechanism in STM vs. phonotactic knowledge in LTM 100
Cause or effect? 102

CONCLUSION **103**

REFERENCE LIST **104**

CHAPTER 4 ..... **111**

BEYOND GLOBAL DIFFERENCES BETWEEN MONOLINGUAL AND BILINGUAL CHILDREN ON THE NONWORD REPETITION TASK: RETENTION SKILLS FOR PHONEMES' IDENTITY AND SERIAL ORDER **111**

ABSTRACT **111**

INTRODUCTION **112**

METHOD **117**

Participants 117
Materials and design 118
Scoring 119
Statistical analyses 124

RESULTS **125**

Vocabulary performance and NRT accuracy 125
Correlation analyses 125
Nonword level performance in monolinguals and bilinguals 127
Phoneme level performance in monolinguals and bilinguals 128
Distribution of error types in monolinguals and bilinguals 130

DISCUSSION **134**

CONCLUSION **140**

REFERENCE LIST **142**

CHAPTER 5 ..... **147**

GENERAL CONCLUSIONS **147**

SUMMARY OF FINDINGS **147**

THE ROLE OF ITEM AND SERIAL ORDER RETENTION IN LEARNING TO READ **148**

MEASURING ITEM AND SERIAL ORDER STM PROCESSING WITH THE NRT **150**

ITEM AND SERIAL ORDER STM PROCESSING IN RELATION TO READING **151**

STM EFFECTS IN CHILDREN'S AND ADULTS' L1 151
STM EFFECTS IN CHILDREN'S L2 153

INTEGRATIVE ACCOUNTS OF SERIAL ORDER PROCESSING PROBLEMS IN LESS SKILLED READERS **154**

STORAGE-BASED PERSPECTIVES 154
The role of phonotactic knowledge *154*
The role of orthographic knowledge *156*
The role of a domain-general sequence mechanism *157*
A RETRIEVAL-BASED PERSPECTIVE 158
CONCLUSION 159

WHAT ABOUT THE PHONOLOGICAL DEFICIT HYPOTHESIS? **159**

CLINICAL IMPLICATIONS AND PRACTICAL RELEVANCE **160**

METHODOLOGICAL LIMITATIONS **161**

FUTURE DIRECTIONS **164**

CONCLUSION **166**

REFERENCES LIST **168**

APPENDIX A **173**

CHAPTER 1

GENERAL INTRODUCTION

When considering today's information-oriented society, learning to read is one of the most important skills a child has to acquire. In contrast to learning to speak, however, learning to read is not a spontaneous process. This difference is reflected in the terminology: speech *acquisition* versus *learning* to read. Educational support is essential for this learning process. Within educational systems, reading instruction is an explicit focus of attention. Nevertheless, a substantial number of both monolingual and bilingual children encounter problems with learning to read and suffer from severe reading and spelling problems, despite normal intelligence, adequate instruction, and intense remedial effort. These problems are known as specific reading impairments or developmental dyslexia (Vellutino, Fletcher, Snowling, & Scanlon, 2004).

What makes learning to read difficult for some children? From a cognitive perspective, emergent literacy is an ongoing process of learning to understand the link between the written and spoken language system. More specifically, learning to read in an alphabetic language primarily entails learning to identify the visual percepts of the alphabetic script as specific graphemes and transform these graphemes into linguistically relevant sounds of that language, i.e., phonemes. To accomplish this, the beginning reader must achieve so-called phoneme awareness, i.e., the ability to recognize, identify, or manipulate speech sound units in the spoken language (Gillon, 2004), or, to put it in other words, the awareness of the internal phonemic structure of the words of a language (e.g. to identify the sounds /c/ /a/ /t/ in the word /cat/). In addition, beginning readers must realize that the orthography or sequence of letters represents this phonemic structure, i.e., must be able to perform grapheme-phoneme conversion. Ultimately, they should be able to acquire meaning from that sequence of written symbols. For some children, however, this learning process is hampered. Research has shown that children with reading problems have less developed phoneme awareness (e.g., Stanovich & Siegel, 1994). In addition, children with reading problems have trouble with phoneme-to-grapheme conversions (e.g., Snowling & Melby-Lervag, 2016).

In the first part of this introduction, we will focus on the importance of phonological skills in reading development. We will also highlight some hypotheses that have been put forward regarding reading problems. Then, we will discuss the development of phonological representations, as well as the use of nonword[1] repetition as a measure of the nature and quality of such representations. Next, we will describe the main findings from nonword repetition studies with monolingual children and sequential bilingual children learning to read in a second language. We will relate these findings to current views on the processing of phonological information in verbal short-term memory (STM). Finally, we will conclude the introduction by describing the overall purpose of this doctoral thesis and giving an overview of the studies that were conducted.

THE IMPORTANCE OF PHONOLOGICAL SKILLS IN MODELS OF READING

EMPIRICAL FINDINGS

For each beginning reader, the challenge par excellence is to develop an orthographic system that allows meanings to be accessed quickly and efficiently from orthography. This process entails three key features: orthography (spelling), phonology (pronunciation), and semantics (meaning) (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). More importantly, all of these features must be fully integrated into one lexical representation in a way that activation of one feature leads to the successful retrieval of the other features from memory (Perfetti & Hart, 2001). In short, for reading success, children must know the spelling of the word ‘bed’, the pronunciation /bɛd/, and the meaning (place to sleep), but, just as importantly, those three elements must be closely integrated so that when children see the word ‘bed’ the pronunciation and meaning are automatically associated with it. For children who learn to read, the link with semantics is not a new one; they have relied on connections between phonological and semantic representations during their childhood. But now, they also need to learn about form, i.e., orthography, phonology, and the links between graphemes and phonemes. Reading development starts with this crucial insight that each grapheme corresponds to a speech sound

[1] Even though the term nonword repetition is consistently used in the literature, the test items in a nonword repetition task are usually pseudowords rather than nonwords. Pseudowords consist of legal phoneme strings in the language of the test whereas nonwords may contain phoneme strings that are illegal in that language, which makes them difficult to pronounce.

or phoneme. Thus, when beginning readers try to read an unfamiliar written word in a language, the word is broken down into individual letters or groups of letters (graphemes) which have to be mapped onto the corresponding speech sound or speech sound sequences in that language, i.e., the process of phonological decoding. More specifically, corresponding phonemic representations are activated and blended and subsequently provide access to the phonological representation of that word. According to the 'self-teaching' hypothesis (Share, 1995, 1999), this phonological decoding process is the first step in the acquisition of an orthographic representation, i.e., the spelling of the word. However, this model proposes that reading a word sufficiently often, i.e., one to four exposures in typically developing readers, establishes an orthographic representation in long-term memory (LTM), which will be mapped onto its phonological representation and its meaning (Perfetti, 2007). As a consequence, this (sufficiently often encountered) written word will be read by sight, i.e., the available orthographic representation in LTM will make it possible to read the word immediately. This process is referred to as orthographic mapping (Erhi, 2014): connections between the spelling, pronunciation, and meaning of specific words are formed by applying both lexical (word level) and sublexical orthographic information (subword level, e.g., a sequence of graphemes such as /–ing/ /ver-/ in Dutch). When the orthographic representation of a word is fully specified in LTM, the impact of phonology will probably diminish but does not disappear (Harris & Perfetti, 2017). Indeed, even fluent silent word reading does not rely solely on orthographic information but involves the activation of the phonology of words. This is revealed for instance by the phonological priming effect. Using the masked priming paradigm, it has been shown that a word such as *feet* will be recognized faster when it is preceded by a homophonic prime word such as *feat* than when it is preceded by a graphemic prime such as *fest* (i.e., a prime with the same number of position-specific graphemes in common with the target word) (see for instance, Humphreys, Evett, & Taylor, 1982). It is important to note that this is not due to phonological priming within the mental lexicon, as the same phonological priming effect has been shown when using pseudohomophones as primes. A word such as *weep* is recognized faster when preceded by a pseudohomophonic prime word such as *weap* than when it is preceded by a graphemic prime such as *werp* (see Perfetti & Bell, 1991; for more recent evidence, see Rastle & Brysbaert, 2006).

Since fluent phonological processing is a necessary process for word reading, it should be no surprise that poor readers experience more problems with processing phonological representations, both implicitly and explicitly. Research has indeed shown that these readers tend to have problems with the access to phonemes in words and with the retrieval of

phonological sequences from both their STM and LTM (Wagner & Torgesen, 1987). The former skill is known as phoneme awareness (PA), i.e., the skill to consciously identify and manipulate individual phonemes in words. The latter skills refer to the retrieval *accurateness* of unfamiliar phoneme sequences from STM (i.e., pseudowords) and the retrieval *speed* of familiar phoneme sequences (i.e., words) from LTM.

First, it has been shown that individuals with reading problems demonstrate poor PA (Swan & Goswami, 1997; Lyon, Shaywitz, & Shaywitz, 2003). There are several tasks to measure PA, e.g., phoneme deletion tasks (omitting a sound in a word, e.g., *block without /b/*) and spoonerism tasks (swopping the onsets of two successive words, e.g., *bunny rabbit* becomes *runny babbit*). It has been claimed that one aspect of PA, i.e., phoneme awareness, is the most widely replicated and accepted correlate of early reading in an alphabetic orthography, such as Dutch (Foy & Mann, 2001; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001).

Second, individuals with reading problems have a poor verbal STM (Beneventi, Tonnessen, & Ersland, 2009). Verbal STM refers to the memory system where verbal (phonological) information can be stored for a short period. This verbal STM plays a crucial role in learning, as it is a temporary storage device before information can be consolidated in LTM. Poor verbal STM in reading disabled individuals has been shown in memory span tasks, in which verbal information such as words, nonwords (phonotactically *illegal* letter strings that are no real words, e.g., /unsd/) or pseudowords (phonotactically *legal* letter strings that are no real words, e.g., /adair/) must be recalled.

Third, there is evidence that individuals with reading problems are hampered in rapid lexical retrieval (Wolf & Bowers, 1999). This ability involves the mapping of visual information (for instance, the picture of a dog) onto a sound-based representation, i.e., the referent's name, by retrieving the appropriate lexical form from LTM. A task that has often been used to measure this ability is rapid automatized naming (RAN), in which participants are asked to name a list of pictures, colors, letters, or digits as fast as they can. A wealth of studies has shown that RAN predicts later literacy performance in so-called transparent orthographies with less orthographic depth (e.g., de Jong & van der Leij, 1999; de Jong & van der Leij, 2002; Lervag, Braten, & Hulme, 2009) as well as in so-called opaque orthographies with more orthographic depth. Transparent and opaque orthographies are the endpoints on a continuum that refers to the phonological transparency of an alphabetic script, which is referred to as orthographic depth (e.g., Georgiou, Parrila, Kirby, & Stephenson, 2008; Kirby, Parrila, & Pfeiffer, 2003). In orthographies with less orthographic depth RAN remains a predictor of reading performance across grades (de Jong & van der Leij, 2003), whereas it appears to be

time-limited in an orthography with increasing orthographic depth such as English (Parrila, Kirby, & McQuarrie, 2004; Roman, Kirby, Parrila, Wade-Woolley, & Deacon, 2009). Furthermore, some recent studies suggest that RAN performance should be considered as a separate and independent risk factor rather than a phonological skill (Rodríguez, van den Boer, Jiménez, & de Jong, 2015).

THE PHONOLOGICAL DEFICIT HYPOTHESIS

DEFECTIVE PHONOLOGICAL REPRESENTATIONS?

Note that all three phonological domains in which reading disabled individuals experience problems involve the storage and retrieval of phonological information in STM or LTM. *Phonological awareness* tasks involve the storage, retrieval, and manipulation of such information in STM. The *repetition of nonwords or pseudowords* requires the storage and retrieval of phoneme strings in STM. Finally, *RAN* requires the (fast) retrieval of stored phonological information from LTM.

The similarities among these problems have led to a widely held hypothesis, which states that disabled readers' phonological processing problems are all due to a basic problem at the deepest level of phonological representation: phonemic representations (Boada & Pennington, 2006; Carroll & Snowling, 2004; Elbro, 1996; Elbro, BorstrØm, & Petersen, 1998). More particularly, these representations are believed to be of poor quality. In other words, sequences of phonemes (real words in RAN, nonwords in a nonword repetition task [NRT]) are believed to be poorly represented in these readers' LTM and STM. Hence, words and nonwords cause problems when they must be retrieved from LTM and STM, respectively, and when manipulations (like phoneme deletion) must be performed on them. In this view, fully developed phonemic representations are at the core of reading development, as they are crucial for the establishment of associations between the orthographic and phonological representations of words, more particularly, by their connections with their constituent graphemes.

Several versions of the phonological deficit hypothesis (PDH) have been proposed. The *phonological representations hypothesis* refers to the view that individuals with dyslexia have degraded phonological representations (e.g., Goswami, 2000). The *distinctness hypothesis* raises the idea that the completeness and accuracy of phonological representations is linked to the amount of phonetic detail that is stored with an item in the mental lexicon, i.e., its

distinctness (e.g., the phonological representation *secretry* is less distinct than *secretary*, e.g., Elbro, 1996). Furthermore, the *sublexical deficit hypothesis* states that the problem is situated at the level of input sub-lexical processes and, hence, also affects the phonological representations of words in the mental lexicon (e.g., Ramus, 2001; de Bree, 2007). In addition, a variety of labels have been employed to refer to the problem of qualitatively poor representations, such as *poorly specified* (Elbro & Jensen, 2005)*, indistinct* (Elbro, 1998), or *less mature* (Boada & Pennington, 2006) phonemic representations. Although the different hypotheses behind these descriptions display subtle differences, their basic assumption is that *poor* phonemic and, hence, phonological representations are key to reading problems.

Defective access to phonological representations?

Other researchers, however, have suggested that the basic problem might not be the nature of the phonological representations but the access to these representations. Ramus and Szenkovits (2008), for example, argued that the phonological representations of people with dyslexia may be normal but might take more time to access. They arrived at this conclusion by observing that the phonological problems are only demonstrable when task requirements increase the processing load on STM (e.g., when longer nonwords have to be recalled), or when conscious awareness or time constraints impose severe time restrictions on lexical access. This position is supported by the results of a study in adults with dyslexia. Combining functional magnetic resonance brain imaging (fMRI) with multi-voxel pattern analysis (MVPA), Boets et al. (2013) concluded that phonetic[2] representations are primarily held bilaterally in the primary and secondary auditory cortices and that, crucially, the quality of these representations does not differ between individuals with dyslexia and typical readers. However, their results also showed that the functional and structural connectivity between the bilateral auditory cortices and the left inferior frontal gyrus, a region involved in higher-level phonological processing, is significantly hampered in individuals with dyslexia. These findings made them suggest that individuals with dyslexia, at least adults, suffer from deficient access to *intact* phonetic representations. Although these two views (phonological representations' *quality* and *accessibility)* slightly differ, they both emphasize the important role of phoneme representations in the reading process: such representations underlie high-quality phonological representations of words and they should be easily accessible. Nevertheless, several fundamental research

[2] The authors use the term 'phonetic' representations but refer to the same type of representation that other researchers describe with the label 'phonemic'.

questions remain unanswered. For instance, as a phonological representation is a sequence of sounds, it involves several phonemic representations. Hence, both the identity of the individual speech sounds and their serial order are potentially important[3]. However, the strong research focus on the quality (or accessibility) of phonemic representations reflects an almost exclusive interest in the representation of phonemes' identity. This doctoral thesis seeks to explore whether the representation of phonemes' serial order is also a determinant of reading success. This issue will be studied in monolingual poor young and adult readers and typically developing bilingual children.

As fast access to high-quality phonological representations is required to become a good reader, it is important to have a detailed picture of these representations. How do they develop during language acquisition and what is their internal structure, more particularly, are these representations segmented into phonemes or not? The next part of this introduction discusses our knowledge of phonological representations in typical development. It also addresses the relation between these (preconscious) representations and a more explicit awareness of the sound structure of words, as indexed via phoneme awareness measures.

PHONOLOGICAL REPRESENTATIONS

Several research findings suggest that preliterate children already develop phoneme-based representations of spoken words, but this does not entail explicit awareness of these units (Swingley & Aslin, 2002; Carroll & Snowling, 2004). Perceptual properties that have a representational reality do not by themselves imply conscious awareness of these properties. For instance, our ability to know what the grammatical subject and object of a sentence are (in our mental representation of the sentence) does not imply that we are also able to consciously identify these grammatical units, i.e., the ability to understand language does not entail metalinguistic awareness. When reading the following paragraphs, it is important to keep in mind this distinction between phonemes' representational importance and a child's conscious access to these speech units.

[3] Note that the difference between identity and serial order applies both to the level of representation and to the process of lexical retrieval. For instance, a problem at the level of phoneme identity might be due to 'poorly developed' phoneme representations but also to difficulties in accessing well-developed phonemic representations – or both.

THE ROLE OF PHONEMES IN THE DEVELOPMENT OF PHONOLOGICAL REPRESENTATIONS IN CHILDREN'S FIRST LANGUAGE (L1)

Phonemes already play a role in the way very young babies perceive speech. It is well documented that newborns are sensitive to phonemic contrasts. In their first year, they gradually shift from language-general to language-specific phoneme perception. By the time they reach their first birthday, they can no longer discriminate non-native phonemic contrasts (Kuhl et al., 2008; Werker & Curtin, 2005). Due to the frequent exposure to their native language phonemes, they develop mental representations of these phonemes and even perceive non-native sounds as variants of such phonemes, due to the 'magnet effect' of their native-language phoneme categories (Kuhl, 1991; Kuhl et al., 2006). For instance, speakers of Japanese have problems hearing the difference between the English /l/ and /r/ because both are perceived as the Japanese /r/. The latter sound lies in-between the English sounds in 'perceptual space' and acts as a perceptual magnet when hearing these sounds, i.e., English /l/ and /r/ are perceived as phonetic variants of Japanese /r/.

However, the fact that very young babies can perceive all phonemic contrasts in the world's languages and gradually zoom in on the phoneme inventory of their native language or native languages (in case of simultaneous bilingualism) does not necessarily imply that phonemes are the basic encoding units in their phonological representations of spoken words. In fact, the exact nature and development of fully developed phonological representations of spoken words is not yet fully understood. The available evidence on the nature of these representations comes from both perception and production studies. There are two main views, representing opposite ideas on the importance of phonemes in young children's phonological representations.

As mentioned above, a large body of evidence suggests the existence of detailed phonological representations, with phonemes as the basic units, even in the early stages of language acquisition (Swingley & Aslin, 2002; Carroll & Snowling, 2004). Claims are based on the observation of infants' increasing sensitivity for native-like phonemes while only *listening* to that language, i.e., before the stage of active word production. In addition, other studies have shown that monolingual infants recognize words to which they have been familiarized but behave differently when these words are mispronounced at the phonemic level, e.g., *teeth* pronounced as *veeth* (Jusczyk & Aslin, 1995; but see Pater, Stager, & Werker, 2004). Research on infant speech production has revealed similar outcomes. Jaeger (1992), for instance, showed that children under the age of two already make slips of the tongue in which

single consonants or single vowels are substituted or swapped, and inferred a segmental, i.e., phoneme-based, organization of words' phonological representations. In sum, many researchers studying early word learning and word recognition/production in young infants assume that these children must have stored phonetically and phonemically detailed representations of words at an early stage in their language development.

In contrast to this idea, other researchers have suggested that infants' initial phonological representations are more robust or holistic, i.e., not segmented into smaller parts at the sub-word level (Coady & Aslin, 2003; Fowler, 1991; Metsala & Walley, 1998; Bonte & Blomert, 2004; Munson, Edwards, & Beckman, 2005). This view is based on the idea that a child's lexicon has a rather limited size with few phonological neighbors[4] in the early stages of language development. Therefore, it is assumed that detailed, i.e., segmented, representations are not needed to distinguish between different lexical items. A gradual segmentation and, thus, refinement of phonological representations would be provoked when vocabulary expands and children have to discriminate between similar-sounding words in their mental lexicon, e.g., the Dutch phonological neighbors /ta:fəl/ (table) and /wa:fəl/ (waffle). As a consequence, smaller *sublexical* units, such as syllables and, finally, phonemes would become part of a word's phonological representation. In this view, vocabulary growth is key to the refinement of initially global phonological representations. Even though this hypothesis can account for segmented phonological representations of words, one major critique is that it does not account for the phonological representations of lexically unique words, i.e., words without similar-sounding neighbor words (e.g., *echo*). These lexically unique words are not forced to be refined into a more segmented form, due to the lack of pressure from similar-sounding words.

For our purpose, the commonality between these two views is more important than their difference. Crucially, they share the claim that a child's phonological representations are segmented into a sequence of phonemes by the time it has acquired a considerable native language vocabulary, i.e., before it learns to read and write. Thus, both emphasize the importance of phoneme-based representations. As has been emphasized above, proponents of neither view claim that children have conscious access to these phonemes. On the contrary, a word's phoneme-based phonological representation is more likely to be a form of tacit knowledge, in a way that makes it quite difficult to achieve conscious access to the constituent phonemes. The only claim is that, by the time the child begins learning to read, the basic infrastructure for developing phoneme awareness is in place, i.e., the representations that are

[4] A phonological neighbor is a word that can be derived from another word by the substitution of a single phoneme, and is, hence, close in phonological space (e.g., *bark-park, cup-cap, meal-meat*).

required to learn that letters (and some letter combinations, like *ea* and *ch*) represent speech sounds, which is the basic skill behind reading and spelling.

Even though the idea of associating graphemes to phonemes suggests that children must first become aware of phonemes and then learn the written symbols for these speech sounds, the evidence suggests a different story. It appears that the letters themselves increase children's awareness that the speech signal contains the spoken counterparts of letters. Thus children learn to become aware that spoken words consist of sounds through their graphemes, which act as a 'lens' and highlight the existence of phonemes, causing explicit phoneme awareness. Indeed, many preliterate children already have (some degree of) phonemic awareness, i.e., the knowledge that letters represent speech sounds (Goswami & Bryant, 1990), as the result of being exposed to letters in kindergarten (e.g., the letters of their own name). However, their explicit phoneme awareness considerably grows as the consequence of learning to read, i.e., the individual sounds in spoken words become salient perceptual units because children are taught that each sound in a spoken word is represented by a letter or letter combination (Liberman, Shankweiler, Fischer, & Carter, 1974; Morais, Cary, Alegria, & Bertelson, 1979). The same applies to illiterate adults. For instance, Morais and colleagues (1979) demonstrated that such adults obtained much higher scores on phoneme awareness tasks after having taken a reading course. In their view, the budding awareness of phonemes in preliterate children and illiterate adults develops into a much more explicit awareness as the result of learning to read and spell.

THE ROLE OF PHONEMES IN THE DEVELOPMENT OF PHONOLOGICAL REPRESENTATIONS IN CHILDREN'S SECOND LANGUAGE (L2)

The number of bilingual children learning to read in a second language is seriously increasing (Berens, Kovelman, & Petitto, 2013). Thousands of children enter schools where instruction is provided in a language different from the one they speak at home. It is commonly reported that many of these *sequential* bilingual children are more challenged and perform worse than their monolingual peers in the early stages of second language acquisition, reading being one of them (Slavin & Cheung, 2005). Research has also revealed the importance of phonemes in their phonological representations of L2 words. Given the role of phonological representations in the literature of learning to read in L1, this does not come as a real surprise. Being monolingual in the first few years of their life, these children have already developed phoneme-sized phonological representations of the spoken words in their L1. However, later in childhood, they acquire a L2, which turns them into bilinguals. Behavioral experiments as well as neuroimaging

studies looking into the neural mechanisms supporting speech perception suggested that two developmental processes can be distinguished (e.g., Archila-Suerte, Zevin, Ramos, & Hernandez, 2013). Apparently, L2 phonemes that are similar to bilinguals' L1 phonemes, are initially assimilated to the phonemic categories of their first language, i.e., due to Kuhl's concept of 'magnet effects' (Best, 1994). This should make it difficult to accurately perceive these phonemes. For instance, Dutch speaking children who acquire English as a second language experience difficulties when having to learn the speech sound that corresponds with the grapheme *th:* a voiced [ð] as in fa*th*er, brea*th*e, and an unvoiced [θ] as in ba*th*, brea*th,* as these sounds do not exist in their L1 (Dutch) and closely approximate the voiced [z] and [v] and unvoiced [s] and [f] in Dutch). However, for all L2 phonemes that are different to bilinguals' L1 phonemes, a new phonemic representation seems to be built. As these children grow older and acquire more linguistic experience in L2, they will increase their capacity to differentiate L2 and L1 phonemes more accurately by making use of increased attention to phonetic information.

In sum, researchers largely agree that, by the time a child learns to read, it will have developed phonological representations of its L1 words that consist of phonemic segments. Children who learn to read in L2 also have developed such phoneme-based representations in L2, even though some phonemes may not be accurately represented, due to the magnet effects of phonemic representations in their L1.

There is also a broad consensus among researchers that the establishment of, and access to precise and well-defined phonemic representations are crucial for achieving phonemic awareness (e.g., Fowler, 1991; Metsala & Walley, 1998; see also Metsala, Stavrinos, & Walley, 2009), and, later, for acquiring literacy skills (Bishop & Snowling, 2004). It follows that high-quality phonemic representations (and access to them) are crucial when it comes to the study of how one becomes a (poor) reader.

THE GOAL OF THIS DISSERTATION: TOWARDS A NOVEL USE OF THE NRT

As the phonological representations of words (and their associated access mechanism) are the core concept in the study of reading problems, it is important to use reliable measures for probing the nature of these representations. Both speech perception-based (receptive) and speech production-based (expressive) measures have been extensively used to assess STM and

LTM phonological representations in children and adults. Examples of receptive measures involve the discrimination between phonological representations, for instance, the ability to detect mispronounced words (Carroll & Snowling, 2004; Sutherland & Gillon, 2005; Rvachew, Ohberg, Grawburg, & Heyding, 2003). Examples of expressive measures are picture naming and nonword repetition (e.g., Elbro, 1996; Swan & Goswami, 1997). In the following paragraphs, we will focus on the NRT as this task will be used in all studies in this doctoral thesis.

PROPERTIES AND PAST USE OF THE NRT

The NRT has gained a great deal of popularity in recent years. The task requires individuals to simply repeat auditorily presented stimuli, more particularly, novel phoneme sequences, differing in length and complexity (e.g., *flimp, mitoc, viluntar*). Consequently, a new phonological representation 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). Hence, the task can be used to study properties of verbal STM and to investigate the ability to encode, maintain, and retrieve a novel sequence of phonemes. Accordingly, the NRT has been popular in several research areas where language problems or verbal STM problems play a central role. It has been extensively used in *monolingual* populations in English, in children with speech disorders (e.g., Shriberg et al., 2005), children with specific language impairments (SLI, e.g., Archibald & Gathercole, 2007), children with genetic disorders such as Down's syndrome (Cairns & Jarrold, 2005), children with cochlear implants (Dillon, Burkholder, Cleary, & Pisoni, 2004), children with fluency disorders (Spencer & Weber-Fox, 2014), and children with reading problems (for a review, see Brady, 1997). The popularity of the task has also led to its widespread use in non-English and *bilingualism* studies, as the NRT has been adapted for speakers of many languages, including Dutch (de Bree, Rispens, & Gerrits, 2007), Swedish (Kalnak, Peyrard-Janvid, Forssberg, & Sahlen, 2014), Finnish (Service, 1992), Spanish (Summers, Bohman, Gillam, Peña, & Bedore, 2010), Turkish (Topbas, Kaçar-Kütükçu, & Kopkalli-Yavuz, 2014), French (Thordardottir & Brandeker, 2013), Icelandic (Thordardottir, 2008), Italian (Dispaldro, Leonard, & Deevy, 2013), Cantonese (Stokes, Wong, Fletcher, & Leonard, 2006), Maltese (Calleja, n.d.), Korean (Lee, Kim, & Yim, 2013), Japanese (Saito, 1995), Brazilian Portuguese (Santos & Bueno 2003), and Greek (Maridaki-Kassotaki, 2002).

It is important to be aware that this task, which taps into a child's underlying phonological representations in STM, can nonetheless be affected by lexical and sublexical knowledge in LTM, depending on the type of stimuli that is used (Brandeker & Thordardottir, 2015). LTM knowledge has been found to ease the burden on STM when the input includes familiar phonological sequences, which facilitate the storage of the phonological input in STM. A first, general indication of the influence of LTM comes from studies showing superior recall of real words relative to nonwords (Gathercole, Pickering, Hall, & Peaker, 2001; Hulme, Maughan, & Brown, 1991). This so-called lexicality effect implies a semantic advantage, i.e., the fact that familiar phonological sequences also have a meaning causes superior recall performance of words. Rather than having to remember an unfamiliar sound string, participants only need to remember a meaning, which considerably reduces the burden on STM. In addition to this, different studies in *monolingual* populations revealed that nonword repetition performance is mediated by the similarity of the nonwords to real words, i.e., a phenomenon known as the word likeness effect (Edwards, Beckman, & Munson, 2004). High-wordlike nonwords seem to be easier to recall compared to low-wordlike nonwords. This indicates that familiar sublexical information is used for encoding purposes, which requires fewer STM resources than when an unfamiliar string of phonemes must be retained in STM (Gathercole, Willis, Emslie, & Baddeley, 1992). Similar evidence comes from *bilingualism* studies, showing better nonword performance in bilinguals when the test is based on their native language compared to their less proficient second language, due to the more extensive vocabulary in their native language and, hence, their native mental lexicon (Gathercole & Baddeley, 1990; Cheung, 1996; Masoura & Gathercole, 1999). Likewise, in studies *comparing monolingual and bilingual* individuals, there is evidence that bilingual children tend to perform worse on a NRT based on their second language compared to monolinguals who speak this language as their native language, and, hence, are more sensitive to the word likeness of the nonwords (Windsor, Kohnert, Lobitz, & Pham, 2010; Messer, Leseman, Boom, & Mayo, 2010).

Together, these studies corroborate the idea that effective nonword repetition performance relies both on lexical and sublexical phonological LTM knowledge of the language on which the test is based. Clearly, such knowledge facilitates the formation, and/or storage, and/or retrieval of the representations of novel phoneme sequences that have to be repeated when performing a NRT. The contribution of LTM can be minimized by using nonwords that have a low degree of word similarity. Items of the latter type offer the best opportunity for testing participants' verbal STM. For this reason, such items will be used in all studies reported in the next chapters. However, even so, the syllables that are used in nonwords will (by

definition) be phonological sequences that also occur in words and are, hence, familiar. This issue will be taken up in the Discussions of the studies reported in Chapters 2-4.

TOWARDS AN IN-DEPTH ANALYSIS OF ERROR PATTERNS IN THE NRT

In *monolingual* populations, most nonword repetition studies yield a consistent outcome, with lower performance *in children at high-risk for reading problems* (e.g., Carroll & Snowling, 2004; Snowling, Gallagher, & Frith, 2003; de Bree et al., 2007), in *less skilled readers and children with dyslexia* (e.g., Baird, Slonims, Simonoff, & Dworzynski, 2011; Robertson & Joanisse, 2010), as well as in *(young) adult less skilled readers and adults with dyslexia* (e.g., Dietrich & Brady, 2001; Ramus et al., 2003). In *bilingual* populations, most of the research using NRTs mainly focused on typically developing bilinguals, language impaired bilinguals, or a comparison of both groups (e.g., Lee & Gorman, 2012; Windsor et al., 2010; Kohnert, Windsor, & Yim, 2006; Girbau & Schwartz, 2008). These studies revealed better NRT performance in typically developing monolinguals compared to typically developing bilinguals and better performance in typically developing bilinguals compared to language impaired bilinguals. However, none of these studies focused on bilinguals' reading ability in their second language. A single study compared bilinguals with dyslexia in their L1 and L2, and revealed that these individuals had a mild difficulty in repeating nonwords in their native language (Kannada), and a larger difficulty in their L2 (English), which was their language of instruction at school (Sunil, Carmel, & Shymala, 2008).

Most often, in nonword repetition studies, an underachievement on the *nonword accuracy score,* i.e., the number of correctly repeated nonwords in the task, is used to demonstrate poor readers' problems with the formation, temporary storage, and/or retrieval of new phonological representations in STM. This consistent underperformance of poor readers is obviously an important finding. However, in order to understand the nature of this problem, one must identify underlying factors that contribute to their larger number of error responses. Hence, an in-depth study of these errors is the major goal of this dissertation. This goal will be pursued by studying reading in L1 and L2. More particularly, the following groups will be studied: (1) adults with and without dyslexia in L1, (2) children who vary on a continuum of reading success in L1, and (3) good young readers in L1 vs. normally developing L2 readers in the same language.

This goal marks a departure from typical NRT studies, which focus on error rates but not on error types. Only a few researchers have paid attention to different error types. For

instance, some have conducted qualitative error analyses, in addition to the main accuracy analyses, to identify different error patterns in at-risk or poor readers and better readers in their native language. Different types of sublexical errors, i.e., syllable and single phoneme errors, such as insertions, omissions, and substitutions, have been analyzed (e.g., de Bree, 2007). Based on their results, it was concluded that length and syllable structure are often maintained in the output, whereas substitution errors are the most prominent error type at the phoneme level. Similar findings have been reported in *bilingualism* studies, although these samples included language impaired children (Girbau & Schwartz, 2008; Ebert, Kalanek, Cordero, & Kohnert, 2008). Taken together, the predominance of substitution errors in these monolingual reading and bilingual language studies is interesting and suggests that the problem is caused by 'forgetting' the target phoneme at one or more positions in the stimulus. Importantly, in contrast to an omission error, which indicates the failure to recall a specific phoneme, a substitution error may reflect two different types of memory failure: the loss of a phoneme's identity or the loss of a phoneme's serial order. In the former case, the phoneme does not appear in the response at all. In the latter case, the target phoneme does appear in the response but not at the correct position.

The distinction between forgetting a phoneme's identity versus its serial order in the novel phoneme string is a crucial one. It lies at the basis of all studies in this dissertation. The distinction highlights the fact that the NRT requires participants to recall two types of information. As a nonword is a string of familiar phonemes in an unfamiliar order (i.e., the phoneme string does not form a word), both the identity and the serial order of each phoneme in the nonword must be retained in and retrieved from verbal STM. Hence, recall errors in the NRT might reflect problems with either type of retention (or both). To our knowledge, no NRT studies in the field of reading research have made this distinction and analyzed the data accordingly. Interestingly, however, recent research on STM processes indicates that analyzing the responses in the light of the distinction between identity and serial order retention – as suggested by a task analysis of the NRT – might shed more light on the source(s) of participants' retention problems in this task.

THE RETENTION OF ITEMS' IDENTITY AND SERIAL ORDER IN STM

The concept of serial ordering is an important asset in human (mental) processing, as almost everything we need to remember consists of sequences: a course of events, sentences within a

narrative, words within sentences, etc. (Acheson & MacDonald, 2009). In reading research, there is partial support for the importance of serial order in phonological processing performance, and thus, reading performance, although these findings are mixed (for a recent review, see Majerus & Cowan, 2016). Research has focused on both serial order in STM and LTM, as well as on their interplay.

SEQUENCE (HEBBIAN) LEARNING MEDIATES BETWEEN STM AND LTM

One influential idea is that STM mechanisms contribute to the long-term learning of phonemic and orthographic representations via Hebbian learning (e.g., Page & Norris, 2009). The Hebb repetition effect refers to a *serial order learning effect* that demonstrates how an item sequence that must repeatedly be stored in STM progressively develops into a stable LTM trace. In the original Hebb study (Hebb, 1961), participants were asked to perform an immediate serial recall task (using digits) in which one specific sequence of digits was repeated every third trial. Hebb found that participants' recall for repeated digit sequences was significantly higher than recall for non-repeated sequences (Hebb, 1961). Driven by the knowledge that new words are made up of sequences of sublexical phonological units (e.g., syllables and phonemes), several recent studies have explicitly linked serial order learning, as observed in the Hebbian paradigm, to new word learning (e.g., Mosse & Jarrold, 2008; Page & Norris, 2009). When linking this idea to early literacy, it is suggested that recurrently processing the same grapheme sequence (i.e., a sequence of individual graphemes) progressively makes its STM representation evolve into an orthographic representation in the mental lexicon, i.e., through Hebbian learning. Concurrently, these orthographic representations would facilitate more automatic processing of unknown letter strings (Szmalec, Loncke, Page, & Duyck, 2011). When linking this idea to reading impairment, it is assumed that individuals with dyslexia suffer from a deficit in (Hebbian) serial order learning. Szmalec and colleagues (2011), for instance, found evidence for impaired serial order learning in adults with dyslexia. Participants were tested under three modalities, using the Hebb learning paradigm: (1) verbal materials were presented visually (i.e., verbal-visual), (2) verbal materials were presented auditorily (i.e., verbal-auditory), and (3) visuospatial materials were presented (i.e., visuospatial). Adults with dyslexia showed a reduced Hebb serial order learning effect compared to matched controls in all three modalities. The observation that the group with dyslexia was also hampered in the visuospatial condition, a condition without any linguistic load, led these researchers to conclude that reading impairment reflects a *domain-general deficit* in serial order learning (Szmalec et al., 2011). These results were confirmed in a recent study on adults with

dyslexia, focusing on the verbal-visual modality (see Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015, but see Staels & Van den Broeck, 2014, who could not replicate these findings).

EVIDENCE FOR A DISTINCT SERIAL ORDER MECHANISM IN STM

In addition to a number of studies highlighting the importance of serial-order learning in STM as a means for developing LTM representations of letter sequences, there is also some evidence suggesting that accurate temporal processing of serial order information in STM, i.e., STM processing that does not lead to the LTM consolidation of the STM memory trace, is key to reading development. Before discussing some of these experiments, we will consider evidence outside the domain of reading. These results suggest the existence of a separate STM mechanism for the retention of serial order information.

Current models of STM suggest that a distinction has to be made between the processing of item information and the processing of serial order information. Even though there is some controversy about the exact way in which serial order information is represented in STM (see Hitch, Fastame, & Flude, 2005; Majerus, Poncelet, Van der Linden, & Weekes, 2008), there is a consensus that a specialized STM system is dedicated to the representation of serial order information, whereas the representation of item information depends directly on the activation and quality of LTM knowledge (i.e., phonological and semantic representations). In other words, there is evidence that item and order information are coded distinctly in STM. This evidence comes from both behavioral and neuroimaging studies. For instance, an increased performance in retrieving item information (fewer item errors) but a decreased performance in retrieving the items' serial order has been shown in tasks involving the recall of phonologically similar words (relative to dissimilar ones). Likewise, an increased performance in retrieving item information but no effect on the retrieval of items' serial order has been shown in tasks involving the recall of semantically related words (relative to unrelated ones) (see Nairne & Kelley, 2004). The fact that both phonological relatedness and semantic relatedness have different effects on STM retention for items and STM retention for serial order and that, moreover, the two variables affect serial order recall differently (while having a comparable effect on item recall) highlights the existence of two types of STM mechanisms.

Evidence from neuroimaging studies is compatible with this conclusion. The data clearly show activation of phonological and semantic processing areas in the bilateral temporal lobes when recalling verbal item information during STM tasks. On the other hand, it has been shown that the retention of serial order information during STM tasks is associated with the activation

of nonlinguistic brain areas in the right intraparietal sulcus (see Majerus et al., 2008). Interestingly, these findings are in line with Szmalec et al.'s (2011) claim that the mechanism for learning serial order information is domain-general, i.e., not restricted to the Hebbian learning of verbal item sequences.

Other studies also suggest the existence of an *ordering mechanism in STM* that underlies performance in STM tasks (Gupta, 2003, 2009; Gupta, Lipinski, Abbs, & Lin, 2005). It is thought that this sequence memory plays an important role in temporarily maintaining and repeating phoneme and syllable sequences that form a (non)word. More particularly, Gupta and coworkers demonstrated primacy and recency effects for syllables in the recall of nonwords, i.e., a better recall for the first and last syllables in the syllable sequence forming a nonword. Importantly, these syllable effects were highly similar to the effects obtained outside the context of nonwords, more particularly, in the serial recall of unrelated items in a list. The latter finding suggests that STM treats the syllables in a nonword like a sequence of unrelated items in a list.

THE RELATIONSHIP BETWEEN SERIAL ORDER STM PROCESSING AND READING PERFORMANCE

Recent behavioral work directly investigated the relationship between a serial order mechanism in STM and reading acquisition (Martinez Perez, Majerus, & Poncelet, 2012a) and reading impairment (Martinez Perez, Majerus, Mahot, & Poncelet, 2012b). These studies aimed to disentangle the retention of the items in a sequence (item STM capacity) from the retention of their serial order (serial order STM capacity). To this end, the authors developed two different tasks. First, they presented an item STM task, in which children had to recall a series of monosyllabic CVC-nonwords after a filled delay. Second, they presented an order STM task, in which children first had to listen to a sequence of animal names with high lexical frequency and a low age of acquisition such as 'cat' and 'dog'. These properties facilitated their retention. Then, the children had to order a sequence of cards with pictures of the auditorily presented animal names, i.e., they did not have to retrieve item information. Based on the results with kindergarteners, they concluded that order STM capacity, but not item STM capacity, predicts independent variance in decoding abilities for reading (Martinez Perez et al., 2012a). Furthermore, in their study comparing children with dyslexia with reading-age matched and chronological-age matched children, Martinez Perez et al. (2012b) observed an inferior performance on the item STM task in children with dyslexia compared to chronological-age

matched controls, but not to reading-level matched controls. In contrast, the performance of children with dyslexia was inferior to that of both chronological-age and reading-level matched controls on the measure for serial order STM capacity.

In the same vein, Hachmann et al. (2014) compared adults with dyslexia and matched controls. Participants were compared in four conditions. In the *verbal item STM* task, pictures were presented on a screen. Upon their disappearance, participants heard a word and had to decide whether the word corresponded to one of the pictures that had been presented by pressing a yes/no button. The *nonverbal-item task* was similar to the verbal task, the only difference being that the items were nonsense drawings and participants had to decide whether the target drawing had been an item in the list. In the *verbal order task* the participants first saw a list of digits on the screen. Upon their disappearance, they had to decide whether another list of digits was presented in the same order by pressing a yes/no button. The *nonverbal-order task* was similar to the verbal order task. However, instead of digits, nonsense drawings were used. The results showed inferior performance for adults with dyslexia on both verbal and nonverbal order tasks. In contrast to Martinez Perez et al. (2012b), these authors could not find a difference between adults with and without dyslexia on the item tasks. They attributed the different outcome on the (verbal) item task between their study and Martinez Perez et al.'s studies to the nature of the verbal item STM task in the latter study. According to Hachmann et al. (2014) CVC monosyllabic nonwords are not suitable to grasp item STM processing independently of order processing, as such items require the retention of the (unfamiliar) serial order of their phonemes[5]. For that reason, they used existing words in their verbal item task.

Even though several studies suggested that individuals with dyslexia experience problems with the retention of serial order information, one study failed to find such an effect. When Staels and Van den Broeck (2014) performed the same analyses as Martinez Perez et al. (2012b) in their study with reading impaired children, they replicated their results: inferior performance for children with dyslexia on both the item STM and order STM tasks. However, a different pattern emerged when they conducted analyses that they claimed to be more appropriate. They argued that one can only demonstrate a specific serial order retention problem when individuals with dyslexia and normal readers score similarly on item retention (Staels & Van den Broeck, 2014). To achieve this goal, they brought the children's item performance under statistical control when analyzing their serial order scores, such that a group effect could not be co-determined by a difference between individuals with dyslexia and controls on item retention. Using state trace

[5] As will become clear in the following chapters, our use of the NRT set out from the same assumption, i.e., that the retention of nonwords by definition requires the retention of the serial order of the phonemes.

analyses and entering performance on the item STM task as a covariate, they were unable to find a specific deficit in serial order STM in children with dyslexia. However, when performance on the serial order task was entered as a covariate, the effect of group remained significant in children's item STM scores. The authors concluded that the item and order tasks that were used by Martinez Perez et al. (2012a, 2012b) do not measure *entirely* independent processes and that both are related to phonological processing (Staels & Van den Broeck, 2014). Moreover, and in contrast to Martinez Perez et al. (2012a; 2012b), they suggested that the partial autonomy of both STM processes cannot be ascribed to a difference in the involvement of a phonological component, but is the result of the role of nonverbal intelligence processes in serial order STM only. This conclusion was based on their finding that only serial order performance was found to correlate with a nonverbal intelligence component (Staels & Van den Broeck, 2014).

Taken together, the overall conclusion that can be drawn from all these studies is that processing item information and serial order information in STM involve, at least partially, distinct processes. With the exception of Staels and Van den Broeck (2014), several studies suggest that adequate serial order processing in STM is important for successful reading and that problems with this ability are related to dyslexia. It remains to be seen how this problem is related to the phonological deficit hypothesis, as several experiments suggest that serial order processing in STM reflects the operation of a domain-general mechanism. However, note that, even if the latter turns out to be the case, individuals with dyslexia and poor readers may still experience problems with the retention of the phonemes themselves, which would point to a problem with the quality of their phoneme representations and/or the access to these representations. In other words, a problem with the retention of serial order information in individuals with dyslexia and poor readers would not by itself discredit the phonological deficit hypothesis.

RATIONALE OF THE CONDUCTED STUDIES

Three properties stand out when considering all aforementioned studies on the importance of serial order retention for successful reading. First, they all used *different* tasks to dissociate the retention of item information and serial order information. Whereas the researchers deliberately used different tasks, more particularly, to minimize serial order processing in the item retention task and to minimize item processing in the serial order task, the use of different materials unavoidably causes a mismatch on several variables (e.g., different materials). This complicates the interpretation of the results. Hence, rather than using a between-task and between-item design

to study the two types of STM retention, we opted for a within-manipulation. More particularly, in all empirical studies reported in the next chapters, we used a single task and a single set of materials to study the two types of STM retention.

Second, the authors of the abovementioned studies focused on the levels of digits, words, or syllables. None of the studies aimed at the phoneme level. When the focus is on the relationship between STM-related abilities and reading performance, this is unfortunate, considering the widely accepted view that individuals with dyslexia have major problems with the representation of and/or access to the representation of phonemes (see the PDH above). Focusing on the phoneme level would make it possible to simultaneously investigate the phonological deficit hypothesis and the hypothesis that individuals with dyslexia have problems with a (possibly domain-general) ordering mechanism in STM. As Gupta and coworkers have shown, the same STM mechanisms are at work when maintaining novel phoneme sequences in STM (i.e., nonwords in a NRT) as when maintaining a list of nonverbal items (see above). Hence, novel phoneme sequences seem to be an ideal instrument for studying the relative importance of both item retention, as reflected in the recall of a nonword's phonemic constituents, and serial order retention, as reflected in the recall of the relative order of the correctly recalled phonemes. If phonological processing problems in poor readers are linked to a basic problem in the formation, access, and/or retrieval of phonemic representations, as suggested by many studies, these readers should encounter problems with the retention of the phonemes themselves. If the problems confronting poor readers are (also) influenced by problems in the retention of serial order information, this should be reflected in their poor performance on the retention of the phonemes' serial position. Note that the two types of retention problems may co-occur.

Third, all previous studies primarily focused on monolingual poor and typically developing readers. However, there are two reasons why it is valuable to study sequential bilingual children learning to read in a second language as well. (1) Given the aforementioned hypothesis that item information depends on the quality of LTM knowledge whereas retention of serial order is thought to be exclusively related to a specialized system in STM (e.g., Majerus et al., 2008), it seems reasonable to predict that monolinguals, who have a high language proficiency in L1, will be better in the retention of the phonemes in nonwords based on their L1 than sequential bilinguals, whose L2 language proficiency is lower compared to native speakers. (2) If retention of serial order is exclusively related to a specialized system in STM (e.g., Majerus et al., 2008) and if poor reading is (also) characterized by problems in the retention of serial order information in STM, one would not expect to find serial order problems when examining typically developing bilingual children learning to read in their second language.

The goal of this dissertation is not to investigate whether the underlying cause of poor NRT performance in individuals with dyslexia is due to their problems with the retention of phonemes (or access to such representations) *or* to their problems with the retention of serial order (or access to its representation). As the evidence reviewed above suggests, it is quite possible that they experience problems with both the recall of phonemes' identity and their serial order. The rationale behind the studies reported in Chapters 2-4 is that the use of nonwords in a NRT makes it possible to assess both the role of phonemic representations and the role of a serial order mechanism in STM. In the next empirical chapters, the NRT will be used for this purpose.

More specifically, in all studies, the Dutch adaptation (Scheltinga, 2003) of the Children's Test of Nonword Repetition (Gathercole, Willis, Baddeley, & Emslie, 1994) will be used. This Dutch NRT test comprises 48 nonwords which are constructed along a 4 (syllable length) x 2 (word complexity) x 2 (phonological similarity) design. Hence, these nonwords are organized into four syllable categories, varying in length from two to five syllables, 12 items per category. Important to note, however, is that in our studies we will not include two-syllable items to guard against ceiling effects in production. For each word length, the test also controls for the influence of word complexity on nonword recall. To this end, half of the nonwords have different consonants at the two sides of the syllable boundary, giving rise to closed syllables (e.g., /lem.ros.pag/). In the other half of the nonwords, the same consonant occurs at both sides of the syllable boundary, giving rise to an open syllable, i.e., the consonant is pronounced only once (e.g., /pip.pok.ket/). Within each half, influence of phonological similarity on nonword recall is also controlled for. To this end, half of the nonwords are phonologically similar, thus constructed of syllables with consonants which share phonetic features (e.g., /nim.mun.naf/ or /nom.lun.fam/). Phonologically dissimilar nonwords contain syllables with contrasting consonants (e.g., /gik.kal.lom/ or /pig.dul.mek/). A vowel occurred only once in each nonword. Appendix A provides an overview of the nonwords, ordered along the above described categories.

Given (a) the differences in pronunciation between Dutch in the Netherlands and Dutch in Flanders and (b) our use of the NRT with Flemish participants, the available Flemish recording, made by De Smedt (see for instance, De Smedt, Verschaffel, & Ghesquière, 2004; De Smedt, 2006) will be used. All nonwords observe the Dutch pattern of placing stress on the penultimate syllable. The nonwords and two test items, pronounced with a consistent rate by a professional speech language pathologist, are recorded on a CD by using Cool Edit. A short beep of 200 ms centered around 440Hz precedes each nonword, followed by one second of silence. Each word is presented only once. In contrast to Gathercole et al. (1994), in this Flemish recording, the nonwords are categorized by syllable length and randomized within these categories, starting with

two-syllable items. This was done to minimize the potential influence of attentional factors on the nonword recall. This recorded list will be used and will be presented auditorily via a laptop. As we will not include two-syllable items in our studies, our recording will start with the two test items, followed by the three-syllable items.

Study 1 will focus on adults with and without dyslexia, whereas study 2 will focus on poor and better third grade readers. Study 3 will focus on typically developing second language readers.

Chapter 2 deals with the question whether adults with and without dyslexia perform similarly on the two types of STM retention discussed above. A sample of adults with dyslexia and a control group performed a NRT to assess their STM performance on novel phoneme sequences. The phonological deficit hypothesis predicts that individuals with dyslexia will have more problems with item retention than the control group. If individuals with dyslexia also encounter problems with a serial order mechanism in STM, their retention of the serial order of phonemes in nonwords should also be worse relative to the control group.

Chapter 3 addresses both retention mechanisms in a sample of poor and better reading third-graders. In this study, a regression design was used rather than a factorial manipulation of the participants' reading level. We used the same NRT as in Chapter 2. The phonological deficit hypothesis predicts worse performance on item retention with decreasing reading performance. If poor reading is (also) related to problems with the retention of serial order information, the retention of phonemes' serial order in nonwords should also decrease with decreasing reading performance.

Chapter 4 compares the two types of retention abilities in typically developing monolingual (Dutch) and sequential bilingual (Turkish-Dutch) children learning to read in their second language. We were interested whether bilingual versus monolingual differences would be reflected in different error rates at the phoneme level. The above prediction that a difference would emerge at the level of phoneme retention but not at the level of phonemes' serial order retention is obviously only valid if test items are used that comply with the phonotactic rules of both languages. Such items will activate knowledge of these rules in LTM, and, hence, assist in the encoding of the nonwords in both groups. Despite some phonotactic differences between both languages used in Chapter 4, the items in the (Dutch) NRT complied with the phonotactic properties of both Turkish and Dutch. As LTM knowledge made it possible for both groups to recognize the nonwords' phonotactic structure and the STM process for storing serial order was thought to be intact in both groups, we did not expect a difference in serial order retention between monolinguals and bilinguals.

In **Chapter 5**, the General Conclusions of the experimental studies are presented. They summarize the main findings and discuss their theoretical and practical implications. In addition, we consider a number of limitations of our studies and outline potential avenues for future research.

In sum, this dissertation addresses the following research questions:

- Study 1:
    - Do adults with and without dyslexia differ in their ability to retain phonemes' identity when processing phonological information in STM?
    - Do adults with and without dyslexia differ in their ability to retain phonemes' serial order when processing phonological information in STM?
- Study 2:
    - Do young readers' literacy skills influence the retention of phonemes' identity when processing phonological information in STM?
    - Do young readers' literacy skills influence the retention of phonemes' serial order when processing phonological information in STM?
- Study 3:
    - Do monolingual and sequential bilingual children differ in their ability to retain phonemes' identity when processing native (monolinguals)/non-native (bilinguals) phonological information in STM?
    - Do monolingual and sequential bilingual children differ in their ability to retain phonemes' serial order when processing native (monolinguals)/non-native (bilinguals) phonological information in STM?

REFERENCE LIST

Acheson, D.J. & MacDonald, M.C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135*, 50-68.

Archibald, L.M.D., & Gathercole, S.E. (2007). The complexities of complex span: Specifying working memory deficits in SLI. *Journal of Memory and Language, 57,* 177-194.

Archila-Suerte, P., Zevin, J. Ramos, A. & Hernandez, A.E. (2013). The Neural Basis of Non-Native Speech Perception in Bilingual Children. *Neuroimage, 67,* 51-63.

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

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

Berens, M., Kovelman, I., & Petitto, L.A. (2013). Learning to read in two languages: Should bilingual children learn reading in two languages at the same time or in sequence? Evidence of a bilingual reading advantage in children in bilingual schools from monolingual English-only homes. *Bilingual Research Journal, 36,* 35-60. doi: 10.1080/15235882.2013.779618

Best, C. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation hypothesis. In J.C. Goodman & H.C. Nusbaum (Eds). *The development of speech perception: the transition from speech sounds to spoken words* (pp. 167–224). MIT Press. Cambridge.

Bishop, D.V.M., & Snowling, M.J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130,* 858–88. doi 10.1037/0033-2909.130.6.858

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., Op de Beeck, H.P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., … Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science, 342,* 1251–1254. doi:10.1126/science.1244333

Bogaerts, L., Szmalec, A., Hachmann, M.W., Page, M.P.A., & Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities, 43–44,* 106–122.

Bonte, M.L., & Blomert, L.P.M. (2004). Developmental changes in ERP correlates of spoken word recognition during early school years: A phonological priming study. *Clinical Neurophysiology, 115,* 409-423. doi: 10.1016/S1388-2457(03)00361-4

Brady, S.A. (1997). Ability to encode phonological representations: An underlying difficulty of poor readers. In B. Blachman (Ed.). *Foundations of Reading Acquisition and Dyslexia* (pp. 21–47). Mahwah, NJ: Lawrence Erlbaum Associates.

Brandeker M., & Thordardottir, E. (2015). Language exposure in bilingual toddlers: Performance on nonword repetition and lexical tasks. *American Journal of Speech-Language Pathology, 24,* 126–138. doi:10.1044/2015_AJSLP-13-0106

Cairns, P., & Jarrold, C. (2005). Exploring the correlates of impaired non-word repetition in Down syndrome. *British Journal of Developmental Psychology, 23,* 401–416. doi: 10.1348/026151005X26813

Calleja, N. (n.d.). Test of Maltese Non-Word Repetition in Noise. *Unpublished assessment.* Malta: University of Malta.

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

Cheung, H. (1996). Nonword span as a unique predictor of second language vocabulary learning. *Developmental Psychology, 32,* 867–873.

Coady, J., & Aslin, R. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language, 30,* 441–469.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

de Bree, E. (2007). *Dyslexia and phonology: A study of the phonological abilities of Dutch children at risk of dyslexia.* Utrecht: LOT dissertation 155.

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

de Jong, P.F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91,* 450-476.

de Jong, P.F., & van der Leij, A. (2002). Effects of Phonological Abilities and Linguistic Comprehension on the Development of Reading. *Scientific Studies of Reading, 6,* 51-77. doi:10.1207/S1532799XSSR0601_03

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. doi:10.1037/0022-0663.95.1.22

De Smedt, B., Ghesquière, P., & Verschaffel, L. (2004, April). *Working memory and arithmetic abilities in seven and eleven year old children.* Poster presented at the Second Working Memory Symposium. Beaune, France.

De Smedt, B. (2006). *Mathematical disabilities in children with velo-cardio-facial syndrome: a further delineation of the cognitive phenotype.* KU Leuven, Faculty of Psychology and Educational Sciences: Unpublished PhD thesis.

Dietrich, J.A., & Brady, S.A. (2001). Phonological representations of adult poor readers: An investigation of specificity and stability. *Applied Psycholinguistics, 22,* 383–418.

Dillon, C., Burkholder, R. , Cleary, M., & Pisoni, D. (2004). Nonword repetition by children with cochlear implants: Accuracy ratings from normal-hearing listeners. *Journal of Speech, Language, and Hearing Research, 47,* 1103–1116.

Dispaldro, M., Leonard, L., & Deevy, P. (2013). Real-word and nonword repetition in Italian-speaking children with specific language impairment: A study of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research, 56,* 323–336.

Ebert, K., Kalanek, J., Cordero, K., & Kohnert, K. (2008). Spanish nonword repetition: Stimuli development and preliminary results. *Communication Disorders Quarterly, 29,* 67–74. doi: 10.1177/1525740108314861.

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.

Ehri, L. (2014). Orthographic Mapping in the Acquisition of Sight Word Reading, Spelling Memory, and Vocabulary Learning. *Scientific Studies of Reading, 18,* 5–21. doi: 10.1080/10888438.2013.819356

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

Fowler, A.E. (1991). How early phonological development might set the stage for phoneme awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97–117). Hillsdale, New Jersey: Erlbaum.

Foy, J.G., & Mann, V.A. (2001). Does strength of phonological representations predict phonological awareness in preschool children? *Applied Psycholinguistics, 22,* 301-325.

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

Gathercole, S.E., Pickering, S.J., Hall, M., & Peaker, S.J. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology, 45A,* 1-30.

Gathercole, S.E., Willis, C., Emslie, H., & Baddeley, A.D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28,* 887 -898.

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

Georgiou, G., Parrila, R., Kirby, J.R., & Stephenson, K. (2008). Rapid naming components and their relationship with phonological awareness, orthographic knowledge, speed of processing and different reading outcomes. *Scientific Studies of Reading, 12,* 325–350. doi: 10.1080/10888430802378518

Gillon, G.T. (2004). *Phonological awareness: From research to practice.* New York, NY: Guilford Press.

Girbau, D., & Schwartz, R.G. (2008). Phonological working memory in Spanish-English bilingual children with and without Specific Language Impairment. *Journal of Communication Disorders, 41,* 124-145.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Goswami, U. (2000). Phonological and lexical processes. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 251-267). Mahwah, New Jersey: Lawrence Erlbaum Associates.

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

Gupta, P. (2009). A computational model of nonword repetition, immediate serial recall, and nonword learning. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 108–135). Hove, England: Psychology Press.

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

Harris, L.N., & Perfetti, C.A. (2017). Individual differences in phonological feedback effects: Evidence for the orthographic recoding hypothesis of orthographic learning. *Scientific Studies of Reading, 21,* 31-45.

Hebb, D.O. (1961). Distinctive features of learning in the higher animal. In J.F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford, UK: Blackwell.

Hitch, G.J., Fastame, M.C., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory, 3-4,* 247–258.

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

Humphreys, G.W., Evett, L.J., & Taylor, D.E. (1982). Automatic phonological priming in visual word recognition. *Memory & Cognition, 10,* 576-590.

Jaeger, J.J. (1992). Not by the chair of my hinny hin hin: Some general properties of slips of the tongue in young children. *Journal of Child Language, 19,* 335–336.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29,*1-23.

Kalnak, N., Peyrard-Janvid, M., Forssberg, H., & Sahlén, B. (2014). Nonword Repetition – A Clinical Marker for Specific Language Impairment in Swedish Associated with Parents' Language-Related Problems. *PLoS ONE, 9*(2), e89544. doi.org/10.1371/journal.pone.0089544

Kirby, J.R., Parrila, R., & Pfeiffer, S.L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology, 95,* 453–464. doi:10.1037/0022-0663.95. 3.453.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with primary language impairment from typical bilinguals? *Journal of Learning Disabilities Research & Practice, 21,* 19–29.

Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophic Transactions of the Royal Society B, 369,* 979-1000.

Kuhl, P.K. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perceptual Psychophysics, 50,* 93-107.

Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9,* F13–F21.

Lee, S.A.S., & Gorman, B.K. (2012). Nonword repetition performance and related factors in children representing four linguistic groups. *International Journal of Bilingualism, 17,* 479–495.

Lee, H.J., Kim, Y.T., & Yim, D. (2013). Non-word repetition performance in Korean-English bilingual children. *International Journal of Speech-Language Pathology, 15,* 375-382.

Lervag, A., Braten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology, 45,* 764–781. doi: 10.1037/A0014132

Liberman, L.Y., Shankweiler, D., Fischer, F.W., & Carter, B.J. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201- 212.

Lyon, G.R., Shaywitz, S.E., & Shaywitz, B.A. (2003). Defining dyslexia, comorbidity, & teachers' knowledge of language and reading. *Annals of Dyslexia, 53,* 1-14.

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

Maridaki-Kassotaki, K. (2002). The relation between phonological memory skills and reading ability in Greek-speaking children: Can training of phonological memory contribute to reading development? *European Journal of Psychology of Education, 17,* 63–73.

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

Masoura, E. & Gathercole, S. E. (1999). Phonological short-term memory and foreign vocabulary learning. *International Journal of Psychology, 34,* 383-388.

Messer, M., Leseman, P., Boom, J., & Mayo, A. (2010). Phonotactic probability effect in nonword recall and its relation with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology, 105,* 306–323. doi: 10.1016/j.jecp.2009.12.006.

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: Lawrence Erlbaum Associates.

Metsala, J.L., Stavrinos, D., & Walley, A.C. (2009). Children's spoken word recognition and contributions to phonological awareness and nonword repetition: A one-year follow-up. *Applied Psycholinguistics, 30,* 101-121. doi:10.1017/S014271640809005X

Morais J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,*323-331.

Mosse, E.K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the acquisition of phonological forms in children. *Quarterly Journal of Experimental Psychology, 61,* 505–514.

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

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)

Nairne, J.S., & Kelley, M.R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50,* 113-133.

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.

Parrila, R.K., Kirby, J.R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development? *Scientific Studies of Reading, 8*, 3-26.

Pater, J., Stager, C.L., & Werker, J.F. (2004). The lexical acquisition of phonological contrasts. *Language, 80,* 361–379.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383. doi: 10.1080/10888430701530730

Perfetti, C.A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backwar masking and priming. *Journal of Memory and Language, 30,* 473-485.

Perfetti, C.A., & Hart, L. (2001). The lexical basis of comprehension skill. In D.S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC, US: American Psychological Association. doi: 10.1037/10459-000

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

Ramus, F. (2001). Dyslexia. Talk of two theories. *Nature, 5,* 393-412.

Rastle, K. & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? Cognitive Psychology, 53, 97-145.

Rayner, K., Foorman, B., Perfetti, C.A., Pesetsky, D., & Seidenberg, M.S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2, 3*1–74.

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

Rodríguez, G., van den Boer, M., Jiménez, J.E., & de Jong, P.F. (2015). Developmental changes in the relations between RAN, phonological awareness, and reading in Spanish children. *Scientific Studies of Reading, 19,* 273-288. doi: 10.1080/10888438.2015.1025271

Roman, A.A., Kirby, J.R., Parrila, R.K., Wade-Woolley, L., & Deacon, S.H. (2009). Toward a comprehensive view of the skills involved in word reading in Grades 4, 6, and 8. *Journal of Experimental Child Psychology, 102,* 96–113.

Rvachew, S., Ohberg, A., Grawburg, M., & Heyding, J. (2003). Phonological awareness and phonemic perception in 4-year-old children with delayed expressive phonology skills. *American Journal of Speech-Language Pathology, 12,* 463-471.

Saito, S. (1995). Effects of pronounceability and articulatory suppression on phonological learning. *Perceptual and Motor Skills, 81,* 651–657.

Santos, F., & Bueno, O. (2003). Validation of the Brazilian Children's Test of Pseudoword Repetition in Portuguese speakers aged 4 to 10 years. *Brazilian Journal of Medical and Biological Research, 36,* 1553–1547.

Scheltinga, F. (2003). *The Dutch Nonword Repetition Test.* University of Amsterdam: Unpublished manuscript.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45,* 21–50.

Share, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Share, D.L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72,* 95–129.

Shriberg, L.D., Lewis, B.A., Tomblin, J.B., McSweeny, J.L., Karlsson, H.B., & Scheer, A.R. (2005). Toward diagnostic and phenotype markers for genetically transmitted speech delay. *Journal of Speech, Language, and Hearing Research, 48*, 834- 852.

Slavin R., & Cheung A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75,* 247–284.

Snowling, M.J., & Melby-Lervag, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, 142, 498-545.

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

Spencer, C., & Weber-Fox, C. (2014). Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence of stuttering. *Journal of Fluency Disorders, 41,* 32-46. doi: 10.1016/j.jfludis.2014.06.001

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

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

Stokes, S., Wong, A., Fletcher, P., & Leonard, L.B. (2006). Non-word repetition and sentence repetition as clinical markers of specific language impairment. The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49,* 219–236.

Summers, C., Bohman, T., Gillam, R., Peña, E., & Bedore, L. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders, 45,* 480–493.

Sunil Kumar, R., Carmel, J., & Shyamala, K.C. (2008). Non-word repetition skills in bilingual (Kannada–English) dyslexic children. *Interdisciplinary Journal of Linguistics, 1,* 95-102.

Sutherland, D., & Gillon, G. (2005). Assessment of Phonological Representations in Children With Speech Impairment. *Language Speech and Hearing Services in Schools, 36,* 294-307.

Swan, D., & Goswami, U. (1997). Phonological Awareness Deficits in Developmental Dyslexia and the Phonological Representations Hypothesis. *Journal of Experimental Child Psychology, 66,* 18-41. doi: 10.1006/jecp.1997.2375

Swingley, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science,* 13, 480–484.

Szmalec, A., Loncke, M., Page, M.P.A., & Duyck, W. (2011). Order or Disorder? Impaired Hebb Learning in Dyslexia. *Journal of Experimental Psychology-Learning Memory and Cognition, 37,* 1270–1279.

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46,* 1-16.

Thordardottir, E. (2008). Language-specific effects of task demands on the manifestation of specific language impairment: A comparison of English and Icelandic. *Journal of Speech, Language, and Hearing Research, 51,* 922–937.

Topbaş, S., Kaçar-Kütükçü, D., & Kopkalli-Yavuz, H. (2014) Performance of children on the Turkish Nonword Repetition Test: Effect of word similarity, word length, and scoring. *Clinical Linguistics & Phonetics, 28,* 602-16. doi: 10.3109/02699206.2014.927003.

Vellutino, F.R., Fletcher, J.M., Snowling, M.J., & Scanlon, D.M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2–40. doi: 10.1046/j.0021-9630.2003.00305.x

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

Werker, J.F., Curtin, S. (2005). PRIMIR: A developmental model of speech processing. *Language Learning and Development, 1,* 197–234.

Windsor, J., Kohnert, K., Lobitz, K.F., & Pham, G.T. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology, 19,* 298–310.

Wolf, M., & Bowers, P. (1999). The "Double Deficit Hypothesis" for the developmental dyslexias. *Journal of Educational Psychology, 91,* 1-24.

CHAPTER 2

SHORT-TERM MEMORY PROBLEMS FOR PHONEMES' SERIAL ORDER IN ADULTS WITH DYSLEXIA: EVIDENCE FROM A DIFFERENT ANALYSIS OF THE NONWORD REPETITION TASK

The content of this chapter has been submitted for publication to Applied Psycholinguistics. Schraeyen, K., Van der Elst, W., Geudens, A., Ghesquière, P., Sandra, D. Short-term Memory Problems for Phonemes' Serial Order in Adults with Dyslexia: Evidence from a Different Analysis of the Nonword Repetition Task.

ABSTRACT

Many studies show that poor readers make more errors in nonword repetition than better readers. Although this finding is generally linked to the lower quality of poor readers' phonological representations in verbal STM, the nature of this poor performance remains unclear. We addressed this issue by focusing on two types of phoneme-related performance in a Nonword Repetition Task (NRT): (a) retention of phonemes irrespective of their serial order (phoneme identity) and (b) retention of correctly recalled phonemes' serial order (serial order). We tested 91 young adults with and without dyslexia. Generalized linear mixed effects models demonstrated that controls outperformed individuals with dyslexia in the retention of phonemes' serial order but failed to detect a difference in the retention of phonemes' identity. These findings are discussed in terms of the nature of or access to phonological representations but also in terms of another concept that has recently been advanced in the literature: a specialized serial order mechanism in verbal STM. We also consider the possibility that individuals with dyslexia may be less sensitive to phonotactic constraints.

INTRODUCTION

Researchers widely agree that poor readers exhibit problems with phonological skills (Vellutino, Fletcher, Snowling, & Scanlon, 2004; Ramus, 2003; Snowling, 2000; Sprenger-Charolles, Colé, & Serniclaes, 2006), resulting in low performance on phonological tasks. One task that has frequently been used to demonstrate problems with the coding and retrieval of phonological information in verbal STM is the *Nonword Repetition Task* (NRT). In the NRT, participants must repeat nonwords varying in phoneme length and complexity. The NRT reflects a sequence of cognitive processes: recognizing and discriminating the sequence of speech sounds in the acoustic signal, temporarily storing the thus derived phonological representation in verbal STM, i.e., a sequence of familiar phonemes in a specific, unfamiliar serial order, and finally retrieving the representation from verbal STM to make an articulatory response (Munson, Swenson, & Manthei, 2005; Rispens & Parigger, 2010). Previous studies with the NRT typically show that individuals with dyslexia or children at risk for dyslexia are less accurate in repeating nonwords, i.e., they repeat significantly fewer nonwords correctly than individuals without dyslexia (with children, e.g., Baird, Slonims, Simonoff, & Dworzynski, 2011; Robertson & Joanisse, 2010; Schraeyen, Geudens, Ghesquière, Van der Elst, & Sandra, 2017; with young adults, e.g., Dietrich & Brady, 2001; Ramus et al., 2003;). Most studies refer to poor performance on the NRT as a demonstration that impaired readers have problems with the coding and/or temporary storage, and/or retrieval of phonological information in verbal STM and thus, argue that there is an association between these skills on the one hand and reading performance on the other hand. However, this finding leaves open the question why individuals with dyslexia find it more difficult to repeat unfamiliar phoneme strings. What is it about the NRT that makes it more difficult for individuals with dyslexia?

To identify potential factors that underlie the observed differences, we set out from the following characterization of the NRT. To correctly reproduce a nonword or novel sequence of speech sounds, participants must have a correct retention of all phonemes appearing in the nonword and of the position of each phoneme in the serial order of these phonemes. Given this characterization, one might wonder whether the extensively reported differences between individuals with and without dyslexia on NRT performance should be linked to their ability to retain phonemes' identity, their serial order, or both. This question is at the heart of the present paper. Are persons with dyslexia more prone to forgetting constituent phonemes' *identity*, phonemes' *serial order*, or *both* pieces of information?

A lot of research seems to suggest that the capacity to *identify phonemes* is involved. Indeed, numerous researchers describe the problems on tasks tapping phonological skills, such as the NRT, due to the nature of underlying phonemic representations of words in individuals with dyslexia (Boada & Pennington, 2006; Caroll & Snowling, 2004; Elbro, 1996; Elbro, BorstrØm, & Petersen, 1998; Ramus & Szenkovits, 2008), which some refer to as *poorly specified* (Elbro & Jensen, 2005), others as *indistinct* (Elbro, 1998), or still others as *less mature* (Boada & Pennington, 2006). These poor-quality representations of words' phonemic structure in the mental lexicon would have a negative effect on the development of the conscious awareness of their constituent phonemes (Foy & Mann, 2001), which would cause problems with the coding, retrieval, and manipulation of phonemic information in verbal STM (McDougall & Hulme, 1994; Torgesen, 1985). As the development of a conscious awareness of words' constituent phonemes is also required for learning to read (Castles & Coltheart, 2004), the poor quality of words' phonemic representations would link poor performance on both reading tasks and the NRT task in individuals with dyslexia.

Other researchers have proposed that the basic problem is not the *nature* of the phonemic representations but the *access* to these representations. Ramus and Szenkovits (2008), for example, argued that the phonemic representations of people with dyslexia are normal, and that the phonological problems they experience are only demonstrable when task requirements increase the processing load on verbal STM, i.e., when only limited resources are available for access purposes. Boets et al. (2013) recently combined functional magnetic resonance brain imaging with multivoxel pattern analysis and functional and structural connectivity analysis. They demonstrated that the neural quality of phonetic representations is intact in adults with dyslexia whereas the functional and structural connectivity patterns between their frontal and temporal language areas is hampered. These results are compatible with the view that persons with dyslexia experience problems with the access to otherwise intact phonemic representations. Even though this is quite a different account of the causal factor behind dyslexia, what matters for our purpose is that it predicts the same behavioral consequence as the idea of poor phonemic representations: problems in tasks where performance relies on the ability to identify phonemes.

Is there any evidence that individuals with dyslexia have problems with the retention of *phonemes' serial order*? Even though there is partial support that these individuals have a weaker ability to recall serial order in memory tasks these findings are mixed. Research has focused on both serial order in STM and LTM. In the domain of STM, Gupta and colleagues provided evidence for a sequence mechanism in STM by showing within-nonword primacy

effects (better recall of first relative to middle syllable positions) and recency effects (better recall of final relative to middle syllable positions) (Gupta, 2003, 2009; Gupta, Lipinski, Abbs, & Lin, 2005). In recent work, Martinez Perez, Majerus, and Poncelet (2012a) and Martinez Perez, Majerus, Mahot, and Poncelet (2012b) further explored the relationship between a basic sequence memory mechanism and reading acquisition and reading impairment. In a longitudinal study, they aimed to disentangle readers' retention ability for individual items (item STM capacity) from their retention ability for the sequential order of these items (serial order STM capacity). To this end, they developed two tasks. Firstly, they presented an item STM task, in which children had to recall as many monosyllabic CVC-nonwords after a filled delay as they could, serial order recall being irrelevant. Secondly, they presented an order STM task, in which children first had to listen to a sequence of animal names with high lexical frequency and low age of acquisition (making item encoding easy), such as 'cat', 'dog'. Afterwards, the children had to order a sequence of cards with pictures of the auditorily presented animal names. The authors found that order verbal STM capacity, but not item verbal STM capacity, predicted independent variance in children's decoding abilities (Martinez Perez et al., 2012a). Furthermore, they observed that children with dyslexia obtained inferior performance than both chronological-age and reading-age matched controls on the measure for serial order STM. They also found inferior performance on the item STM task for children with dyslexia compared to chronological-age matched controls, but not to reading-level matched controls (Martinez Perez et al., 2012b). Together, these findings suggest that the problem in children with dyslexia is not situated at the level of item retention but at the level of serial order retention. Hachmann, Bogaerts, Szmalec, Woumans, Duyck, and Job (2014) found similar results, problems with serial order verbal STM, but not with item verbal STM in young adults with dyslexia. However, Staels and Van den Broeck (2014) claimed they could not replicate these results. Using state trace analysis, these authors argued to perform a more direct comparison of serial order STM performance after equating participants with and without dyslexia on item STM performance. Based on these analyses, they were unable to find a specific deficit in serial order STM in children with dyslexia.

In the domain of LTM, some recent studies revealed a significant contribution of serial order memory in differentiating typical and impaired readers, using a Hebbian learning paradigm. Hebbian learning refers to a *serial order learning effect* that reveals how a sequence of individual units (i.e., digits, phonemes, graphemes) is learned in STM and progressively develops into a more stable LTM representation. Using this paradigm, Szmalec, Loncke, Page, and Duyck (2011) and Bogaerts, Szmalec, Hachmann, Page, and Duyck (2015) concluded that

adults with dyslexia suffer from a deficit in (Hebbian) serial order learning. The observation that individuals with dyslexia were also hampered in serial learning when using non-linguistic stimuli (i.e., nonsense drawings), led these researchers to conclude that reading impairment reflects a *domain-general deficit* in serial order learning, i.e., a deficit that is not restricted to the encoding of phonemic representations (Szmalec et al., 2011; Bogaerts et al., 2015; but see Staels & Van den Broeck, 2015).

However, it is important to note that, in studies addressing the retention of phonological information, all measures of item or serial order retention adopted so far mainly focused on the level of whole items (like nonwords) or syllables. We are not aware of any study that targeted these two types of retention abilities at the phoneme level, except for one recent study by Schraeyen et al. (2017). This is somewhat surprising, as the first line of research discussed above established the importance of accurate phoneme identification in individuals with and without dyslexia. Hence, studying participants' retention ability of individual phonemes might still reveal that individuals with dyslexia experience a problem with item verbal STM capacity for phoneme-sized units. Additionally, or alternatively, they may also have problems with the immediate recall of phonemes' serial order.

In our research, we wanted to investigate two hypotheses that we derived from our earlier characterization of the NRT: (a) individuals with dyslexia have problems with the retention of phonemes' identity and (b) individuals with dyslexia have problems with the retention of phonemes' serial order. Possibly, they encounter both problems. Recently, Schraeyen et al. (2017) reported a similar experiment with third-graders representing a continuum of literacy skills. They found that literacy skills do not predict children's capacity for retaining phonemes' identity (after partialling out the variability due to serial order recall). However, literacy skills did predict the capacity for retaining phonemes' serial order (after partialling out the variability due to identity recall).

Hence, our purpose was to investigate the problem(s) underlying poor phoneme-related skills in individuals with dyslexia. Firstly, we compared the performance of individuals with dyslexia and controls on their retention scores for entire nonwords, i.e., the kind of comparison that has been done previously when a NRT is used. We expected our findings to replicate previous findings. Secondly, the effect of group was investigated on the two phoneme-related factors that underlie overall performance, i.e., the retention of individual phonemes' identity and the retention of phonemes' serial order. In addition, we explored the relationship between both types of retention ability.

METHOD

PARTICIPANTS

Two groups of native Dutch-speaking adults took part in this experiment: 45 adults with dyslexia (mean age = 18.0, SD = 8 months) and 46 typical adult readers (mean age = 18.03, SD = 6 months). The group of individuals with dyslexia had been diagnosed by specialists in recognized centers for reading disabilities, which means that they met all clinical criteria used in Flanders (scoring below percentile 10 on both a standardized word reading test and a nonword reading test, taking both accuracy and speed into account; normal intelligence; resistance to intervention; exclusion of alternative explanations). The participants in the control group were age-matched (group-wise, within a year band) young adults with a similar academic background and no history of reading difficulties or other language problems.

MATERIALS AND DESIGN

The Dutch adaptation (Scheltinga, 2003) of the Children's Test of Nonword Repetition (Gathercole et al., 1994) was used. Although the original Dutch test consists of 48 nonwords, organized in four categories on the basis of the number of syllables in the nonwords, we only used 36 nonwords (see Appendix A) organized in three categories, varying in length from three- to five-syllables, 12 items per category. All items comply with Dutch phonotactic rules. Neither the nonwords themselves nor their constituent syllables correspond to existing words, to minimize the use of LTM representations and, hence, to focus mainly on the retention process for sublexical material in verbal STM. All nonwords followed the Dutch pattern of placing stress on the penultimate syllable (Booij, 1995).

PROCEDURE

Participants were tested individually. A digitalized recording, made by a female native speaker of Flemish, were presented once to the participant through headphones. All responses were recorded (with Audacity software, version 1.2.4) to be able to listen again to the audio files when necessary and to exclude that the scoring of responses was influenced by memory factors of the experimenter.

The participants were told that they would hear a nonword and asked to repeat it back immediately. We started with the three-syllable items, which were followed by the four-syllable items, which were in turn followed by the five-syllable items. Each response was phonetically transcribed by the experimenter and by a second judge using SAMPA (Speech Assessment Methods Phonetic Alphabet). An interrater agreement of 94.6% was found, based on the ratio of the number of identically transcribed phonemes on the total number of transcribed phonemes. As far as the scoring of the responses is concerned (see below), there was also a high agreement between the scores at the two linguistic levels under study, i.c., 99.4% at the nonword level and 92.8 % at the phoneme level. In case of disagreement, the experimenter and the second judge decided together how the responses were scored.

SCORING

In a first set of global analyses we wanted to answer the following questions: (1) Is there an effect of Group (individuals with versus without dyslexia) on the retention of a nonword? (2) Is there an effect of Group on the retention of phonemes' identity when repeating a nonword? (3) Is there an effect of Group on the retention of phonemes' serial order when repeating a nonword? To that end, we assigned a binary score to each response, at the three different levels of analysis, i.c., the nonword level and the two phoneme levels. Binary scores were used for two reasons. Firstly, in line with previous researchers who scored overall NRT performance (nonword level performance) with a binary coding system (correct response or not) (e.g., Martens & de Jong, 2006), we decided to use binary scores for the phoneme-related variables as well. This made for a consistent scoring procedure. Secondly, we wanted the phoneme level scores to reflect whether participants could repeat the nonword (a) *without any phoneme identity errors*, irrespective of serial order errors (score of 1) and (b) *without any serial order errors among the correctly recalled phonemes*, i.e., irrespective of phoneme identity errors (score of 1). Hence, the *nonword score* reflected whether both the identity and the serial order of all phonemes were correctly recalled, the *phoneme identity score* reflected whether the identity of all phonemes was correctly recalled and the *serial order score* reflected whether the serial order of all (correctly recalled) phonemes was correctly recalled. This coding system made it possible to distinguish between the two types of phoneme retention that are unavoidably conflated in the way NRT data are typically analyzed.

Note that it is important not to restrict the analyses at the phoneme level to the set of incorrect responses. As our questions with respect to the impact of reading skill on participants'

ability to retain phonemes' identity and serial order involves the scores at a lower phonological level than the nonword scores (i.e., the phoneme level), it is important to include the scores on phoneme identity and serial order for all responses, i.e., for both correct and incorrect nonword repetitions. Indeed, correct nonword repetitions are responses where no problems arise at the phoneme level, neither with respect to phoneme identity performance nor with respect to serial order performance, and, hence, provide equally relevant information as incorrect nonword repetitions, which may reflect a problem at the level of phonemes' identity, phonemes' serial order, or both. Leaving out correct nonword responses from the analyses at the phoneme level would be a decision that results in a considerable loss of data at the phoneme-level variables, and, hence in a reduction of the statistical power of our analyses. Thus considered, the phoneme-level analyses in a sense 'split up' the nonword analysis, which conflates performance at the levels of phoneme identity and serial order, into two separate analyses on the same set of responses. A more detailed description of our coding system is given below. The same scoring system was used in recent a study on the relationship between NRT performance and decoding skill in a group of 10-year-old children (Schraeyen et al., 2017).

Retention of the nonword (nonword level score). A response was counted as correct and given a binary score of 1 if the entire nonword was correctly reproduced, i.e., no errors were made at the word stress level, the syllable, and the phoneme levels. For example, for the target nonword /lem.rós.pag/, the response /lem.rós.pag/ would be scored as 1. Any other response was given a binary score of 0. Note that this first analysis is the analysis that is typically reported in NRT experiments. Although it generally reflects differences in the overall retention performance of individuals with dyslexia and controls (individuals with dyslexia performing worse), it cannot address the question whether differences in identity retention or in serial order retention are responsible for this difference. As described above, it is our explicit purpose to address this question. This is where the two retention scores at the phoneme level come in.

Retention of phonemes (phoneme level scores). To evaluate participants' ability to retain the identity and serial order of the phonemes in the nonword item, each response was assigned two binary values at the phoneme level. These values reflected performance on the two phoneme-related variables. All possible combinations of the binary values on these two variables resulted in four types of responses. *Response type 1*: a *correct response* received a binary score of 1 on both variables. In such a response, there were no problems with the retention of either identity or serial order. *Response type 2:* a response with *only phoneme identity errors* received a binary score of 0 on the dependent variable Phoneme Identity and a binary score of 1 on the dependent variable Serial Order. In such a response, there were only

problems with identity retention. *Response type 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. In such a response, there were only problems with serial order retention. *Response type 4*: a response with *both phoneme identity and serial order errors* was given a binary error score of 0 on both the Phoneme Identity and Serial Order variables. In such a response, there were problems with both identity and serial order retention. The different response types are described in more detail below.

(1) Correct response (no identity or serial order errors): all target phonemes appeared in the same serial order in the response, for example, target: /lem.rós.pag/, response: /lem.rós.pag/.

(2) Phoneme identity errors. Phoneme identity errors refer to errors in which a participant (a) added and/or (b) substituted and/or (c) omitted at least one phoneme in the response while respecting the correct serial order of all correctly recalled phonemes (see below). For instance, target: /lem.rós.pag/, response: /lem.rós.pags/ (addition of /s/); target: /lem.rós.pag/, response: /lem.vós.pag/ (substitution of /r/ by /v/); target: /lem.rós.pag/, response: /lem.ró.pag/ (omission of /s/); target: /lem.rós.pag/, response: /lem.kós.pa/ (substitution of /r/ by /k/ and omission of /g/).

(3) Serial order errors. Serial order errors refer to errors in which a participant shifted at least one of the *correctly recalled* phonemes. We did not score a phoneme's *absolute* serial order but its position *relative* to the other correctly recalled phonemes of the target word. Hence, phoneme substitutions, phoneme omissions, and phoneme additions did not necessarily have a negative effect on serial order scores. For example, the substitution error in the response /lem.pós.pag/ for the target /lem.rós.pag/ does not affect the serial position of the *correctly recalled* phonemes. Note that it is obvious, both from a logical point of view and from an analytical perspective, that the retention of a phoneme's serial order can only be determined if the phoneme itself has been correctly recalled, i.e., if the phoneme does not give rise to an identity error. First, it makes no sense to attribute positional properties to phonemes that are absent in the output (omissions) or did not appear in the input (additions, substitutions). Second, if each phoneme identity error also counted as a serial order error, the former errors would be a subset of the latter, which would be at odds with our intention to separate these two error types.

To determine whether the serial order of phonemes was respected, we used the procedure suggested by McKelvie (1987). In a first step, all phoneme strings of two or more phonemes were considered correct if they also appeared in the input nonword, whether or not

these were in the correct absolute position when counting from the beginning or the end of the nonword sequence. In a second step, we checked whether any remaining single phoneme was in the correct absolute position when counting either from the beginning or the end of the string, but not both. For example, for the target word /lem.rós.pag/ the response /rem.los.pag/ would be checked as follows: all nine phonemes are correctly identified (i.e., no phoneme identity errors). Following the McKelvie method, the phonemes in the strings /em/, /os/, /pag/ are correctly ordered. However, the single phoneme /l/ as well as the single phoneme /r/ are swapped between positions 1 and 4, violating the rule that single phonemes have to be in the correct absolute position when counting either from the beginning or the end of the string. For the target word /lem.rós.pag/ the response /ler.só.pa/ would be checked as follows: seven phonemes are correctly identified (i.e., omission of /m/ and /g/). Following the McKelvie method, the phonemes in the strings /le/, /pa/ and /o/ are correctly ordered. However, the single phonemes /r/ and /s/ are shifted, violating the rule that single phonemes have to be in the correct absolute position when counting either from the beginning or the end of the string.

(4) *Combined errors.* These errors refer to errors in which a response contained both phoneme identity and serial order errors. For example, target /lem.rós.pag/, response: /rem.ló.pag/. The target phonemes /r/ and /l/ are swapped, giving rise to a serial order error, whereas the /s/ is omitted, giving rise to a phoneme identity error.

Table 1 provides an overview of the four response types described above and their specific subtypes. In addition, their associated binary values on the two phoneme-related dependent variables are provided.

Table 1

Overview of Response Types and their Associated Binary Values with respect to Identity and Serial Order Performance

Response type	Identity Performance	Serial order Performance
Correct responses	1	1
Phoneme identity errors		
Phoneme addition	0	1
Phoneme omission	0	1
Phoneme substitution	0	1
Phoneme substitution and omission	0	1
Phoneme substitution and addition	0	1
Serial order errors		
Phoneme shift	1	0
Combined errors		
Phoneme shift and phoneme substitution	0	0
Phoneme shift and phoneme omission	0	0
Phoneme shift and phoneme substitution and phoneme omission	0	0

Note. Combined errors were incorporated in the statistical analyses of both phoneme identity performance and serial order performance (see text). Note that these errors (final three rows) could occur in combination with or without phoneme additions (not specified in the table to reduce the number of rows).

STATISTICAL ANALYSES

To start with, we computed Pearson correlation coefficients to explore the (linear) association between the different dependent variables in this study: nonword level performance, phoneme identity performance and serial order performance. We calculated total sum scores (i.e., total number of correct responses on each dependent variable) per participant.

Next, we performed additional analysis in which the effect of the independent variables was examined at the level of the individual item responses. To this end, we used General Linear Mixed effects Models (GLMMs) with a logit link function (because the outcomes are binary) (see for instance Molenberghs & Verbeke, 2005). These models make it possible to simultaneously assess the variance explained by the random effect of participants and the

random effect of items. The latter one is crucial to avoid the language-as-a-fixed-effect fallacy (Clark, 1973). In our initial models, we included Group (coded as group with dyslexia = 0, group without dyslexia = 1), Syllable Length (a factor with levels 3, 4 and 5; this variable was dummy coded with 2 dummies and syllable length = 3 as the reference category) and the Group by Syllable Length interaction as fixed effects. Additionally, Phoneme Identity performance was used as a covariate in the analysis of Serial Order performance (and vice versa) to bring the collinearity between both factors under statistical control, i.e., to make sure that the effect of Group was not due to the shared variance between the two phoneme-related variables but to the unique variance that it caused on the outcome variable (see Table 2). The use of such an approach is quite common (see Verbeke & Molenberghs, 2000). Subsequently, it was evaluated whether removing the interaction term led to a significant decrease of model fit. We used the Maximum Likelihood method to estimate all model parameters. An advantage of this is that likelihood-ratio tests can be carried out to compare the fit of the model with the interaction term and the model without this term. The final model was then refitted using the Restricted Maximum Likelihood (REML) estimation, as this yields more accurate estimates of the variance components in smaller samples. All models were fitted using R (packages lme4, Bates & Maechler, 2009; language R, Baayen, 2008). Throughout the paper, the significance level was set at $p \leq .05$.

RESULTS

Table 2 shows the mean percentages for all response types as a function of both Group and Syllable Length.

Table 2

Mean Percentages for each Response Type for Each Syllable Length in the Dyslexic and Non-Dyslexic Groups

	Dyslexic group			Non-dyslexic group		
	3-syllable items	4-syllable Items	5-syllable items	3-syllable items	4-syllable items	5-syllable Items
No phoneme error (%)	24.4	14.6	5.7	26.9	17.6	7.2
Phoneme identity error (%)	8.1	16.4	26.9	5.9	14.1	24.2
Serial order error (%)	2.2	9.4	17.5	.9	6.5	14.6
No response (%)	.2	.06	.25	.06	.4	1.1
Total (%)	34.9	40.5	50.4	33.8	38.6	47.1

Note. Combined errors are not analyzed as a separate error type. Since they count as both identity errors and serial order errors, they have also been included in the error rates for these two error types.

CORRELATIONAL ANALYSES

Table 3 reveals high positive correlations between performance at the levels of whole nonword retention, phoneme identity retention, and serial order retention (i.e., all rs $\geq .76$, all ps $< .0001$). This should not come as a surprise, as three of the four response types[6] imply identical binary values on the nonword level variable and one phoneme level variable. For the correlation involving performance on phoneme identity, these are correct responses, phoneme identity errors, and combined errors (representing 96.10% and 95.90% of all responses in the dyslexic and non-dyslexic groups, respectively). For the correlation involving the performance on phonemes' serial order, these are correct responses, serial order errors, and combined errors (representing 73.80% and 73.70% of all responses in the dyslexic and non-dyslexic groups, respectively). Hence, only the proportion of responses of the fourth response type (phoneme

[6] The 4 response types are: correct responses, phoneme identity errors, serial order errors and combined errors.

serial order error in the case of phoneme identity variable; phoneme identity error in the case of serial order variable), weaken the correlation. In the absence of these responses, there would be a perfect correlation between nonword performance and performance at the phoneme level variables. The high positive correlation between the two phoneme-related variables is not surprising either, as correct responses and combined errors imply identical scores on these two variables (representing 70.41% and 71.16% of all responses in the dyslexic and non-dyslexic groups, respectively). In contrast, the identity and serial order errors weaken the correlation, such that the distribution of the two categories of response types determines the exact value of the correlation.

In addition, we found a significant positive correlation between Group and both nonword level performance ($r = .22$, $p < .05$) and serial order performance ($r = .25$, $p < .05$), but not between Group and phoneme identity performance ($r = .20$, $p > .05$).

However, the correlations between Group and each of the three outcome variables do not allow us to infer that adults with dyslexia have significantly more problems with serial order than controls but perform equally well with respect to the retention of phoneme identity. In a more fine-grained analysis we will have to take the (strong) correlation between phoneme identity performance and serial order performance into account, i.e., partial out the effect of serial order (phoneme identity) when assessing the effect of Group on phoneme identity (serial order) to isolate unique variance on the dependent variable of interest (see also Staels & Van den Broeck, 2014).

Table 3

Pearson's Correlations 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	.98**	-	
3. Serial Order Performance	.80**	.76**	-
4. Group	.22*	.20	.25*

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

NONWORD LEVEL PERFORMANCE

The results of the GLMM that was used to analyze participants' nonword level performance are summarized in Table 4. The two-way interaction Group by Syllable Length was not significant. Hence, in the final model, the interaction term was removed.

Table 4

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Nonword Level Performance from Group and Syllable Length

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	1.32	.38	3.48	.0005
Non-Dyslexic Group	.48	.22	2.15	.03
Syllable Length 4	-1.64	.49	-3.37	.0008
Syllable Length 5	-3.65	.50	-7.33	< .0001

Note. Reference levels for the generalized linear mixed model are Syllable Length 3 and Dyslexic Group.

As expected, we observed a main effect of Group (p = .03), revealing that the probability of a correct nonword repetition was higher for participants without dyslexia compared to participants with dyslexia. In addition, for both participant groups (i.e., no significant interaction), the probability of a correct response was higher for three-syllable nonwords compared to four-syllable nonwords, p = .0008, and five-syllable nonwords, p < .0001, and higher for four-syllable nonwords compared to five-syllable nonwords, β = -2.10, p < .0001 (the latter β estimate was obtained by fitting the same model as shown in Table 4 but using Syllable Length 4 as the reference level; data not shown).

Figure 1 visualizes the probability of a correct response as a function of Syllable Length and Group. Probabilities are the values that are predicted by the statistical model, more particularly by back-transforming the beta value for each combination of Group and Syllable Length (a value on the logit scale) to the number whose natural logarithm is beta. This number expresses the ratio between the probability of success (p) over the probability of failure (1-p). Solving the equation for p yields the proportion of correct performance for that cell in the design. Across all syllable length conditions, adults with dyslexia repeated about 7% fewer nonwords correct than controls (44.66% vs. 51.71%). The performance of both groups linearly dropped between three-syllable items (77.12%), four-syllable items (48.21%), and five-syllable items (19.22%), each additional syllable resulting in a drop of around 30% correct nonword

repetitions (averaged across groups). For the sake of comparison, Figure 1 also visualizes the observed proportions (black lines).

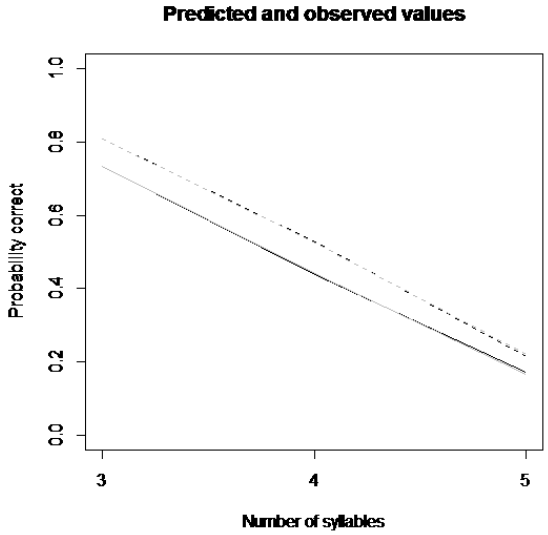

Figure 1. Predicted and Observed Probabilities of Correct Nonword Repetition as a Function of Group and Syllable Length.

Note. Predicted values (grey) and observed values (black) for individuals with dyslexia (solid line) and individuals without dyslexia (dashed line). Although the lines for the predicted and observed values virtually coincide, this does not mean that the model makes perfect predictions. The predicted probabilities in Figure 1 are averaged across many predicted individual responses, which may or may not differ from the actual responses. The model does a good job in predicting mean response behavior.

PHONEME LEVEL PERFORMANCE

We also used GLMMs to analyze the retention of (1) phonemes' identity, and (2) serial order. Recall that *all* responses were included in these two analyses. As explained, a participant's response was assigned two binary scores at the phoneme level, expressing whether all phonemes were recalled correctly or not (identity scores of 1 and 0, respectively), and whether all correctly recalled phonemes had been placed in the correct serial order or not (serial order scores of 1 and 0, respectively).

We found no significant Group by Syllable Length interaction effect in the analysis of phoneme identity performance. After removing the interaction term from the model, we obtained the results that are summarized in Table 5. There was a significant effect of serial order performance on phoneme identity performance ($p < .0001$), which is in line with the relatively strong correlation between these two variables (see Table 3). As explained above, serial order performance was included in the model to ensure that the effect of the other predictors accounted for unique variance on phoneme identity performance. The final GLMM revealed no main effect of Group ($p = .12$), i.e., participants with and without dyslexia did not differ in their retention of phonemes' identity. In addition, phoneme identity performance was better on in three-syllable nonwords compared to four-syllable nonwords ($p = .0004$) and five-syllable nonwords ($p < .0001$). Similarly, phoneme identity performance was better in four-syllable nonwords compared to five-syllable nonwords ($\beta = -1.8$, $p < .0001$) (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). The effect of Syllable Length was the same for both groups, as there was no Group by Syllable Length interaction.

Figure 2 visualizes the model-predicted probabilities of correct phoneme identity performance as a function of Syllable Length and Group. On average, performance on phoneme identity retention was about 6.5% worse in the dyslexic group compared to the control group (48.23% vs. 54.81%), a non-significant effect. Performance in both groups dropped about 25% between three-syllable items (78.81%) and four-syllable items (53.86%), and about 32% between four-syllable items and five-syllable items (21.89%). For the sake of comparison, Figure 2 also visualizes the observed proportions (black lines).

Table 5

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Phoneme Identity Performance from Group and Syllable Length when Statistically Controlling for Serial Order Performance

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	-.17	.33	-.52	.61
Non-Dyslexic Group	.32	.21	1.55	.12
Syllable Length 4	-1.12	.39	-2.90	.004
Syllable Length 5	-2.91	.40	-7.30	< .0001
Serial Order Performance	1.73	.14	12.28	< .0001

Note. Reference levels for the generalized linear mixed model are Syllable Length 3 and Dyslexic Group.

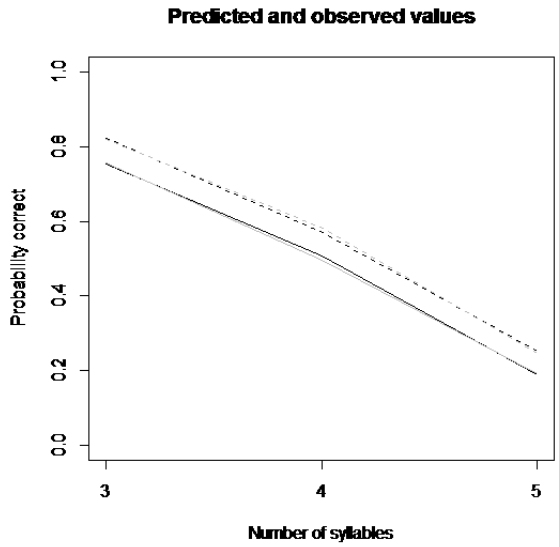

Figure 2. Predicted and Observed Probabilities of Correct Phoneme Identity Retention as a Function of Group and Syllable Length

Note. Predicted values (grey) and observed values (black) for individuals with dyslexia (solid line) and individuals without dyslexia (dashed line).

Table 6 summarizes the results of the final model for the analysis of serial order performance. We found no significant Group by Syllable Length interaction effect. After removing this interaction term from the final GLMM, the model revealed a significant effect of Phoneme Identity ($p < .0001$). Hence, phoneme identity performance was a predictor of serial order performance at the level of individual responses (see also Table 3), these responses being the input to the GLMM. We found a significant effect of Group ($p = .006$). Adults without dyslexia outperformed the adults with dyslexia in the retention of the serial order of correctly recalled phonemes. In addition, we observed a main effect of Syllable Length across both groups (no Group by Syllable Length interaction), indicating that serial order performance was better in three-syllable nonwords compared to four-syllable nonwords ($p < .0001$), and five-syllable nonwords ($p < .0001$). Serial order performance was also higher in four-syllable nonwords compared to five-syllable nonwords, $\beta = -2.7$, $p < .0001$ (this β estimate was obtained by fitting the same model as shown in Table 6 but using Syllable Length 4 as the reference level; data not shown).

Figure 3 visualizes the model-predicted probabilities of repeating the correctly recalled phonemes in their correct serial position, as a function of Syllable Length and Group. Note that the percentages are higher than in Figures 1 and 2 because serial order recall is a conditional probability, i.e., it expresses the probability that no serial order errors are made within the set of correctly recalled phonemes. On average, performance on serial order retention was about 7% worse in the dyslexic group compared to the control group (70.72% vs. 77.43%), a highly systematic effect (p = .006). The drop in performance was on average about 20% between three-syllable items (95.43%) and four-syllable items (75.95%), and about 25% between four-syllable items and five-syllable items (50.84%). For the sake of comparison, Figure 3 also visualizes the observed proportions (black lines).

Table 6

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Serial Order Performance from Group and Syllable Length when Statistically Controlling for Phoneme Identity Performance

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	2.31	.35	6.68	<.0001
Non-dyslexic Group	.44	.16	2.77	.006
Syllable Length 4	-1.79	.43	-4.21	<.0001
Syllable Length 5	-2.87	.43	-6.68	<.0001
Phoneme Identity Performance	1.76	.14	12.72	<.0001

Note. Reference levels for the mixed model are Syllable Length 3 and Dyslexic Group.

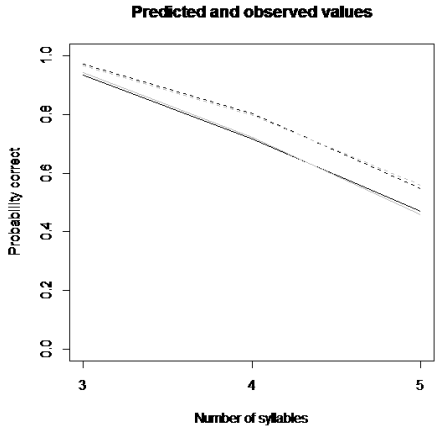

Figure 3. Predicted and Observed Probabilities of Correct Serial Order Retention as a Function of Group and Syllable Length

Note. Predicted values (grey) and observed values (black) for individuals with dyslexia (solid line) versus individuals without dyslexia (dashed line).

PAIRWISE COMPARISONS BETWEEN RESPONSE TYPES

In the mixed model analyses of phoneme identity performance and serial order performance reported above, we included both correct and incorrect responses. These analyses shed light on abilities individuals with and without dyslexia to recall (a) the identities and (b) the positions of the phonemes in a nonword. In these analyses, we statistically controlled for the correlation between the recall of phonemes' identity and their serial order. However, the fact that part of the variability on the dependent variable can be explained by both identity and serial order recall (the part that is kept under statistical control), reduces the amount of variability that can be uniquely explained by each independent variable.

In a different type of analysis, we did not keep the collinearity between the two types of recall under statistical control, but kept one of the two recall measures constant while varying the other. More particularly, we first studied the *effect of Group on the probability of making a serial order error* by focusing on the contrast between pure identity errors and a combination of identity and serial order errors (keeping the presence of identity errors constant). Secondly, we studied the effect of *Group on the probability of making a phoneme identity error* by focusing on the contrast between correct responses and pure identity errors (keeping the absence of serial order errors constant).

We did not use the data for the pure serial order errors because such errors were very rare, both in the non-dyslexic group (5.4 % of all errors) and in the dyslexic group (6.2 %), as can be seen in Table 7. That table gives the distribution of the three error types – (1) pure identity errors, (2) pure serial order errors, and (3) combined error – as a function of both Group and Syllable Length. Hence, when participants could reproduce all phonemes of the nonword, they almost invariably reproduced their serial order as well, i.e., made a correct response. Moreover, the probability of making a pure serial order error did not differ between individuals with and without dyslexia. A GLMM in which the effect of Error Type (pure serial error vs. other error) was predicted by Group yielded no effect of this fixed factor: $\beta = .07$ (SE $\beta = .27$), $p = .80$. Consequently, we removed this error type from further analyses.

Table 7

Mean Percentage Errors for each Error Type as a Function of Group and Syllable Length

	Dyslexic group			Non-dyslexic group		
	3-syllable items	4-syllable items	5-syllable items	3-syllable Items	4-syllable items	5-syllable items
Pure Phoneme identity error	11.8	16.9	18.0	11.4	18.9	22.4
Pure Serial order error	1.0	4.2	1.0	0.9	2.6	1.9
Combined error	3.0	13.1	31.0	1.2	11.4	29.3
Total		100			100	

Note. The sum of all percentages in each group represents 100% of all error responses in that group.

Table 8 summarizes the results of the GLMM predicting the probability of a serial order error when a phoneme identity error has been made (Group and Syllable length being the only two fixed factors). A pure identity error was coded as 0, whereas a combined error was coded as 1. The model revealed a main effect of Group, $p = .01$. When a phoneme identity error was made, the probability of making a serial order error as well (i.e., a combined order error) was smaller in participants without dyslexia compared to participants with dyslexia. In addition, we observed a main effect of Syllable Length indicating that the probability of a combined error increased with increasing syllable length.

Table 8

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Phoneme Serial Order Errors in the Set of Pure Identity Errors and Combined Errors from Group and Syllable Length

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	-2.27	.39	-5.84	< .0001
Non-dyslexic Group	-.43	.17	-2.55	.01
Syllable Length 4	1.86	.48	3.91	< .0001
Syllable Length 5	2.87	.47	6.06	< .0001

Note. Reference levels for the mixed model are Syllable Length 3 and Dyslexic Group.

Figure 4 visualizes the model-predicted probabilities of a failure in serial order retention when a phoneme identity error has been made. On average, this failure of serial order retention in the presence of a phoneme identity error was about 7% higher in the dyslexic group compared to the control group (38.98% vs. 31.82%). The increase of errors was on average about 26% between three-syllable items (11.60%) and four-syllable items (37.61%), and about 19% between four-syllable items and five-syllable items (56.99%). For the sake of comparison, Figure 4 also visualizes the observed proportions (black lines).

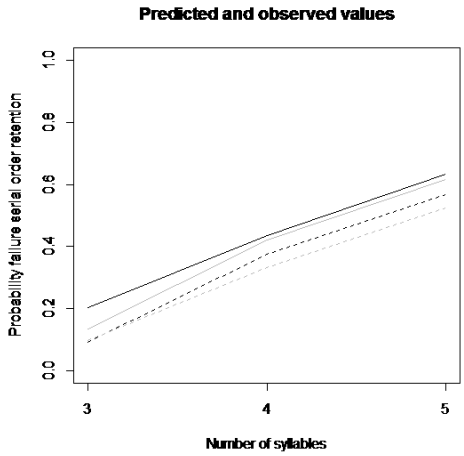

Figure 4. Probabilities of Serial Order Retention Failures when a Phoneme Identity Error has been made, as a Function of Group and Syllable Length.

Note. Predicted values (grey) and observed values (black) for individuals with dyslexia (solid line) and individuals without dyslexia (dashed line).

In a second analysis, we studied the effect of Group on the recall probability of phonemes' identity, more particularly, by contrasting the sets of correct responses and pure identity errors. No serial order errors were made in these two response types. A correct response was coded as 0, whereas a pure identity error was coded as 1. As can be seen in Table 9, the GLMM revealed no effect of Group, $p = .10$, i.e., when no serial order errors were made, both groups were equally likely to make a phoneme identity error. An identity error was higher for five-syllable nonwords compared to three-syllable nonwords ($p < .0001$) and four-syllable nonwords ($p = .0001$), and higher in four-syllable nonwords compared to three-syllable nonwords ($p = .004$).

Table 9

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Phoneme Identity Errors (in the Set of Pure Identity Errors and Correct Responses) from Group and Syllable Length

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	-1.58	.32	-4.91	< .0001
Non-dyslexic Group	-.34	.20	-1.66	.10
Syllable Length 4	1.20	.41	2.91	.004
Syllable Length 5	2.82	.43	6.57	< .0001

Note. Reference levels for the mixed model are Syllable Length 3 and Dyslexic Group.

Figure 5 visualizes the model-predicted probabilities of a failure in phoneme identity order retention in the absence of a serial order error. On average, this failure was about 5.5% worse in the dyslexic group compared to the control group (45.81% vs. 40.22%), a non-significant effect. The increase of errors was on average about 21,5% between three-syllable items (19.19%) and four-syllable items (40.65%) and about 29% between four-syllable items and five-syllable items (69.21%). For the sake of comparison, Figure 5 also visualizes the observed proportions (black lines).

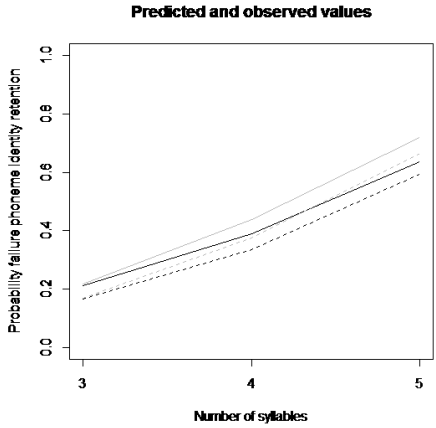

Figure 5. Probabilities of Phoneme Identity Retention Failures when no Serial Order Error has been made, as a Function of Group and Syllable Length.

Note. Predicted values (grey) and observed values (black) for individuals with dyslexia (solid line) and individuals without dyslexia (dashed line).

DISCUSSION

The central aim of this study was to further our understanding of the well-documented worse performance of individuals with dyslexia versus typical readers on the Nonword Repetition Task (NRT). We attempted to explain the cause of lower performance on the NRT in individuals with dyslexia by focusing on their retention of phonemes' identity and their retention of phonemes' serial order when repeating a nonword.

We conducted three types of analyses. First, we calculated global correlations between the main variables of interest, using the participant total scores on the NRT. Second, we performed generalized linear mixed effects model analyses at the level of individual (item) responses, in which the probabilities of phoneme identity performance and serial order performance were predicted by Group (dyslexia vs. no dyslexia), the items' Syllable Length, and the correlation between the two dependent measures (identity and serial order recall). This analysis included all responses, i.e. correct and incorrect ones, to maximize the power of the statistical analysis. The third type of analysis assessed the effect of Group on two response types only, i.e., the collinearity between identity and serial order recall was removed by keeping one response type constant. All analyses converged on the same conclusions.

We found a significant positive correlation between group and overall nonword level performance (see Table 3): the total number of correctly recalled nonwords was smaller for adults with dyslexia than for controls. We found the same results when analyzing the data at the level of individual responses, using a GLMM. The probability of a correct response was higher in adults without dyslexia compared to adults with dyslexia. These results corroborate previous findings that reading difficulties in adults are indexed by difficulties in accurate nonword repetition (e.g., Dietrich & Brady, 2001).

Our next aim was to detect the cause of this overall problem with the repetition of nonwords. There are only two possible causes: a problem in recalling a nonword's phonemes or a problem in recalling the serial order of correctly recalled phonemes. The correlation analyses based on participants' overall scores on these measures revealed a significant correlation between group and phoneme serial order performance, but no correlation between group and phoneme identity performance (see Table 3). This rough exploratory analysis suggests that the ability to retain phonemes' serial order is dependent on participant's reader group (individuals with dyslexia performing worse than individuals without dyslexia) whereas the ability to retain phonemes' identity is not. However, the observation of a strong positive correlation between participants' retention of phonemes' identity and phonemes' serial order (see Table 3) urges caution. A correct interpretation of the data is only possible when the collinearity between these two measures is brought under statistical control.

To this end, in our GLMM analyses, one type of phoneme performance (phoneme identity or serial order) was included as a covariate when predicting performance on the other phoneme-related variable (see also Staels & Van den Broeck, 2014, for an analysis that is based on the same rationale). When thus analyzing phoneme identity performance, no effect of Group was found: individuals with dyslexia were as successful as individuals without dyslexia in retaining phonemes' identity when repeating a nonword. In contrast, an effect of Group was found when predicting serial order performance: adults without dyslexia outperformed adults with dyslexia with respect to the retention of phonemes' serial order, irrespective of the number of syllables in the nonword, i.e., the interaction between group and syllable length was not significant.

Separate analyses, in which the correlation between phoneme identity and phoneme serial order was not brought under statistical control but removed from the data by keeping responses on one measure constant (e.g., all responses contained phoneme identity errors), confirmed these findings. An effect of Group was found when the data on pure identity errors and combined errors were compared, i.e., the probability of making a serial error as well was

larger for adults with dyslexia than adults without dyslexia. In contrast, no effect of Group was found when the data on correct responses and pure identity errors were compared, i.e., the probability of making an identity error did not differ between the two groups. The latter analysis shows that the effect of group on serial order performance in the GLMM on all data was not due to a reduction in the variability of the data after statistical controlling for the correlation between the two phoneme-related outcome variables.

Overall, our results clearly show that (a) a failure of the retention mechanism for serial order (i.e., serial order performance), defined as a failure to reproduce the serial order of all correctly recalled phonemes of the nonword, occurred more often in adults with dyslexia than in adults without dyslexia, and that (b) a failure of the retention mechanism for phoneme identity, defined as a failure to reproduce all phonemes of the nonword, occurred equally often in both groups. These results are in line with a recent study by Schraeyen et al. (2017) in young readers with different literacy skills as well as other recent studies that focused on the recall of serial order in individuals with dyslexia (for a review, see Majerus & Cowan, 2016). Our results also confirm our hypothesis that serial order retention is a crucial factor for describing the widely reported differences in overall NRT performance between adults with and without dyslexia.

Even though the statistical analyses suggest that adults with dyslexia perform equally well as typical readers on the retention of phonemes' identity, one should be cautious in interpreting this result. The very high correlation between participants' nonword repetition performance and their performance on phoneme identity recall makes one expect the same effect of Group on both measures. Still, the control group outperformed the dyslexia group on the nonword measure but not on the phoneme identity measure. This failure to obtain an effect of Group on phoneme identity recall may be due to a lack of statistical power in the analyses. First, in the analysis of all responses (correct and incorrect ones), the collinearity between the recall of phonemes' identity and their serial position was brought under statistical control. This statistical procedure comes with a cost: it considerably reduces the amount of variability on the dependent measure that can be uniquely accounted for by the distinction between individuals with dyslexia and controls. Second, when analyzing the effect of Group on the difference between pure identity errors and combined errors (i.e., differing on the serial order measure only) the number of responses in the dataset is considerably smaller than in the entire dataset, i.e., the basis for the effect of Group on the probability of a correct nonword response. As both analyses have less power than the analysis of the effect of Group on nonword recall, our results do not allow us to doubt the possibility that individuals with dyslexia have more problems than

controls when having to recall the phonemes in a nonword. We emphasize that this does not undermine our conclusions. Indeed, despite this loss of statistical power, we still found a solid effect of Group on serial order recall in the two analyses mentioned above. This highlights problems with serial order recall in individuals with dyslexia even more. At the same time, this loss of power and the high correlation mentioned above imply that our null effect of Group on the recall of phoneme identity is still compatible with the established idea that individuals with dyslexia have problems at the level of phonemic representations. In what follows we will focus on our finding that individuals with dyslexia and controls strongly differ on their ability to recall phonemes' serial order. This finding corroborates the outcome of several recent studies that individuals with dyslexia experience problems in recalling serial order information.

The question now is how we can interpret the difference between the two groups with respect to serial order performance when performing a NRT task. We suggest four possible accounts. The first three accounts situate the effect at the level of STM storage. Two of these storage accounts attribute the effect to the nature of a linguistic representation in STM, whereas the third one links it to a non-linguistic, i.e., domain-general, STM ordering mechanism. The fourth account situates the effect at the level of retrieval from STM.

SERIAL ORDER PROBLEMS IN A PHONOLOGICAL REPRESENTATION IN STM

A first account explains problems in individuals with dyslexia as the result of their poorer knowledge of phonotactic regularities in the language. Phonotactics pertains to the permissible orderings of phonemes within syllables and words. For instance, in English as well as in Dutch, no word begins with the sequence /tl/, though the same sequence is allowed word-internally, as in *atlas* (Geudens & Sandra, 2003). Phonotactic knowledge also involves the implicit 'awareness' of the transitional probability between successive phonemes. This account can be directly linked to the development and structure of the phonological representations of words, as serial order is obviously an important property of these representations. Even though many researchers do not take this factor into account within the domain of phonology, it does play a crucial role in the development of phonological representations of words, as supported by the largely documented role of phonotactics in language development studies (Jusczyk, Luce, & Charles-Luce, 1994; Storkel, 2001). Note that the idea that individuals with dyslexia might have poorer phonotactic knowledge can easily be linked to the concept that they have more problems in representing accurate phoneme representations, the building blocks of phonotactic sequences.

Bonte, Poelmans, and Blomert (2007) used an oddball paradigm to investigate phonotactic sensitivity in children with dyslexia by analyzing ERP responses to phonologically legal nonwords with either high (e.g., /not.sel/) or low (e.g., /not.kel/) phonotactic probability. The mismatch negativity, i.e., the difference between the ERP activity induced by the same nonword when presented as the standard (the frequently repeated item of the pair in a block) and when presented as the deviation (the infrequently repeated item), was larger for the nonword with large phonotactic probability – but only for the control group. The individuals with dyslexia even showed the reversed tendency. The authors attributed this finding to a difference in sensitivity to phonotactic probabilities, suggesting that children with dyslexia are less sensitive to quantitative changes in phonotactic probability. As phonotactics deals with serial order restrictions on phonemes and phonotactic probability with differences in the transitional probabilities between successive phonemes, evidence suggesting inferior phonotactic sensitivity in poorer vs. better readers may be related to their different abilities in coding, retaining, or accessing the correct serial order of phonemes. Hence, these results seem to support our finding that individuals with dyslexia performed worse on the recall of phonemes' serial position than controls in a NRT.

Note, however, that Szenkovits, Darma, Darcy, and Ramus (2013) could not replicate a phonotactic sensitivity effect when studying adults with dyslexia. The authors suggested that "it could be perfectly possible that children with dyslexia, owing to their phonological deficit, incur a delay in their tuning to phonotactic probabilities, but that in adulthood they eventually reach normal phonotactic sensitivity, albeit later than controls" (p. 11). Another possibility is that the source of the difference between Bonte et al.'s and Szenkovits et al.'s studies is due to the magnitude of the phonotactic variations that are used in the two studies. Szenkovits et al. suggested that poorer sensitivity may be so subtle in individuals with dyslexia that it is difficult to measure. This would explain why this lower sensitivity surfaced in, for instance, Bonte et al.'s study, where larger phonotactic variations were used, but not in their own.

Obviously, our findings cannot shed light on this discussion. The only case we want to make is that there is evidence to suggest that poorer performance in individuals with dyslexia in recalling phonemes' serial order may be an intrinsic phonological effect. If so, it is plausible to trace it back to the vital role of phonotactics in phonological development. Findings like the ones in the oddball paradigm suggest that individuals with dyslexia have a poorer phonotactic knowledge, due to their weaker phonemic skills. This makes sense, as a good knowledge of phonemes should make it easier to detect (high-probability) phonotactic sequences. In turn, such sequences are likely to increase the size of phonological encoding units, making it possible

to use both phonotactic chunks and individual phonemes for the purposes of encoding and representation in STM. It follows that individuals with dyslexia would have fewer possibilities than controls for encoding, retaining, or accessing a sequence of phonemes in a NRT, which would explain their poorer recall of phonemes' serial order.

In short, the effect of Group on the retention of phoneme order may be explained in terms of higher-order phonological units, i.e., phonotactic sequences that are retrieved from LTM for encoding and storage in STM. In this account, there is no need for a special STM mechanism for ordering individual phonemes because subsequences of the nonwords match LTM representations with an internal order.

One observation in our study is compatible with (at least a partial) role of phonotactic knowledge: the extremely low proportion of pure serial order errors. This indicates that, when participants can recall all phonemes, they almost invariably also recall their serial order. At first sight, the fact that all phonemes' serial order is easier to recall when more phonemes can be recalled seems paradoxical. However, it makes perfect sense if the successful recall of these phonemes depends on the ability to encode the nonword in terms of phonotactic units. Such units are familiar phoneme sequences, characterized by high-probability transitional probabilities, such that positional information needs to be encoded, stored, and retrieved for fewer information units, the phonotactic chunks. Note that this account is not discredited by the null effect for the retention of phoneme identity (see power issue above).

SERIAL ORDER PROBLEMS IN AN ORTHOGRAPHIC REPRESENTATION IN STM

A second possibility is that the effect is non-phonological in nature but originates at the level of orthographic or graphemic knowledge. Previous work indeed shows that performance on the NRT is affected by orthographic factors (Nation & Hulme, 2011). Acting as a graphic support, the possible use of orthographic representations in the NRT could facilitate participants' memory for short-lived phonemic segments (Geudens, 2006). Within such a framework, the question remains how orthographic knowledge could explain our observation that the retention of serial order turns out to be a discriminating factor between the two groups. If anything, it suggests that persons with dyslexia have problems keeping a generated sequence of orthographic segments in their correct order. Note that this does not by itself imply that individuals with dyslexia can map phonemes onto graphemes equally fast as individuals without dyslexia. This would be an unlikely consequence in the light of the many demonstrations that individuals with dyslexia have problems with such mappings. Moreover,

recall that our measure of serial order recall is a conditional one: it expresses participants' success in serial order recall, given their set of correctly recalled phonemes, i.e., the number of correctly recalled phonemes does not play a role in the metric. Yet the data do suggest an equally fast mapping process of a phonemic representation onto an orthographic representation of the auditory nonwords in both groups. However, as mentioned above, we should be careful with this null effect. Our failure to obtain an effect of Group on the recall of phoneme identity may be due to low statistical power. As a matter of fact, this is quite likely in the light of the high correlation between nonword recall and phoneme identity recall (see above). Following the above line of reasoning, our findings are compatible with the idea that individuals with dyslexia are worse than controls in recalling the order of the letters in the orthographic representation of the auditorily presented nonword.

SERIAL ORDER PROBLEMS IN A DOMAIN-GENERAL SEQUENCING MECHANISM IN STM

A third account at the level of STM storage proposes the existence of a mechanism that is specifically dedicated to the retention of (any kind of) serial order information in STM. In previous work, such a mechanism has been suggested in the context of storing sublexical phonological representations (more particularly, syllables) in verbal STM (VTSM) (Gupta, 2003, 2009; Gupta et al., 2005). In these studies, a computational model for verbal STM has been proposed in which *sequence memory*, i.e., a language-independent memory component for keeping track of sequential information, is thought to be involved in temporarily maintaining the serial order of (sublexical) units. Participants found it easier to recall initial and final syllables in nonwords, suggesting the importance of serial order in verbal STM.

As highlighted in the Introduction, a few other recent studies elaborated on this view by making a distinction between item and serial order STM and exploring the relationship between a basic *sequence memory* mechanism, reading acquisition and reading impairment in children and adults (e.g., Martinez Perez et al.,2012a;2012b; Hachmann et al., 2014; Bogaerts et al., 2015). These studies provided evidence for a serial order problem in individuals with dyslexia. Importantly, these studies used both verbal and non-verbal stimuli, suggesting that the ordering mechanism is a domain-general one and is, hence, not restricted to verbal STM but is used for each type of information that must be stored in STM.

The distinction between *item* STM and *order* STM in all aforementioned studies is to some extent analogous to our distinction between the short-term retention of phoneme *identity*

and the short-term retention of *position* in the NRT. Hence, it is tempting to interpret our results in terms of these explanatory concepts and attribute problems with the retention of phonemes' serial order in individuals with dyslexia to a more fundamental problem with a basic domain-general ordering mechanism in STM, specifically dedicated to retaining the order of (any kind of) information units.

Yet, there is a major difference between our study and previous work. We captured the distinction between identity and order information by using the same (nonword) items and focusing at the intra-syllabic *phoneme* or *segmental* level (the crucial level according to the phonological deficit hypothesis), whereas all previous studies used two types of stimulus materials for studying the distinction between the recall of these two types of information. One might argue that the use of a single set of materials targeting the recall of phonemes involves a risk if one wishes to demonstrate the existence of two different STM processes (e.g., Majerus et al., 2012b; Hachmann et al., 2014; but see Staels & Van den Broeck, 2014). It is certainly true that, in our data, the correlation between identity recall and serial order recall indicates that the two mechanisms that are responsible for encoding, storing, and/or retrieving both types of information were not independent from each other. At the same time, it should be stressed that we found worse performance in serial order recall in individuals with dyslexia in two analyses where this correlation could not affect this outcome.

Moreover, this methodological choice might also be perceived as a strength. The fact that phonemes can be forgotten without losing the serial information for the correctly recalled phonemes and that phonemes can be recalled while losing their serial order makes it possible that two STM mechanisms, i.e., item STM and order STM, can be studied independently with a single set of materials. The fact that we could analyze our data without any contamination by the correlation between item and order recall demonstrates that our study did benefit from the strength offered by a single set of materials. In short, the effect of Group on the retention of phonemes' serial order is at least compatible with the concept of a domain-general, i.e., non-language-specific, ordering mechanism for information in STM.

SERIAL ORDER PROBLEMS DURING RETRIEVAL FROM STM

The three accounts above all concern the encoding or storage (representation) of a nonword's representation. The observed problem with the serial order recall of phonemes in individuals with dyslexia might reflect (a) a problem in encoding/representing the correct order of the item's phonemes, either in terms of of phonotactic chunks or in terms of graphemes (i.e., a language-

specific problem), or (b) a problem with a domain-general sequence mechanism that associates positional information with each phoneme in the nonword (i.e., a non-language-specific problem). One might propose a fourth account, in which the *nature* of the representations itself is not the key problem but rather the way in which these representations are *accessed* or retrieved. Such a view would mean that individuals with dyslexia have difficulties with the order in which they have to *retrieve* the specific information from the nonword's representation, not with the temporary storage of the nonword itself. This would imply that responses in the NRT do not reflect the underlying representation of a nonword but the process of *reconstructing* this representation by serially accessing its phonemes (or graphemic counterparts) or phonotactic units in a temporary STM representation. Hence, they would reveal the operation of a serial order mechanism at the moment of retrieval, rather than during encoding and, hence, at the level of the STM representation itself. In line with this account, Ramus and Szenkovits (2008) presented a detailed analysis of performance in adults with dyslexia on a range of phonological tasks, and concluded that their problems were not related to their phonological representations as such but strongly depended on task requirements, such as the processing load on STM, conscious awareness, or time constraints. Hence, they proposed that the basic problem of poor readers is a problem in *accessing* phoneme-based representations, not a problem involving the quality or intrinsically 'poor' nature of these representations (see also Van den Broeck & Geudens, 2012). Recent neuroimaging findings support this hypothesis. Boets et al. (2013) demonstrated that the neural quality of the phonetic representations in adults with dyslexia was intact but that they were hampered by a dysfunctional connection between frontal and temporal language areas. This suggests that problems in the fluent manipulation of speech sounds in individuals with dyslexia stem from their problems in accessing otherwise intact phonemic representations.

Future studies are required to shed further light on this matter and elucidate whether the serial phoneme order problem in individuals with dyslexia reflects a problem with the intrinsic quality of phonemic representations in STM, which do not adequately represent the correct serial order of the phonemic segments, or a problem with the *access* to the phonemic information in STM in the correct serial order, i.e., the composition or reconstruction of the phonological representation at the moment of retrieval. If the former possibility is correct, the problem might be rooted in a language-independent sequence memory problem in STM or be due to those individuals being less competent in the use of phonotactic and/or orthographic knowledge for retaining unfamiliar phoneme strings. If the second possibility is correct, the problem also involves the malfunctioning of a domain-general serial ordering mechanism but

one that does not regulate the retention process in STM itself but the retrieval of information from STM. Note that different skills in the use of phonotactics and/or orthography might also play a role if the problem is situated at the level of retrieval.

Even though our data are compatible with different accounts, the major demonstration of our study is that individuals with dyslexia have a problem in recalling the serial order of phonemes in a novel phoneme string, i.e., a nonword. We could demonstrate this with a single set of nonwords, even though there was a correlation between the two dependent measures, i.e., the recall of phonemes' identity and the recall of their serial order. The problem with serial order recall in individuals with dyslexia is so big that it is stronger than their problem with the recall of phonemes' identity. The latter problem was even not observed, even though the phonological deficit hypothesis made us expect it. However, we emphasize that the above-mentioned correlation and the consequent loss of power in the statistical analyses may be responsible for this, i.e., may have caused a Type II error. Hence, the absence of an effect of Group on identity recall cannot be construed as counter-evidence for the phonological deficit hypothesis. A second contribution of our study is that it is, to our knowledge, the first attempt to link reading success to two current views on the cause of dyslexia: the importance of the quality of phonemic representations in STM and the role of serial order in STM.

CONCLUSION

Our results show that the distinction between two types of retention errors – *phoneme identity* retention errors and *serial order* retention errors – contribute to a more precise understanding of NRT performance and the importance of the two task components in determining reading success. *Serial order* turns out to be a key factor in explaining the difference in the NRT error patterns of readers with and without dyslexia. Interestingly, Schraeyen et al. (2017), who also used phonemes' identity and serial position recall scores on a single set of items in a recent NRT study, arrived at the same conclusions for a group of young readers varying in reading success (regression design). We conclude that adult readers with dyslexia are less sensitive to phonemes' serial order compared to typical readers and that this finding imposes restrictions on any account of dyslexia.

REFERENCE LIST

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

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

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

Bogaerts, L., Szmalec, A., Hachmann, W.M., Page, M.P.A., Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities, 43–44,* 106–122. doi:10.1016/j.ridd.2015.06.01

Bonte, M.L., Poelmans, H., & Blomert, L. (2007). Deviant neurophysiological responses to phonological regularities in speech in dyslexic children. Neuropsychologia, 45, 1427-1437. doi: 10.1016/j.neuropsychologia.2006.11.009

Booij, G. (1995). *The Phonology of Dutch*. Oxford: Oxford University Press.

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

Castles, A., & Coltheart, M. (2004). Is There a Causal Link from Phonological Awareness to Success in Learning to Read? *Cognition, 91,* 77-111. doi: 10.1016/S0010-0277(03)00164-1

Clark, H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335-359. doi: 10.1016/S0022-5371(73)80014-3

Dietrich, J. A., & Brady, S. A. (2001). Phonological representations of adult poor readers: An investigation of specificity and stability. *Applied Psycholinguistics, 22,* 383-418.

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, word learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*, 375-384. doi: 10.1111/j.1467-9450.2005.00468

Foy, J. G., & Mann, V. (2001). Does strength of phonological representations predict phonological awareness? *Applied Psycholinguistics*, *22*, 301-325. doi:10.1017/S0142716401003022

Hachmann, W.M., 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

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

Geudens, A. (2006). Phonological awareness and learning to read a first language: Controversies and new perspectives. *LOT Occasional Series, 6,* 25–43.

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

Gupta, P. (2009). A computational model of nonword repetition, immediate serial recall, and nonword learning. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 108–135). Hove, England: Psychology Press.

Jusczyk, P. W., 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

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, article 1522. doi: 10.3389/fpsyg.2016.01522.

Martens, V.E., & de Jong, P.F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language, 98*, 140-149.

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

McDougall, S., & Hulme, C. (1994). Short-term memory, speech rate and phonological awareness as predictors of learning to read. In: Hulme, C., Snowling, M. (Eds.), *Reading development and dyslexia* (pp. 31–44). Chichester: Wiley.

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

Molenberghs G. and Verbeke G. (2005). *Models for discrete longitudinal data.* New-York: Springer-Verlag.

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.1467-7687.2010.01008

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*, 212-218. doi: 10.1016/S0959-4388(03)00035-7

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

Scheltinga, F. (2003). *The Dutch Nonword Repetition Test.* University of Amsterdam: Unplublished manuscript.

Schraeyen, K., Geudens, A., Ghesquire, P., Van der Elst, W., & Sandra, D. (2017). Poor performance on the retention of phonemes' serial order in short-term memory reflects young children's poor reading skills. *The Mental Lexicon, 12 (1),* 129-158.

Snowling, M.J. (2000). *Dyslexia* (2nd ed.). Oxford: Blackwell.

Sprenger-Charolles, L., Colé, P., & Serniclaes, W. (2006). *Reading acquisition and developmental dyslexia.* New York: Psychology Press.

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

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)

Szenkovits, G., Darma, Q., Darcy, I., & Ramus, F. (2013). *Exploring dyslexics' phonological deficit II: phonological grammar.* Manuscript submitted for publication. Retrieved from www.lscp.net/persons/ramus /docs/phonogrammar08

Szmalec, A., Loncke, M., Page, M., & Duyck,W. (2011). Order or dis-order? Impaired Hebb learning in dyslexia. *Journal of Experimental Psychology Learning Memory and Cognition, 37,*1270 – 1279. doi: 10.1037/a0023820

Torgesen, J.K. (1985). Memory processes in reading disabled children. *Journal of Learning Disabilities*, *18*, 350–357. doi: 10.1177/002221948501800608

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

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychiatry*, *45*, 2-40. doi: 10.1046/j.0021-9630.2003.00305

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

CHAPTER 3

POOR PERFORMANCE ON THE RETENTION OF PHONEMES' SERIAL ORDER IN SHORT-TERM MEMORY REFLECTS YOUNG CHILDREN'S POOR READING SKILLS

The content of this chapter has published as: Schraeyen, K., Geudens, A., Ghesquière, P., Van der Elst, W., & Sandra, D. (2017). Poor performance on the retention of phonemes' serial order in short-term memory reflects young children's poor reading skills. The Mental Lexicon, 12 (1), 129-158.

ABSTRACT

The present study aimed to identify crucial factors that underlie phonological representations in 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.

INTRODUCTION

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 between *rapid automatized naming* (RAN), i.e., the capacity to retrieve phonological codes for digits, objects etc., stored in 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* (verbal STM) tasks has also been shown, better reading skills being associated with superior verbal STM. This is, for instance, the case in memory span tasks such as a nonword repetition task (NRT), in which nonwords have to be recalled immediately from verbal STM (e.g., Beneventi, Tonessen, & Ersland, 2009). verbal STM 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.

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 readers and children with dyslexia (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 in- formation 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 dis- similar/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 reten- tion 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 inde- pendent variance in decoding abilities for reading (Martinez Perez et al., 2012a). In a second study, in which they used the same tasks to compare children with dyslexia 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 children with dyslexia 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. Individuals with dyslexia 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 ad- justments, arguing that 'a specific problem in serial order retention can only be proven by directly comparing individuals with and without dyslexia 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 children with dyslexia 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[7]. 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

[7] 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.

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 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, individuals with dyslexia 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 individuals with and without dyslexia. If so, the problem with the retention of phonemes' serial order in individuals with dyslexia 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., persons with dyslexia. 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 ex- posed to reading instruction from first grade in primary school onwards, i.e., they had received 20 to 21 months of explicit reading instruction[8]. We did not make a distinction between two reader groups, i.e., children with poor readings skills or dyslexia 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 having dyslexia, 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 children without dyslexia but severely impaired in those with dyslexia.

To determine children's literacy level, we administered a standardized oral word reading

[8] 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.

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 perfor- mance 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, i.e., a z-score, as our measure for Literacy.

MATERIALS AND DESIGN

We used the Dutch adaptation (Scheltinga, 2003) of the Children's Test of Nonword Repetition (Gathercole et al., 1994). More specifically, we used 36 nonwords (see Appendix A) of the original 48 nonwords as we did not include two-syllable items. Items in this NRT test were manipulated for three phonological factors: (a) syllable length (3-4-5, 12 items each), (b) nonword complexity (18 nonwords had the same consonants at each side of the syllable boundary, e.g., /*pip.pok.ket*/, whereas the other 18 nonwords had different consonants at each side of the syllable boundary, e.g., /*lem.ros.pag*/), and (c) phonological similarity (nonwords whose syllables contained two phonologically distinct versus phonologically similar consonants, e.g., /nim.mun.naf/ vs. /gik.kal.lom/) (see Appendix A). All items were phonologically legal in Dutch. All nonwords observed the Dutch pattern of placing stress on the penultimate syllable. Note that the latter two manipulations were not related to issues that were of theoretical importance for the current experiment. The only manipulation that was relevant, and will be included in the analyses, was the items' length in number of syllables, as this variable pertained to a phonological level in-between the level of the entire nonword and its constituent phonemes, the two levels that we focused on.

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 immedi- ately. 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.

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);

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 omit- ted 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.

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.

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 sub- types, 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. 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 were 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) were included. Each analysis addressed the question whether participants' literacy level determined 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 enabled 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 com- puted 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,..., N and j = 1,...,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,β,φ) = exp *f* ()ı{φ^(−1) [y_ij θ_ij-ψ(θ_ij)] + c(y_ij,φ)}, with η(μ_ij) = x_ij^' β+z_ij^' b_i for a known link function η(.), 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 ob- tained 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 sig- nificance.

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
5. Nonword Level Performance	-		
6. Phoneme Identity Performance	.93**	-	
7. Serial Order Performance	.77**	.69**	-
8. Literacy	.22*	.14	.23*

*p < .05, **p < .0001

This should not come as a surprise, as three of the four response types[9] imply identical binary values on the nonword level variable and one phoneme level variable. For the correlation

[9] The 4 response types are: correct responses, phoneme identity errors, serial order errors and combined errors.

involving performance on phoneme identity, these are correct responses, phoneme identity errors, and combined errors (representing 95.50% of all responses). For the correlation involving the performance on phonemes' serial order, these are correct responses, serial order errors, and combined errors (representing 67.40% of all responses). Hence, only the proportion of responses of the fourth response type (phoneme serial order error in the case of phoneme identity variable; phoneme identity error in the case of serial order variable), weaken the correlation. In the absence of these responses, there would be a perfect correlation between nonword performance and performance at the phoneme level variables.

The high positive correlation between the two phoneme-related variables is not surprising either, as correct responses and combined errors imply identical scores on these two variables (representing 63% of all responses). In contrast, the identity and serial order errors weaken the correlation, such that the distribution of the two categories of response types determines the exact value of the correlation.

The correlations between literacy and the other three outcomes were substantially lower and can obviously not be predicted from the scores at the nonword level. 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. However, firm conclusions concerning the association between Group and each of the two phoneme-related variables can only be drawn from an analysis in which the correlation between these two variables is kept under statistical control. This will be done in the next paragraph.

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 (β = −1.47 for Syllable Length 5, p = .007) (the latter β estimate is ob- tained 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.

The average increase in successful nonword repetition per z-score unit on the literacy measure was 3.24% for three-syllabe items, 2.05% for four-syllable items, and 0.65% for five-syllable units[10]. Successful performance was 45% worse on four-syllable items than on three-syllable items, and 13% worse on five-syllable items than on four-syllable items.

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

[10] Even though the effect of Literacy is constant across the three nonword lengths in the *logit* estimates of the model, this is not reflected in the *percentages.* As the proportion correct performance is derived by the formula p = exp(x)/(1+exp(x)), where x is the estimated logit effect of a condition on the basis of the linear combination of the relevant betas, the constant effect in one component of x cannot be reflected in the output of this formula.

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

Note. 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 com- pared 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[11] (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 Discussion).

In the model-predicted scores for items where no serial order errors were made, the effect of Literacy on successful recall of all phonemes from a nonword was on average 1.70% per z-score unit on the Literacy variable for three-syllable nonwords, 1.73% for four-syllable nonwords, and 0.79% for five-syllable nonwords. For items on which serial order errors were made, these average effects were 1.64%, 1.78%, and 0.84% for three-syllable, four-syllable, and five-syllable items, respectively. The statistical analysis indicated that Literacy did not have a significant effect. Successful performance was 37.33% worse on four-syllable items than on three-syllable items, and 21.20% worse on five-syllable items than on four-syllable items.

[11] 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.

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.

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 or- der 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 'un- contaminated' 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 Discussion).

In the model-predicted scores for items where no phoneme-identity errors were made, the effect of Literacy on successful recall of the serial order of the correctly recalled phonemes was on average 0.54% per z-score unit on the Literacy variable for three-syllable nonwords, 3.22% for four-syllable nonwords, and 4.79% for five-syllable nonwords. For items on which serial order errors were made, these average effects were 0.20%, 1.52%, and 3.07% for three-syllable, four-syllable, and five-syllable items, respectively. Successful performance was 11.91% worse on four-syllable items than on three-syllable items, and 14.57% worse on five-syllable items than on four-syllable items.

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.

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 readers and individuals with dyslexia 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 or- der 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 associat- ed 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 individuals with dyslexia 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 adults with dyslexia suggests that serial order problems are not exclusively associated with individuals with dyslexia 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 individuals with dyslexia. 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 children with dyslexia 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 those with dyslexia 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 individuals with dyslexia (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 STM. 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 individuals with dyslexia 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 persons with dyslexia (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.

REFERENCE LIST

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

Binamé, 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

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 timinganalysis. *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

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

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.), *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' sensitiv- ity 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

Scheltinga, F. (2003). *The Dutch Nonword Repetition Test.* University of Amsterdam: Unpublished manuscript.

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

CHAPTER 4

BEYOND GLOBAL DIFFERENCES BETWEEN MONOLINGUAL AND BILINGUAL CHILDREN ON THE NONWORD REPETITION TASK: RETENTION SKILLS FOR PHONEMES' IDENTITY AND SERIAL ORDER

The content of this chapter has published as: Schraeyen, K., Van der Elst, W., Geudens, A., Ghesquière, P., & Sandra, D. (2017). Beyond global differences between monolingual and bilingual children on the nonword repetition task: retention skills for phonemes' identity and serial order. Bilingualism: Language and Cognition, 1-16.

ABSTRACT

This study compared NRT performance in monolingual Dutch and sequential bilingual Turkish-Dutch third-graders using a Dutch Nonword Repetition Task (NRT). Several novel response analyses at the phoneme level were applied to further understand the earlier reported overall accuracy differences in NRT performance between bilinguals and monolinguals. Analyses in which the retention of phonemes and the retention of their serial order were disentangled revealed that monolinguals outperform bilinguals with respect to the retention of the phonemes themselves. However, both groups did not differ in their retention of the serial order of correctly recalled phonemes. Furthermore, this study confirms that expressive vocabulary skills do affect overall NRT performance. The results are discussed in light of current STM models and the role of long-term phonological knowledge in NRT tasks.

INTRODUCTION

In recent years, nonword repetition performance has received a great deal of attention, both in language and literacy research. In a nonword repetition task, the participant first listens to a nonword, then temporarily stores the novel phonological information in short-term memory, and finally reproduces it. Although there is no consensus yet on what the NRT actually measures (for a review, see Coady & Evans, 2008), this task has been explicitly used in language and literacy research as a measure of the storage, coding, and retrieval of phonological information, i.c., phonological representations in STM. It has been documented that there is a positive correlation between NRT performance and vocabulary knowledge (Gathercole, 2006; Munson, Kurtz & Windsor, 2005) and literacy level (Conti-Ramsden & Durkin, 2007).

A considerable amount of literature has been published on the association between NRT performance and lexical (i.e., long-term) knowledge in typically developing and language impaired *monolingual children* (Edwards, Beckman, & Munson, 2004; Gathercole, 2006; Thordardottir, Kehayia, Lessard, Sutton, & Trudeau, 2010). Different types of studies demonstrated that effective STM encoding of nonwords relies on long-term knowledge of lexical properties, and thus, makes active use of many cues related to both lexical-semantic features (i.e., meaning) and sublexical phonological/phonotactic features (i.e., form) of words. For instance, studies have shown that words are recalled significantly better than nonwords. This semantic advantage in STM has been called the lexicality effect (Gathercole, Pickering, Hall, & Peaker, 2001; Hulme, Maugha, & Brown, 1991). Additionally, it has been demonstrated that structural knowledge of the words in a language (restrictions on possible word forms, such as phonotactic regularities) also facilitates immediate recall. This is known as the word-likeness effect, i.e., the finding that nonwords with high phonotactic probability are recalled better than nonwords with low phonotactic probability (Edwards et al., 2004).

Another indication of an association between long-term linguistic knowledge and the efficacy with which STM encodes phonological information comes from studies in *bilingual populations* (e.g., Thorn & Gathercole, 1999; Cheung, 1996; Yoo & Kaushanskaya, 2012). Several studies have shown that NRT performance in sequential bilingual children is associated with their vocabulary knowledge in that language (e.g., with Cantonese-English sequential bilinguals, Cheung, 1996; with Greek-English sequential bilinguals, Masoura & Gathercole, 1999). In line with this finding, bilingual children perform worse on a NRT based on their second language than monolinguals who speak this language as their first language (e.g., with

Spanish-English sequential bilinguals, Windsor, Kohnert, Lobitz, & Pham, 2010; with Turkish-Dutch (early) sequential bilinguals, Messer, Leseman, Boom, & Mayo, 2010).

These conclusions were mainly based on the overall accuracy scores in NRT tasks, i.e., scores based on a dichotomy between fully correct vs. incorrect repetitions of nonwords (see for instance Messer et al., 2010; Summers, Bohman, Gillam, Pena, & Bedore, 2010; Thordardottir & Brandeker, 2013). In such studies the unit of analysis is the number of nonwords that a participant recalls correctly. Even though a lot has been learnt by using this metric, analyses of this type can only reveal *that* there is a difference between two groups, not *why* there is a difference.

A first step towards addressing the latter question has been made in some bilingual studies, which also included (more) detailed analyses by relying on proportional scores (see for instance Lee & Gorman, 2012; Gutiérrez-Clellen & Simon-Cereijido, 2010). These scores were obtained by dividing the number of phonemes repeated correctly by the number of input phonemes. This type of scoring is more fine-grained than the use of overall accuracy scores, making it possible to focus on types of errors such as phoneme omissions, additions and deletions (see also Sorenson Duncan and Paradis (2016) which, for instance, considered onsets versus codas and Topbaş, Kaçar-Kütükçü and Kopkalli-Yavuz (2014), who considered consonants versus vowels). By comparing the proportion of correctly recalled phonemes one can find out whether the group difference at the level of the overall scores is matched by a difference in the proportion of correctly recalled phonemes. If so, this would suggest that differences in phoneme retention lie at the basis of the group effect in overall NRT performance.

The Gutiérrez-Clellen & Simon-Cereijido study, which compared young Spanish-English bilinguals (English dominant, Spanish dominant and fluent in both languages), demonstrates that, in order to better understand the performance pattern on a NRT, it is important to perform more fine-grained analyses of the responses, i.e., to explore the sublexical processing of nonwords in more detail by focusing at the phoneme level. However, even though proportional scores shed more light on the question why groups differ in their NRT performance, there are two reasons why an even more sensitive measure is desirable. First, the proportional scores mentioned above only reflect participants' skill in remembering the identity of phonemes. Other phoneme properties (see below) are not captured. Second, by analyzing the average proportion of correctly recalled phonemes (aggregation across items) a lot of information is lost, more particularly, how systematically participants of both groups differ with respect to their recall of phoneme identities.

A more systematic study of nonword repetition performance is to make a task analysis. Which processes do participants have to engage in when repeating nonwords, i.e., which types of information do they have to encode and retrieve? In order to correctly reproduce a nonword, participants must have a correct retention of (1) all phonemes appearing in the target nonword, i.e., the phonological content of the nonword (henceforth: 'item information') and (2) the serial position associated with each phoneme, i.e., the serial order of the phonemes in the nonword (henceforth: 'serial order information').

Interestingly, this distinction between processing item information and serial order information has already received some attention in current models of STM, based on memory span tasks in which a series of stimuli (i.e., words, letters, and digits) have to be recalled (for a review, see Majerus & Cowan, 2016). Most of this research has been carried out in a monolingual context (Burgess & Hitch, 1999, 2006; Page & Norris, 1998; Gupta, 2003) with the exception of two studies, one behavioral study including monolingual English as well as simultaneous and sequential English-Spanisch bilinguals (Majerus, Poncelet, Van der Linden, & Weekes, 2008) and one fMRI-study comparing German-French bilinguals who differed in their second language proficiency (Majerus et al., 2008). Although there is some controversy about the exact way in which serial order information is represented (Hitch, Fastame, & Flude, 2005; Majerus et al., 2008), there seems to be a consensus that a specialized STM system is dedicated to the representation of the serial order between items, whereas representation of item information depends directly on the activation of the language system (i.e., LTM knowledge).

Evidence for the distinction between these two types of information comes, for instance, from studies indicating an increased performance in retrieving item information (fewer item errors) but not in retrieving the items' serial order. This was observed in tasks revealing higher recall of semantically related vs. unrelated words, i.e., the semantic similarity effect, or higher recall of high vs. low frequency words, i.e., the word frequency effect (Poirier & Saint-Aubin, 1996; Nairne & Kelley, 2004). Research from Gupta and colleagues suggests the existence of an ordering mechanism in STM (Gupta, 2003, 2009; Gupta, Lipinski, Abbs, & Lin, 2005). Evidence for this idea was furnished by their finding of syllable primacy and recency effects in a NRT task (Gupta, 2003; Gupta et al., 2005). Syllable primacy effects refer to the finding that first syllables in polysyllabic words are better recalled than syllables in mid-word positions, whereas syllable recency effects refer to the finding that final syllables are better recalled than syllables in mid-word positions. Based on their results, and the fact that primacy and recency effects are the hallmark characteristics of immediate serial recall, Gupta et al. (2005) postulated the existence of a "sequence memory" mechanism that underlies performance in STM tasks.

They argue that this mechanism plays an important role in temporarily maintaining and repeating syllable sequences that form a nonword.

Where does this leave us with respect to understanding NRT performance in bilinguals? The aforementioned effects suggest the importance of distinguishing between item and serial order retention. However, as explained above, the focus in NRT experiments on participants' number of correct responses or on their average proportion of correctly recalled phonemes makes it impossible to distinguish between the retention of a nonword's phonemes and the retention of their serial order. Indeed, item and serial order information are by definition collapsed in correct responses, as these are instances where all phonemes are recalled in their correct serial order. Analyzing the proportion of correctly recalled phonemes is already a first step in the direction of an error analysis, but is restricted to participants' ability to recall phonemes' identity, not their serial position. As nonword repetition requires the retention of both item and serial order information we need a way to distinguish between errors that result from a failure to recall all phonemes in a nonword and errors that result from a failure to recall phonemes' serial position (some errors will probably reflect both types of failure). The explicit aim of the current study is to disentangle participants' retention of the phonemes themselves (item STM processing) from their retention of the phonemes' serial order (serial order STM processing).

Recently, Schraeyen, Geudens, Ghesquière, Van der Elst and Sandra (2017) conducted a NRT study in which they compared more skilled and less skilled young monolingual readers. They analyzed both nonword level performance (i.e., number of correct nonwords responses) and phoneme level performance (i.e., an analysis of participants' recall performance at the subitem level). More precisely, they distinguished between phoneme identity performance, defined as the number of responses without phoneme identity errors (but with possible serial order errors), and serial order performance, defined as the number of responses without serial order errors (but with possible phoneme identity errors). Literacy skill was not linked to the ability to retain phonemes (item information). However, literacy skill was linked to the retention of phonemes' serial order, higher serial order scores being significantly associated with higher literacy scores).

In the current study, we will use the same technique to compare NRT performance in bilinguals and monolinguals. We want to assess whether participants' language group (monolinguals vs. bilinguals) predicts their phoneme identity and serial order performance when repeating nonwords. In order to accomplish this goal we will first investigate whether we can detect an effect of Language Group on the retention of nonwords, as has been shown in

many previous studies (e.g., Lee & Gorman, 2012). Next, we will investigate whether Language group has an effect on the retention of the identity of phonemes and/or on the retention of their serial order. Finally, we will analyze whether the two groups differ with respect to the distribution of the different types of error responses (only phoneme identity errors, only serial order errors, both error types combined).

Our major objective is to find out whether the difference between bilinguals and monolinguals will be reflected in different NRT performance patterns at the phoneme level, more particularly, in their retention of phonemes' identity and their serial order. Given the idea that item information depends on the activation of the language system (cf. Majerus et al., 2008), it is reasonable to hypothesize that participants with a better knowledge of Dutch vocabulary (monolinguals) will have a better retention of phonemes compared to participants with a smaller vocabulary in Dutch (bilinguals). At the same time, we want to know whether this expected difference will hold for the retention of phonemes' serial order as well. If retention of serial order is related to a specialized system in STM, as suggested by previous literacy research (e.g., Majerus et al., 2008), one would expect a negative answer. There is no reason to assume that typically developing monolinguals and bilinguals will differ in the way this STM mechanism operates.

However, the prediction with respect to serial order retention is complicated by the fact that the retention of an unfamiliar phoneme sequence (a nonword) is not quite identical to the retention of a sequentially ordered set of random items. For the purposes of coding and retention of phonemes' serial order, linguistic LTM knowledge may also play a role, more particularly, knowledge of the combinatorial possibilities of the phonemes in the language (so called phonotactic restrictions). Knowledge of these (language-specific) phonotactic restrictions might be beneficial for encoding and retaining phonemes' serial order. As this LTM knowledge is likely to be better developed in monolingual children, these are likely to outperform bilingual children on the retention of phonemes' serial order as well.

Participants in our monolingual group speak Dutch as their native language. Children in the bilingual group speak Turkish as their native language and acquired Dutch as a second language (for a detailed review of Dutch/Turkish speech acquisition, see McLeod, 2007). Even though the role of phonotactic knowledge must be considered when two such different languages are involved, this problem was ruled out by the specific stimulus set for the NRT. Despite some differences between the phoneme inventories of Turkish and Dutch, especially in the vowel set, our nonwords, constructed according to the phonological rules in Dutch, happened to respect the phonotactic restrictions in the bilinguals' native language (see Materials

and Design Section; for a review of Dutch/Turkish orthography, see Bosman, de Graaff, & Gijsel, 2006; Öney & Durgunoglu, 1997). For instance, in Turkish, consonant clusters, a hallmark of phonotactics, rarely consist of more than two consonants. In addition, they never occur at the onset of native words (Topbaş & Kopkalli-Yavuz, 2008). None of these phonotactic properties of the bilinguals' native language were violated in our item set. Hence, we do not expect a difference between the monolingual and bilingual groups with respect to serial order recall, as the participant groups are not expected to differ with respect to their STM sequencing memory (the literacy development of all children being normal) and the items' phonotactic constraints are not expected to be more difficult for the bilingual group.

A second objective of our study is to evaluate the impact of prior language knowledge, more particularly, expressive vocabulary knowledge, to the overall performance on a NRT, i.e., the number of correct responses. If pre-existing language knowledge (LTM knowledge) is critical to overall NRT performance, participants with better expressive vocabulary knowledge in Dutch (monolingual group) should outperform participants with a smaller vocabulary size in Dutch (bilingual group).

METHOD

PARTICIPANTS

Monolingual and bilingual participants were recruited in two primary schools located in the central region of Belgium[12]. Eighteen Turkish-Dutch bilinguals participated in the study (mean age: 8 years 11 months, SD = 5 months, range 8 years 3 months – 9 years 6 months). All participants spoke Turkish as L1 and had acquired Dutch as L2 during their early school years, with a mean age of acquisition of 3 years 8 months (SD = 7 months, range 2 years 8 months – 4 years 2 months). All bilingual participants acquired their home language first and were exposed to Dutch upon entering Dutch day care, preschool, or kindergarten settings. At the time of testing, they communicated primarily in Dutch with siblings and peers at school and used their native language in their home environment. Twenty Dutch-speaking monolinguals participated in this study (mean age: 8 years 9 months, SD = 3 months, range 8 years 4 months -9 years 3 months). The monolingual participants had not been exposed to any other foreign language, except Dutch. The groups did not differ in age (p = .11). All children were in third

[12] In the description of these two groups all children who repeated their year have been removed.

grade and developed normally according to parent reports and teacher interviews and reports. They had no speech, language, hearing problems, or other serious concerns in any developmental area.

MATERIALS AND DESIGN

Children were tested individually by a native speaker of Dutch. They had to perform a vocabulary test and a NRT.

Vocabulary test. The Dutch standardized version of the CELF-4-NL (Semel, Wiig, & Secord, 2008) was administered. The subtest 'Expressive vocabulary' was used to evaluate the participants' productive lexical knowledge in Dutch[13]. This task requires participants to name an object, person, or activity portrayed in an illustration.

Nonword Repetition Task (NRT). We administered the Dutch adaptation (Scheltinga, 2003) of the Children's Test of Nonword Repetition (Gathercole et al., 1994). However, we did one small adaptation, as we did not include the two-syllable items that were included in the original version. We ended with 36 nonwords (see Appendix A). As mentioned in the introduction all items respected the phonotactic properties of Turkish as well. All test items were prerecorded on a CD by a female native speaker of Dutch and presented once to the participant through headphones. Responses were recorded via a head-mounted microphone-for subsequent offline scoring of repetition accuracy. Corrective feedback was provided for two practice items. The child was then presented with the 36 test items and asked to repeat the presented nonword. The test items were blocked on the basis of their syllable length, starting with three-syllable items and systematically moving up to five-syllable items.

Each nonword repetition response was phonetically transcribed by the experimenter using SAMPA (Speech Assessment Methods Phonetic Alphabet) and scored with respect to retention at the nonword level (global analysis) and the phoneme level (in-depth analysis). For the purpose of reliability checking, all responses were transcribed and scored by a second judge, who was blind to the original transcriptions and scoring. For the transcription, an inter-judge agreement of 92.3% was found, based on the ratio of the number of identically transcribed phonemes on the total number of transcribed phonemes. For the scoring, there was also high agreement between the scores at each linguistic level, i.c. 98.6% at the nonword level-and 93.9

[13] It is important to note that in case of bilinguals, measuring their lexical knowledge in only one of their languages will differ from their total vocabulary (in both languages).

% at the phoneme level. In case of disagreement, the experimenter and second judge decided together how the responses were scored.

SCORING

In a first set of analyses, we evaluated each response in three different ways. First, we checked whether the item was correctly recalled or not (nonword level score). Second, we scored the response with respect to the two phoneme related variables: (1) phoneme identity and (2) serial order (phoneme level scores). Hence, we used the following scoring scheme:

Nonword level score. This score reflects the overall correctness of the response, i.e., nonword level performance. A response was counted as correct and given a binary score of 1 if the entire nonword was correctly recalled, i.e., no errors were made at the stress level, the syllable level, or the phoneme level. For example, for the target nonword /*pig.dúl.mek*/ the response /*pig.dúl.mek*/, would receive a score of 1. At this level, any other response was scored as 0, i.e., an error response.

Phoneme level scores. To evaluate children's ability to retain phonemes' identity and serial order when repeating a nonword, each response was given a binary score on these two phoneme-related variables. As mentioned earlier nonword repetition requires the recall of both the phonemes themselves and their serial order. Given our binary coding of each response on these two variables four different response types can be distinguished: (a) a correctly repeated nonword was given a binary score of 1 on both phoneme-related variables; (b) an incorrect response containing only one or more phoneme identity errors, was given a binary score of 0 on phoneme identity retention and a score of 1 on serial order retention; (c) an incorrect response containing only one or more serial order errors, was given a binary score of 0 on serial order retention and a score of 1 on phoneme identity retention; (d) an incorrect response containing both phoneme identity and serial order errors, was given a binary score of 0 on both phoneme-related variables.

Our classification of responses in these four categories was based on the following criteria:

(1) *No identity or serial order errors* (correct responses): all target phonemes appeared in the same serial order in the response, i.e., target: /*pig.dúl.mek*/, response: /*pig.dúl.mek*/.

(2) *Phoneme identity errors.* (a) *Phoneme addition*: one or more phonemes were added in the response compared to the original nonword. For example, target: /*pig.dúl.mek*//, response: /*pig.dúls.mek*/. (b) *Phoneme substitution*: one or more phonemes in the target were replaced by

other phonemes in the response, e.g., target: /*pig.dúl.mek*/, response: /*pig.dól.mek*/. Here, the target phoneme /*u*/ was replaced by the phoneme /*o*/. (c) *Phoneme omission*: one or more phonemes in the target were omitted in the response, for example, target: /*pig.dúl.mek*/, response: /*pig.dú.mek*/, the phoneme /*l*/ was omitted in the response. (c) *Phoneme substitution in combination with phoneme omission*, for example, target: /*pig.dúl.mek*/, response: /*pis.dú.mek*/. In the response, the target phoneme /*g*/ was replaced by the phoneme /s/ whereas the target phoneme /*l*/ was omitted.

Phoneme identity errors do not by themselves disturb the serial order of the correctly recalled target phonemes, i.e., their relative position in the nonword stimulus is preserved (for the distinction between relative position vs. absolute position, see further). They can obviously co-occur with serial order errors (see below).

(3) *Serial order errors.* Note that it is obvious, both from a logical and from an analytical perspective, that the retention of a phoneme's serial order can only be determined if the phoneme itself has been correctly recalled. Indeed, it makes no sense to attribute positional properties to phonemes that are absent in the output (omissions) or did not appear in the input (additions, substitutions). Moreover, if each phoneme identity error also counted as a serial order error, the former errors would be a subset of the latter (potentially a large subset, depending on the number of responses with serial order errors and no phoneme identity errors). This would be at odds with our intention to separate these two retention variables.

To determine whether the serial order of the correctly recalled phonemes was respected, we used the procedure suggested by McKelvie (1987). Using this approach, input phoneme strings of at least 2 phonemes that appear in the response are considered correct irrespective of their absolute position when counting from the beginning or the end of the nonword sequence. Next, all remaining single phonemes that appear in the target are handled. The serial position of a single phoneme is correct if it appears in the correct serial order when counting either from the beginning or the end of the target nonword, but not both. For example, given the nonword /*pig.dúl.mek*/ and the response /*pil.dúg.mek*/, the response would be checked as follows. With respect to phoneme identity all 9 phonemes are correctly identified. With respect to serial order the phonemes /*pi*/, /*du*/, /*mek*/ are correctly ordered when using the McKelvie method. However, the single phonemes /*g*/ and /*l*/ are swopped between the third to sixth positions, 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 target nonword. Such *phoneme shifts* were the only error type containing a serial order error without an identity error: each target phoneme

appeared in the response, but some phonemes were recalled at another (i.e., incorrect) serial position,

(4) *Combination of identity and serial order errors. (a) Phoneme shift in combination with phoneme substitution*, e.g., target: /*pig.dúl.mek*/, response: /*pil.dúg.nek*/, target phonemes /*g*/ and /*l*/ are swopped (giving rise to a serial order error) whereas target phoneme /*m*/ is replaced by /n/ (giving rise to an identity error). *(b) Phoneme shift in combination with phoneme omission, e.g.,* target: /*pig.dúl.mek*/, response: /*pug.dí.mek*/, target phonemes /*i*/ and /*u*/ are swopped (giving rise to a serial order error) and target phoneme /*l*/ is omitted (giving rise to an identity error). *(c) Phoneme shift in combination with a phoneme substitution and a phoneme omission:* for example, target: /*pig.dúl.mek*/, response: /*pug.dí.nek*/, target phonemes /*i*/ and /*u*/ are swopped (giving rise to a serial order error), target phoneme /*m*/ is replaced by /*n*/ and target phoneme /*l*/ 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 different response types and their binary values on both retention variables at the phoneme level.

Table 1

Overview of response types and their associated binary values with respect to phonemes' identity and serial order performance.

Error type	Identity Performance	Serial order Performance
Correct response	1	1
Phoneme addition	0	1
Phoneme substitution	0	1
Phoneme omission	0	1
Phoneme substitution and omission	0	1
Phoneme substitution and addition	0	1
Phoneme shift	1	0
Phoneme shift and phoneme omission	0	0
Phoneme shift and phoneme substitution	0	0
Phoneme shift and phoneme addition	0	0
Phoneme shift, phoneme substitution and phoneme omission	0	0
Phoneme shift, phoneme substitution and phoneme addition	0	0

Note. Binary values (0: error, 1: correct) 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 (final five rows) were incorporated in the statistical analyses of both phoneme identity performance and serial order performance (see text).

Phoneme Identity and Serial Order were treated as binary variables because we wanted to find out whether participants were able to repeat the nonword without making any phoneme identity error and/or without making any serial order error. This scoring technique is quite similar to the way responses are scored at the nonword level, i.e., have ALL phonemes been repeated in their *correct* serial position (hence, a correct response) or not? As a matter of fact, this question is split up in two separate questions, which makes it possible to probe into the phoneme level: (a) have all phonemes been repeated? (phoneme identity retention) and (b) have all correctly repeated phonemes been placed in their correct serial position? (serial order retention). The use of a binary coding system is not new in STM research. Note that most STM tasks and NRT

tasks make use of binary scores (see for instance Martens & de Jong, 2006; Grainger, 1990). Recently, Schraeyen, Geudens, Ghesquière, Van der Elst, & Sandra (2017) used the binary coding scheme described above to investigate the relationship between young children's literacy skill and their retention performance at the phoneme level. By using binary scores at the phoneme level we address the following question at the conceptual level: has the participant been able to make a response without any failure of the identity and/or serial order retention mechanism? As mentioned above we hypothesize that this ability will be stronger in monolinguals than in bilinguals for the retention of phonemes' identity but not for their serial order. It might be tempting to suggest the use of proportional scores when analyzing the retention of phonemes' identity and serial order, e.g., the proportion of phonemes from the nonword that has been correctly repeated. However, proportions create a non-trivial problem when the data are analyzed with a linear mixed effects model, i.e., a model in which both the theoretically relevant predictors and several random variables (here: participants and items) are included. First, it is well-known that proportions often give rise to marginal residuals that are not normally distributed, whereas this is a basic assumption of such models. Marginal residuals are the difference scores between the observed values and the values that are predicted by the fitted model. Second, proportions have the undesirable property that predicted values are sometimes larger than the theoretical maximum of 1. Obviously, it is impossible to observe a proportion that is larger than 1. These two problems actually occurred when scoring the responses in the form of proportional data. Importantly, it can be shown that these two disadvantages do not occur for binary data, which are analyzed with a generalized linear mixed effects model with a logit link (see for instance Cohen, Cohen, West, & Aiken, 2003). Hence, the only way to study the responses at the phoneme level in a statistically sound way is by scoring these responses with respect to two binary variables. As argued above, these analyses run quite parallel to the way responses are analyzed at the nonword level and, moreover, can be translated in sensible questions at the conceptual level.

In a second set of analyses we wanted to study the effect of language background (monolingual vs. bilingual) on the distribution of the three error types. To this end we contrasted pure identity errors (i.e., responses with one or more failures to retain a phoneme's identity but without serial order errors) with all other errors, (2) pure serial order errors with all other errors and (3) combined errors with all other errors.

STATISTICAL ANALYSES

At first, Pearson correlation coefficients were calculated between vocabulary scores and the three dependent variables: nonword level performance, phoneme identity performance, and serial order performance. These correlations enabled us to first explore the linear association between the different measures used in this study. For this purpose, we used total sum scores, i.e., the total number of correct responses (with respect to each dependent variable) per participant, aggregated across all items.

Next, we performed a more fine-grained analysis in which the effect of the independent variables was examined at the level of the *individual item responses*. To this end, the data were analyzed with generalized linear mixed-effects models (GLMMs). A logistic link function was used because the dependent variables were binary. GLMMs make it possible to simultaneously assess the variance explained by the random effect of participants and the random effect of items. Inclusion of the latter is necessary to avoid the language-as-a-fixed-effect fallacy (Clark, 1973). In all initial models, we included Language Group (coded as bilingual = 0, monolingual = 1), Syllable Length (3, 4, and 5 syllables, coded with 2 dummies and syllable length = 3 as the reference category), and the Language Group x Syllable Length interaction as fixed effects. In addition, Serial Order performance was used as a covariate in the analysis of Phoneme Identity performance (and vice versa) to statistically control for the collinearity between these factors (see Table 3). It is common practice to use such an approach (see Verbeke & Molenberghs, 2000). During the model-building phase, it was examined whether removing the interaction term significantly decreased the model fit. All model parameters were estimated using the Maximum Likelihood method, which has the advantage that likelihood-ratio tests can be conducted to formally compare the fit of the different (nested) models (e.g., the model with the interaction term and the model without this term). Once the final model was obtained, it was refitted using the Restricted Maximum Likelihood approach because the latter approach yields more accurate estimates of the variance components in smaller samples. For the mathematical details of GLMMs the interested reader is referred to Molenberghs and Verbeke (2005). All models were fitted using R (packages lme4, Bates & Maechler, 2009; language R, Baayen, 2008). Throughout the paper, the significance level was set at $p \leq .05$.

RESULTS

VOCABULARY PERFORMANCE AND NRT ACCURACY

Table 2 presents the means and standard deviations for the expressive vocabulary measure and the overall accuracy scores on the NRT measure. Independent-samples t-tests for each measure separately revealed that children in the monolingual group had higher expressive vocabulary scores than those in the bilingual group, $t(36) = 8.48$, $p < .0001$. In addition, the monolingual group had higher overall accuracy NRT-scores than the bilingual group, $t(36) = 2.75$, $p = .01$.

Table 2

Descriptive statistics of the Vocabulary Measure and the overall accuracy scores on the NRT-measure for the Monolingual and Bilingual Groups

	Monolingual Group (n = 20)			Bilingual Group (n = 18)		
	M	SD	Range	M	SD	Range
Vocabulary (raw scores)	45.7	5.8	30-53	26.7	7.9	16-42
Nonword Repetition (raw scores)	15.1	5.9	5-26	10.3	4.9	2-17

CORRELATION ANALYSES

As can be seen in Table 3, we found high positive (Pearson) correlations between the participant totals in all pairwise combinations of the three dependent variables: nonword retention, phoneme identity retention, and serial order retention (all rs $\geq .86$, all ps $< .0001$). This should not come as a surprise, as three of the four response types[14] imply identical binary values on the nonword level variable and one phoneme level variable. For the correlation involving performance on phoneme identity, these are correct responses, phoneme identity errors, and combined errors (representing 94.43% and 93.83% of all responses in the monolingual and bilingual groups, respectively). For the correlation involving the performance on phonemes' serial order, these are correct responses, serial order errors, and combined errors (representing 71.80% and 62.33% of all responses in the monolingual and bilingual groups, respectively).

[14] The 4 response types are: correct responses, phoneme identity errors, serial order errors and combined errors.

Hence, only the proportion of responses of the fourth response type (phoneme serial order error in the case of phoneme identity variable; phoneme identity error in the case of serial order variable), weaken the correlation. In the absence of these responses, there would be a perfect correlation between nonword performance and performance at the phoneme level variables. The high positive correlation between the two phoneme-related variables is not surprising either, as correct responses and combined errors imply identical scores on these two variables (representing 69.13% and 59.57% of all responses in the monolingual and bilingual groups, respectively). In contrast, the identity and serial order errors weaken the correlation, such that the distribution of the two categories of response types determines the exact value of the correlation.

Table 3

Pearson's Correlations and Levels of Significance (Two-Sided Tests) among the Main Variables of the Study

	1	2	3	4
1. Nonword Level Retention	-			
2. Phoneme Identity Retention	.98***	-		
3. Serial Order Retention	.88***	.86***	-	
4. Language Group	.41**	.39*	.15	-

Note. The bilingual and monolingual language groups were assigned the binary values of 0 and 1, respectively.
* $p < .05$, ** $p < .01$, *** $p < .001$

The correlations between Language Group and the three dependent variables were substantially lower. Language Group was significantly positively correlated with nonword level performance ($r = .41$, $p < .01$) and phoneme identity performance ($r = .39$, $p < .01$), but not with serial order performance ($r = .15$, $p > .05$). These correlations are in line with our hypotheses: monolinguals were predicted to outperform bilinguals on the overall accuracy of their responses and on the retention of phonemes' identity but not on the retention of their serial order. Below we will report the results of the GLMMs in which the effect of Language Group on each of our three dependent measures was analyzed (based on item level scores).

NONWORD LEVEL PERFORMANCE IN MONOLINGUALS AND BILINGUALS

The results of the GLMM with overall nonword performance as the dependent variable are summarized in Table 4. It turned out that the two-way interaction Language Group x Syllable Length was not significant, $p > .05$. Therefore, in the final model, this interaction term was removed.

Table 4

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Nonword Level Performance from Language Group and Syllable Length

Parameter	β	SE(β)	z	p (>|z|)
(Intercept)	.12	.41	.29	.77
Monolingual Group	.97	.35	2.79	.005
Syllable Length 4	- 1.64	.48	- 3.40	.0007
Syllable Length 5	- 3.40	.51	- 6.70	< .0001

Note. Syllable Length was dummy coded, using syllable length 3 as the reference category, and Language Group was dummy coded, using the bilingual group as the reference category (bilinguals = 0, monolinguals = 1).

As expected, a main effect of Language Group was observed, $p = .005$, demonstrating that the probability of a correct response (i.e., a correct nonword repetition) was higher for the monolingual group than for the bilingual group. We also observed a main effect of Syllable Length, indicating that the probability of a correct response was higher for three-syllable nonwords than for four-syllable nonwords, $p = .0007$, and was higher for three-syllable nonwords than for five-syllable nonwords, $p < .0001$. In addition, the probability of a correct response was higher for four-syllable nonwords than for five-syllable nonwords, $\beta = -1.76$ for syllable length 5, $p = .0004$ (this β estimate is obtained by fitting the same model as shown in Table 4 but using syllable length 4 as the reference level; data not shown).

The model-predicted scores showed that monolinguals made 15% more correct nonword repetitions than bilinguals. Four-syllable nonwords yielded 37% fewer correct responses than three-syllable items, whereas five-syllable nonwords gave rise to 21% fewer correct responses than four-syllable items.

PHONEME LEVEL PERFORMANCE IN MONOLINGUALS AND BILINGUALS

At the level of phoneme performance, we distinguished between the retention of (a) phoneme identity and (b) serial order. In Table 5, the mean error rates for the different response types are presented as a function of both language group and syllable length.

Table 5

Mean Error Percentage for Each Syllable Length in Monolinguals and Bilinguals

	Monolinguals			Bilinguals		
	3-syllable items	4-syllable Items	5-syllable items	3-syllable items	4-syllable items	5-syllable Items
No phoneme error (%)	67.9	42.1	15.8	53.7	24.1	7.9
Phoneme identity error (%)	29.6	53.3	74.6	42.6	72.2	81.0
Serial order error (%)	7.1	29.6	52.9	10.6	38.4	52.3
Combined phoneme error (%)	5.4	25.8	50.4	7.4	36.1	49.5
No response (%)	.8	.8	7.1	.5	1.4	8.3

Note. The total adds up to more than 100% as the combined errors were also counted as errors on phoneme identity and errors on serial order. They are included to show how often these error types co-occurred in a response.

Note that in the GLMM analyses below, all responses were included, as each response (i.e., correct or incorrect) could be scored with respect to the two phoneme-related variables. In effect, each response was a demonstration of the child's ability (a) to recall all phoneme identities correctly or not and (b) to recall their serial order correctly or not.

Retention of phoneme Identity. Table 6 summarizes the results of the final model for the analysis of children's phoneme identity performance. The interaction Language Group x Syllable Length was not significant and, therefore, removed in the final model. The final GLMM revealed a main effect of Language Group after statistically controlling for serial order performance, $p = .003$. Monolingual children made significantly more responses in which all phonemes were correctly recalled. The effect of serial order performance on the retention of phoneme identity was significant ($p < .0001$), which is due to the strong correlation between the two phoneme-related variables (see also Table 3). Importantly, the effect of Language Group on the retention of phoneme identity was significant when this correlation was brought

under statistical control. As the effect of Language Group could not be contaminated by this collinearity it only accounts for the variance in phoneme identity retention that could not also be explained by serial order retention. In addition, we observed a main effect of Syllable Length for both monolinguals and bilinguals (no interaction with Language Group). Phoneme identity performance was higher in three-syllable nonwords compared to four-syllable nonwords, $p = .002$, and compared to five-syllable nonwords, $p < .0001$. Similarly, phoneme identity performance was higher in four-syllable nonwords compared to five-syllable nonwords, $\beta = -1.24$ for syllable length 5, $p = .0002$ (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 the model-predicted scores, monolinguals were 15% better in correctly repeating all phonemes of a nonword. Successful performance on four-syllable nonwords was 21% worse than on three-syllable items, whereas it was 25% worse on five-syllable nonwords than on four-syllable items.

Table 6

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Phoneme Identity Performance from Language Group and Syllable Length when Statistically Controlling for Serial Order Performance

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	.42	.32	1.28	.20
Monolingual Group	.84	.29	2.96	.003
Syllable Length 4	- 1.18	.38	- 3.10	.002
Syllable Length 5	- 2.41	.41	- 5.89	< .0001
Serial Order Performance	1.64	.22	7.52	< .0001

Note. Syllable Length was dummy coded, using syllable length 3 as the reference category, and Monolingual Group was dummy coded, using the bilingual group as the reference category (bilinguals = 0, monolinguals = 1).

Retention of serial order. Table 7 summarizes the results of the final model for the analysis of serial order performance. The interaction Language Group x Syllable Length was not significant and, therefore, removed in the final model. The final GLMM revealed no main effect of Language Group, $p = .63$, indicating that there was no significant influence of Language Group on the retention of serial order. We observed a significant main effect of Syllable Length (all $p < .0001$). This indicates that serial order retention decreased with increasing syllable length. For both monolinguals and bilinguals (no interaction with Language Group), serial order

performance was higher in three-syllable nonwords than in four-syllable nonwords, $p < .0001$, and five-syllable nonwords, $p < .0001$. Similarly, serial order performance was higher in four-syllable nonwords than in five-syllable nonwords, $\beta = -.96$ for syllable length 5, $p < .0001$ (this β estimate is obtained by fitting the same model as shown in Table 7 but using syllable length 4 as the reference level; data not shown). Finally, children's phoneme identity performance significantly decreased as their serial order decreased ($p < .0001$, see also the high correlation value in Table 3). As the GLMM brings this correlation under statistical control the effect of Language Group on serial order performance is uncontaminated by this collinearity in the data.

In the model-predicted scores, the monolinguals performed only 1% better than the bilinguals. In the above analysis, Language Group was a non-significant effect. Successful performance was 6% higher for three-syllable nonwords than for four-syllable nonwords and was 9% higher for five-syllable nonwords than for four-syllable nonwords.

Table 7

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Serial Order Performance from Language Group and Syllable Length when Statistically Controlling for Phoneme Identity Performance

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	3.81	.42	9.04	<.0001
Monolingual Group	.13	.28	.48	.63
Syllable Length 4	- 1.89	.45	- 4.18	<.0001
Syllable Length 5	- 2.84	.46	- 6.16	<.0001
Phoneme Identity Performance	1.65	.22	7.58	<.0001

Note. Syllable Length was dummy coded, using syllable length 3 as the reference category, and Monolingual Group was dummy coded, using the bilingual group as the reference category (bilinguals = 0, monolinguals = 1).

DISTRIBUTION OF ERROR TYPES IN MONOLINGUALS AND BILINGUALS

All preceding analyses were performed on all data and, hence, included the four type response types: correct responses, pure identity errors, pure serial order errors, and combined errors. In the analysis below we will look at the effect of Language Group on the distribution of the error types only. When there is a retention failure, the two groups may differ in the association strength between the two retention mechanisms, i.e., the probability that, if one error type is made, the other one is also made. For instance, one group might make more pure identity errors

than the other group, whereas the opposite pattern might occur for combined errors. This would reveal that the two retention mechanisms are more strongly associated in one group than in the other. Table 8 provides an overview of the number (and percentage) of error responses per error category.

Table 8

Number and percentage of error responses per error category.

	Monolingual Group (n = 20)		Bilingual Group (n = 18)	
	N	%	N	%
Pure Phoneme Identity Error	182	45.8	222	50.3
Pure Serial Order Error	19	4.8	18	4.1
Combined Error	196	49.4	201	45.6

Clearly, pure serial order errors, i.e., a problem with the retention of phonemes' serial order when the retention of all phonemes' identity is perfect, occur very seldom, both in the monolingual group (4.8 %) and bilingual group (4.1 %). Not surprisingly, a GLMM in which the probability of such an error was predicted by Language Group revealed no effect of this predictor, $\beta = .21$ (SE $\beta = .18$), $p = .66$. Considering these facts, we decided to remove this error type from the analysis.

Hence, the purpose of the analysis is to investigate the effect of Language Group on the distribution of pure identity errors and combined errors, both representing more than 40% of the total set of error responses in the two groups. Interestingly, these two error types differ only in the presence or absence of a serial order error (combined errors vs. pure identity errors, respectively). Hence, the distribution of the two error types is determined by the likelihood of making a serial order error if a phoneme identity error has been made. The more serial errors are made, the stronger the co-occurrence of the two error types. If Language Group has a significant effect on the serial order scores in this subset of the data (the identity score being constant), this will reveal a different association strength between identity errors and serial order errors. Table 9 summarizes the results of the final model for the analysis of these two error types (pure identity vs. combined errors).

Table 9

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Serial Order Errors in the Set of Pure Identity Errors and Combined Errors from Language Group and Syllable Length

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	- 2.26	.42	- 5.37	<.0001
Monolingual Group	- .012	.29	- .04	.97
Syllable Length 4	2.05	.47	4.34	<.0001
Syllable Length 5	2.86	.48	5.97	<.0001

The interaction Language Group x Syllable Length was not significant and, therefore, removed in the final model. The final GLMM revealed no main effect of Language Group, $p = .97$, indicating that the probability of making a serial order error, given that an identity error is made, is the same for the monolingual and bilingual groups. Hence, the association strength between the two types of retention errors is the same in both language groups. In addition, for both groups, the probability of making a serial order error, given that an identity error is made, increases with increasing syllable length.

The above contrast between two response types amounted to an investigation of the effect of Language Group on a conditional probability, i.e., the probability of making a serial order error if an identity error has been made. It would have been interesting to make the complimentary analysis: the effect of Language Group on the retention of phoneme identity if a serial error has been made. However, the small percentage of pure serial order errors makes this a hazardous exercise in a statistical test (GLMM) that takes the individual response as the unit of analysis. However, the more robust chi-square test that compares the two language groups with respect to the distribution of the two error types, i.e., pure serial order errors and combined errors, reveals that this distribution does not differ between the two groups ($\chi^2 = .43$, $p = .51$). Incidentally, the complimentary chi-square test, assessing the effect of language group on the effect of serial order if an identity error has been made was non-significant as well, confirming the outcome of the above GLMM ($\chi^2 = 1.03$, $p = .31$). Together, these results suggest that the association strength between retention errors at the level of phoneme identity and at the level of serial order was the same for the monolingual and bilingual children. There was a strong correlation between the two error types (Table 3) but the conditional probability

of making one phoneme-related error if an error on the other phoneme-related variable had been made did not differ between groups.

As mentioned above, the technique of varying the binary value on one phoneme-related variable while keeping the value on the other variable constant could not be used for pairwise comparisons involving the minority response group of pure serial order errors. However, we used the technique to assess the effect of language group on the comparison between correct responses and pure identity errors (both representing a considerable percentage of the data), i.e., a contrast in which all responses were correct with respect to the retention of serial order but differed with respect to the retention of phonemes' identity. This analysis makes it possible to study the effect of Language Group on identity retention if NO serial error has been made. In this analysis the effect of language group on identity retention can be studied without the need to statistically control for its correlation with serial order retention (as in the earlier reported analysis on all responses), as the latter variable is restricted to the response subset in which no serial order errors are made.

Table 10 summarizes the results of the final model for the analysis of phoneme identity performance in the absence of serial order errors. The interaction Language Group x Syllable Length was not significant and, therefore, removed in the final model. The final GLMM revealed a main effect of Language Group, $p = .0005$. Monolingual children made significantly more responses in which all phonemes were correctly recalled. In addition, we observed a main effect of Syllable Length for both monolinguals and bilinguals (no interaction with Language Group). Phoneme identity performance was higher in three-syllable nonwords than in four-syllable nonwords, $p = .006$, and five-syllable nonwords, $p < .0001$. Similarly, phoneme identity performance was higher in four-syllable nonwords than in five-syllable nonwords, $\beta = -1.22$ for syllable length 5, $p = .003$ (this β estimate is obtained by fitting the same model as shown in Table 10 but using syllable length 4 as the reference level; data not shown).

Table 10

Estimated Fixed Effects of the Generalized Linear Mixed Model Predicting Phoneme Identity Performance from Language Group and Syllable Length in the Absence of Serial Order Errors.

Parameter	β	SE(β)	z	P(>|z|)
(Intercept)	.34	.31	1.11	.27
Monolingual Group	.92	.26	3.47	.0005
Syllable Length 4	- 1.03	.38	- 2.72	.006
Syllable Length 5	- 2.25	.42	- 5.37	< .0001

Note. Syllable Length was dummy coded, using syllable length 3 as the reference category, and Monolingual Group was dummy coded, using the bilingual group as the reference category (bilinguals = 0, monolinguals = 1).

DISCUSSION

The present study was designed to further understand earlier reported overall accuracy differences in NRT performance between bilinguals and monolinguals by means of an in-depth analysis of their nonword repetition responses. By taking retention ability at the *phoneme level* into account, we collected more detailed information on the determinants of typically developing bilinguals' STM capacity for the encoding and retrieval of non-native phonological information than has been revealed in previous studies.

In bilingual language assessment, a STM task such as a NRT has been proposed as a valuable component of a language assessment protocol. However, so far, NRT performance has been mainly assessed by determining the percentage of correctly repeated nonwords (number of correct responses on the total number of items in the test) or the percentage of correctly recalled phonemes in the participant's responses (Gutiérrez-Clellen & Simon-Cereijido, 2010), except for some recent studies, which, for instance, considered onsets versus codas (Sorenson Duncan & Paradis, 2016), or consonants versus vowels (Topbaş, Kaçar-Kütükçü, & Kopkalli-Yavuz, 2014). In addition, Girbau and Schwartz (2008) examined NRT performance in sequential bilingual children with and without language impairment who had Spanish as their first language and English as their second language. They analyzed substitution, omission, and addition errors. Even though these studies made an important step towards a more fine-grained analysis of NRT responses in bilinguals, none of them probed into the constituent retention processes that are required to perform the NRT. That was the goal of the current study.

In this study we differentiated between participants' retention of phonemes' identity (phoneme identity performance) and the retention of phonemes' serial order (serial order performance). This was based on the idea that correct nonword repetition depends on the correct retention of the item's constituent phonemes and the serial order of these phonemes. These two types of retention ability can fail independently of each other. Hence, studying these two retention skills separately makes it possible to formulate more precise claims on the processes that underlie a participant's overall NRT performance. Note that, by taking these phoneme-related variables into account, error responses in a NRT also become an important piece of information. Indeed, such responses also provide valuable information, as they can be incorrect with respect to one phoneme-related variable but correct with respect to the other. In other words, the introduction of phoneme-related variables makes it possible to dissect the retention of nonwords into component retention processes, each of which sheds light on a distinct memory process that is required for adequate task performance.

Earlier findings suggest that retention of item information depends on language knowledge stored in LTM, whereas retention of serial order information is related to the quality of a STM device specifically dedicated to the storage of serial order information (Gupta, 2003). We expected bilinguals to have more problems than monolinguals with the retention of the phonemes in a nonword in their L2, as the result of a smaller vocabulary and the fewer possibilities to extract well-defined phonemic representations. In contrast, we did not expect the bilinguals to have problems with the retention of phonemes' serial order as the result of a deficit in their STM serial ordering device (as has been reported for individuals with dyslexia, for instance, see Hachmann, Bogaerts, Szmalec, Woumans, Duyck, & Job, 2014). Indeed, none of the bilingual (or monolingual) children in our study had developmental problems. However, when the serial order of *phonemes* in nonwords must be retained, group differences might also originate in different degrees of phonotactic knowledge, which is stored in LTM. These are likely to result from differences in the familiarity with the combinatorial possibilities of the phonemes in the target language, i.e., the language's phonotactic restrictions. Despite differences between Dutch and Turkish, the bilinguals' native language, we did not expect phonotactic knowledge to cause any differences in the serial retention scores of the two language groups. Indeed, the phonotactic restrictions that characterized the items in our NRT materials also occurred in the bilinguals' native language, making them useful for the bilingual children as a familiar encoding scheme for the phonemes in the nonword stimulus. Hence, we did not predict differences between the two language groups with respect to serial order retention.

At first, we investigated whether we could find the pattern that is typically reported in previous NRT research, that is, an association between NRT performance and prior language knowledge. We found a significant positive correlation between language group and overall nonword level performance ($r = .41$, $p < .01$, see Table 3), i.e., the effect of language group on the total number of nonword responses per participant. The same effect was found when analyzing the data at the level of individual responses, i.e., by means of a GLMM. The probability of a correct response was higher in monolinguals compared to bilinguals. This finding is in line with previous studies, indicating that monolinguals outperform bilinguals on a NRT task in bilinguals' second language (e.g., Windsor et al., 2010; Messer et al., 2010; Kohnert, Windsor, & Yim, 2006; Lanfranchi & Swanson, 2005) and confirms the well-established idea that vocabulary knowledge, which significantly differed between the two groups (see Table 2), is associated with STM performance (e.g., Michas & Henry, 1994). In addition, we found that the number of syllables in a nonword predicted the probability of a correct nonword repetition, an effect that did not differ between groups. This outcome is also consistent with a body of literature, indicating better recall of shorter nonwords as compared to longer nonwords in both monolinguals (Gathercole, 2006) and bilinguals (Girbau & Schwartz, 2008). This replication is not surprising, as it seems obvious that longer nonwords require more STM resources than short ones.

Secondly, we were interested in finding out whether the probability of making no phoneme identity errors, i.e., flawless phoneme identity performance, and the probability of making no serial order errors on the correctly recalled phonemes, i.e., flawless serial order performance, is predicted by participants' language group (monolingual vs. bilingual). The correlation analyses on the basis of participants' total number of correct nonword repetitions revealed a significant correlation between language group and phoneme identity performance ($r = .39$, $p < .0001$, an outcome that is equivalent to an independent samples t-test). In contrast, we found no significant correlation between language group and serial order performance ($r = .15$, $p > .05$). Thus, a rough, exploratory analysis suggests that the ability to retain phonemes' identity seems to be dependent on participants' language background, but the ability to retain phonemes' serial order does not.

However, the fact that there was a strong positive correlation between participants' phoneme identity performance and their serial order performance ($r = .86$, $p < .0001$) makes it difficult to interpret these correlations with language group. To assess the effect of language group on each phoneme-related variable, the collinearity between these two variables must be statistically removed from the variance that language group has to predict. This makes it

possible to assess the unique effect of language group on each phoneme-related variable. Therefore, we opted for GLMM analyses in which one type of phoneme performance (phoneme identity or serial order performance) was included as a covariate when predicting performance on the other phoneme-related variable. For instance, we predicted phoneme identity performance on individual responses from participants' language group and nonword syllable length while statistically controlling for serial order performance (see also Staels & Van den Broeck, 2014). The GLMM in which *phoneme identity performance* was analyzed, revealed an effect of language group after controlling for serial order performance (p = .003). Note that this is a strong effect. Even when statistically controlling for the very strong correlation between the two phoneme-related variables, the effect of language group is still significant. In other words, the difference between monolinguals and bilinguals still significantly predicts the variance in the retention of phoneme identity that cannot be explained by serial order retention. Bilingual children performed at a level below that of their monolingual age peers in three-, four- and five-syllable nonwords with respect to the retention of phonemes. The interaction between language group and syllable length was non-significant. However, the GLMM in which both monolinguals and bilinguals were compared on their *serial order performance* did not yield an effect of language group after removal of the variability on serial order retention that was caused by phoneme identity retention. Bilinguals were as successful as their monolingual peers in retaining the correct serial order of correctly recalled phonemes when repeating a nonword.

Thirdly, we focused on participants' error responses to check whether the distribution of error types (pure phoneme identity errors, pure serial order errors, or a combination of both errors) was similar in both groups. We removed the pure serial order errors from the data, i.e., responses on which only serial order errors were made, as these accounted for a very small minority of the data (about 5% of the responses in both groups). A GLMM assessing the effect of language group on the distribution of the remaining two error types revealed that the distribution of these two main error types was the same for both groups. Importantly, as pure identity errors and combined errors only differ in the absence or presence of a serial order error, respectively, (a phoneme identity error being made in both) the failure to find an effect of language group in these error data is in line with the outcome of the global analysis (including all responses: correct and incorrect ones). Language group does not appear to affect children's serial order retention ability. Despite this convergence between the two analyses, there is a crucial difference between them, besides the different response types that are included in each analysis. The overall analysis investigates the effect of language group after the collinearity between the two phoneme variables has been statistically removed from the variability in the

serial error scores. In contrast, the analysis of the effect of language group on serial order retention in the above set of error responses investigates whether language group can account for the variability in the actual serial order data. Indeed, the constant value for the retention of phonemes' identity (i.e., retention failure) removes the need to statistically control for this type of retention. Hence, the latter analysis shows that the failure to find an effect of language group on serial order performance in the global GLMM was not due to a reduction in the variability of the data after statistical controlling for the correlation between the two phoneme variables.

The analysis also demonstrates that the associative strength between the mechanisms that are responsible for identity retention and for serial order retention is the same in both groups, at least in one direction. More particularly, the conditional probability of making a serial order retention error if a phoneme identity error has been made is the same for both language groups. The chi-square analysis comparing the distribution of pure serial errors and combined errors in both language groups (a GLMM being unfit for this purpose, given the small number of pure serial order errors) suggested that the associative strength between the two error types is the same for both groups in the opposite direction as well. The non-significant interaction between the error type variable and language group in this analysis indicates that the probability of making a phoneme identity retention error if a serial error is made is the same for both groups. Finally, we generalized the rationale of keeping the value on one phoneme variable constant while varying the other. Given the small percentage of pure serial order errors, the only remaining contrast of this type that is amenable to a GLMM analysis involved the correct responses and the pure identity errors (two sizeable response sets in both language groups). These response types share the property that no serial order error has been made. The outcome of this analysis, too, was in line with the analysis on all responses: if no serial order error was made, bilinguals made significantly more identity errors than monolinguals.

In sum, the most important conclusion is that (a) a failure of the retention mechanism for phoneme identity (i.e., phoneme identity performance), operationally defined as a failure to reproduce all phonemes from the nonword stimulus, occurred more often in bilinguals than in monolinguals, and that (b) a failure of the retention mechanism for serial order occurred equally often in both groups. Thus, the results confirm our hypothesis that phoneme identity retention but not serial order retention is a crucial factor for describing the widely reported differences in overall NRT performance between typically developing monolingual and bilingual children (with the proviso that the nonwords' phonotactic constraints are in line with the phonotactic restrictions in the bilinguals' native language).

As this type of analyses in a single NRT task are novel, our results cannot directly be compared with previous research in BILINGUAL children. However, it is interesting to consider our results in the light of the recent conceptualization of item and order retention in current STM models.

Given the observation that the Turkish-Dutch bilingual children in our study showed significantly lower expressive vocabulary skills in Dutch, it might not be surprising that phoneme identity performance was better in monolinguals compared to bilinguals. Indeed, a possible explanation, which is in line with previous research, is that bilinguals are less acquainted with phonemes in their second language, because they know fewer words in that language, and, hence, have fewer opportunities for inducing the phonemes in the language. As a result, they might have more problems to flawlessly segment the non-familiar phoneme input of a nonword. This could be the consequence of less well-developed or less accessible phoneme representations, both possibilities being plausible outcomes of the less frequent exposure to the phonemes of the language from which the nonwords are derived (see also Metsala & Walley, 1998). As a consequence, the presented nonwords would not as readily 'fall apart' into their constituent phonemes as is the case for monolinguals. Although previous STM research did not directly distinguish item and serial order retention in a single NRT task, our results are compatible with the idea proposed by current STM models that storage of ITEM information (i.e., phoneme identity performance in our study) strongly depends on the richness of the language knowledge system (e.g., Cumming, Page, & Norris, 2003; Gupta et al., 2005; Majerus et al., 2008).

In these STM models the retention of SERIAL ORDER information is linked to a specialized STM system, so called 'sequence memory' (Gupta, 2003; Gupta et al., 2005), which makes it possible to store the succession of novel sequences. The existence of such a mechanism would not predict problems associated with the retention of phoneme sequences, i.e., phonemes' serial order, in our typically developing (monolingual and bilingual) children, who have no history of language problems or phonologically processing problems. However, as serial order was defined over phonemes in the current study, participants' knowledge of the ordering possibilities among phonemes may have played a role in the retention of the nonword items. More particularly, LTM knowledge of phonotactic constraints on the phonemic 'items' is likely to have been a determinant of the retention of phonemes' serial order. However, as the nonwords did not violate phonotactic constraints in Turkish (the bilinguals' native language) this would not cause the bilinguals to perform worse on serial order retention than the monolinguals either. Our results supported this hypothesis.

Note that the conclusion with respect to serial order retention is contingent on the convergence of Dutch and Turkish phonotactic constraints in the specific set of nonword items in our experiment. It is quite possible (even plausible) that monolingual children will have a better serial order performance than their bilingual peers when the nonwords do not respect the phonotactic restrictions of the bilinguals' native language. Indeed, it is expected that one's familiarity with the words of a language will not only entail a sensitivity to the phonemes in the language but also to these phonemes' combinatorial possibilities, i.e., the phonotactic regularities in the language, which determine legal phoneme sequences (serial order of phonemes). This sensitivity will increase as vocabulary size increases, which is likely to have a beneficial effect on participants' ability to encode, retain, and recall sequences of phonemes that do not form a real word (i.e., items in the NRT). Further experiments will have to put this hypothesis to the experimental test.

CONCLUSION

The present study examined monolingual and bilingual children's retention abilities for phonemes' identities and their serial order, in an attempt to better understand the mechanisms that underlie the often observed differences in the overall NRT performance in both groups. To this end, we developed measures for the retention of the phonemes themselves and the retention of their serial order. This method has not been used by previous researchers, who always focused on the proportion of correct responses or the proportion of correctly recalled phonemes. In line with previous research, the primary finding confirms that overall NRT performance is influenced by the amount of prior exposure to the target language: the proportion of correct responses was higher for monolinguals than for bilinguals. More importantly, the analyses at the phoneme level shed light on factors that may explain this difference. Bilinguals turned out to be less sensitive to retain phonemes than monolinguals, presumably as a consequence of less familiarity with the words in the target language, which we inferred from bilinguals' significantly smaller expressive vocabulary knowledge. A smaller vocabulary size is likely to cause more poorly developed or less accessible phoneme representations, which are required to adequately encode novel phoneme strings. In addition, bilinguals were equally sensitive to phonemes' serial order as monolinguals when repeating nonwords. The latter finding, being a null effect, does not by itself confirm the presence of a special sequence memory in current STM-models, but is compatible with such a mechanism. Indeed, as all children developed

normally, no problems with a sequence memory in STM were expected in the monolingual and bilingual groups. However, as the retention of nonwords integrates the retention of item and serial order information, serial order information for phonemes is also likely related to participants' familiarity with the way in which phonemes can be serially ordered in the target language, i.e., to LTM language knowledge. In contrast to previous studies, serial order memory was not tested by having participants recall the random order of familiar items but by having them recall the non-random, i.e., phonotactically constrained, order of (familiar) items, i.e., phonemes. Despite the differences between Dutch and Turkish at the phonological level, the phonotactic principles of both languages happened to coincide in our set of nonwords, which is likely to account for the finding that bilinguals did not perform worse on the serial order measure than monolinguals.

In sum, the findings of the present study enhance our understanding of mechanisms that underlie inferior NRT performance in bilinguals' L2. No doubt, more research needs to be conducted on the relationship between NRT performance and L2 language exposure. Based on the current findings, some recommendations for future studies can be made. For instance, it would be interesting to study the supportive role of phonotactic knowledge in the retention of serial order in children's L2. This can be achieved by designing a NRT that consists of nonwords whose phonotactic patterns are dominant in bilinguals' second language but are absent in their first language.

REFERENCE LIST

Bates, D., & Maechler, M. (2009). *lme4: Linear mixed-effects models using S4 classes* [Computer software manual]. Available: http://CRAN.R-project.org/package=lme4.

Bosman, A., de Graaff, S., & Gijsel, M. (2006). Double Dutch: The Dutch Spelling System and Learning to Spell in Dutch. In R.M. Joshi & P.G. Aron (Eds.), *Handbook of orthography and literacy* (pp. 135-150). Mahwag, NJ: Lawrence Erlbaum.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review, 106,* 551–581.

Burgess, N. & Hitch, G.J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. Journal of Memory and Language, 55, 627-652.

Cheung, H. (1996). Nonword span as a unique predictor of second language vocabulary learning. *Developmental Psychology, 32,* 867–873.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

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.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences, 3rd edition.* Lawrence Erlbaum Associates.

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.

Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory, 11,* 43-63.

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.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. J. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology, 45A,* 1-30.

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

Gathercole, S. E. (2006). Complexities and constraints in nonword repetition and word learning. *Applied Psycholinguistics, 27,* 599-613.

Girbau, D., & Schwartz, R. G. (2008). Phonological working memory in Spanish-English bilingual children with and without Specific Language Impairment. *Journal of Communication Disorders, 41,* 124-145.

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.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P-H. (2005). Serial Position Effects in Nonword Repetition. *Journal of Memory and Language, 53,* 141-162.

Gupta, P. (2009). A computational model of nonword repetition, immediate serial recall, and nonword learning. In A. Thorn & M. Page (eds.), *Interactions between short-term and long-term memory in the verbal domain*, pp. 108–135. Hove, England: Psychology Press.

Gutiérrez-Clellen, V., & Simon-Cereijido, G. (2010). Using Nonword Repetition Tasks for the Identification of Language Impairment in Spanish-English-Speaking Children: Does the Language of Assessment Matter? *Learning Disabilities Research & Practice, 25,* 48-58.

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.

Hitch, G.J., Fastame, M.C., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory, 3-4,* 247-258.

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

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-processing tasks separate children with language impairment from typical bilinguals? *Learning Disabilities Research & Practice, 21,* 19–29.

Lanfranchi, S., & Swanson, H.L. (2005). Short-term memory and working memory in children as a function of language-specific knowledge in English and Spanish. *Learning and Individual Differences, 15,* 299-319.

Lee, S.A.S., & Gorman, B.K. (2012). Nonword repetition performance and related factors in children representing four linguistic groups. *International Journal of Bilingualism, 17,* 479-495.

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short- term memory for serial order and item information. *Journal of Experimental Child Psychology, 93,* 95-119.

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.

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A.L., Martinez, T., Weekes, B., & Marquet, P. (2008). Neural networks for short-term memory for order differentiate high and low proficiency bilinguals. *Neuroimage, 49,* 1698-1713.

Masoura, E. & Gathercole, S. E. (1999). Phonological short-term memory and foreign vocabulary learning. *International Journal of Psychology, 34,* 383-388.

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

McLeod, S. (Ed). (2007). *The international guide to speech acquisition.* Clifton Park, NY: Thomson Delmar Learning.

Messer, M.H., Leseman, P.P.M., Mayo. A.Y., & Boom, J. (2010). Phonotactic probability effects in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology, 105,* 306-323.

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: Lawrence Erlbaum Associates.

Michas, I.C., & Henry, L.A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology, 12,* 147–163.

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

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.

Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50,* 113-133.

Öney, B. & Durgunoglu, A.Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics, 18,* 1-15.

Page, M.P.A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761-781.

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology, 50,* 408-412.

Scheltinga, F. (2003). *The Dutch Nonword Repetition Test.* University of Amsterdam: Unpublished manuscript.

Schraeyen, K., Geudens, A., Ghesquire, P., Van der Elst, W., & Sandra, D. (2017). Poor performance on the retention of phonemes' serial order in short-term memory reflects young children's poor reading skills. *The Mental Lexicon, 12 (1),* 129-158.

Semel, E., Wiig, E.H., & Secord, W.A. (2008). *Clinical Evaluation of Language Fundamentals (CELF-4NL): handleiding.* Amsterdam: Pearson.

Sorenson Duncan, T., & Paradis, J. (2016). English Language Learners' Nonword Repetition Performance: The Influence of Age, L2 Vocabulary Size, Length of L2 Exposure, and L1 Phonology. *Journal of speech, language, and hearing research, 59,* 39-48.

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.

Summers, C., Bohman, T.M., Gillam, R.B., Pena, E.D., & Bedore, L.M. (2010). Bilingual performance on NWR in Spanish and English. *International Journal of Language and Communication Disorders, 45,* 480-493.

Thordardottir, E., Kehayia, E., Lessard, N., Sutton, A. & Trudeau, N. (2010). Typical performance on tests of language knowledge and language processing of French-speaking 5-year-olds. *Canadian Journal of Speech Language Pathology and Audiology, 34,* 5-16.

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46,* 1-16.

Thorn, A.S.C., & Gathercole, S.E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *Quarterly Journal of Experimental Psychology, 52A,* 303–324.

Topbaş, S., & Kopkalli-Yavuz, H. (2008). Reviewing sonority for word-final sonorant + obstruent consonant cluster development in Turkish. *Clinical Linguistics and Phonetics, 22,* 871–880.

Topbaş, S., Kaçar-Kütükçü, D., Kopkalli-Yavuz, H. (2014). Performance of children on the Turkish Nonword Repetition Test: Effect of word similarity, word length, and scoring. Clinical linguistics & phonetics, 28, 602-16.

Windsor, J., Kohnert, K., Lobitz, K., & Pham, G. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology,19,* 298-310.

CHAPTER 5

GENERAL CONCLUSIONS

The central aim of this doctoral thesis was to examine verbal STM as measured by nonword repetition in monolingual good and poor readers and in typically developing sequential bilingual children learning to read in a second language. Poor verbal short-term memory span is, next to poor phoneme awareness and slow lexical retrieval, one of the most consistently reported areas of (cognitive) difficulty for children and adults with dyslexia (Snowling, 2000). Apart from the rather consistent finding of lower overall nonword repetition performance in impaired readers, so far, little was known about underlying factors of this underperformance. By separating the performance on two types of phoneme-related retention in the analysis of NRT data, i.e., retention of phonemes' identity (item STM retention) and retention of phonemes' serial order (order STM retention), we attempted to find out how these two types of retention in verbal STM processing impact upon reading (dis)ability.

To this end, we carried out three independent studies using the NRT. In study 1 (Chapter 2) and study 2 (Chapter 3), we compared item and serial order STM retention performance when participants had to recall nonwords conforming to the phonotactics of their native language. Study 1 focused on adults with and without dyslexia. Study 2 investigated children who varied in literacy skills. In study 3 (Chapter 4), we extended this line of research by comparing monolingual and sequential bilingual children with adequate literacy skills in their second language. Thus, we investigated the role of lexical knowledge on both types of STM retention.

SUMMARY OF FINDINGS

In this paragraph, we will briefly summarize the effects observed in each study. In the next paragraphs we will discuss these findings in more detail.

Study 1 (Chapter 2) showed that the retention of phoneme identity was unaffected in adults with dyslexia. In contrast, we provided evidence for decreased serial order retention in adults with dyslexia compared to adults without dyslexia. In other words, adults with dyslexia were less likely to recall a sequence of phonemes in their correct order.

Study 2 (Chapter 3) showed that children's literacy level predicted their performance in the retention of serial order but not their performance in the retention of phoneme identity. More particularly, better literacy skills were associated with a higher probability of repeating a nonword without serial order errors.

The data presented in study 3 (Chapter 4) supports the idea that the amount of prior exposure to the target language plays a role in nonword repetition. Monolingual children outperformed sequential bilingual children on overall NRT performance when the items conformed to the phonotactics of the former group's native and the latter group's non-native language. Sequential bilinguals turned out to be less sensitive to the retention of phoneme identity compared to monolinguals. However, they were equally sensitive to the retention of serial order.

In this final chapter, we further discuss our findings, relate them to the literature on verbal STM and literacy, and outline their theoretical implications. At the end of this chapter, some methodological limitations and future research directions will be discussed.

THE ROLE OF ITEM AND SERIAL ORDER RETENTION IN LEARNING TO READ

It is quite easy to understand why verbal STM, and more precisely item and serial order STM retention, might play a crucial role in learning to read. Verbal STM refers to the *storage* and *retrieval* of phonological representations in STM. When a child plans to sound out a word, it can use STM as a short-term storage device for the ordered sequence of phonemes matching the word's letters. Two types of information must be represented: the words' individual phonemes (i.e., item STM retention) and their relative position in the phonological chain that matches the printed letter sequence (i.e., serial order STM retention). Apart from subserving the reading process itself, verbal STM is also crucial for overall word learning, as it mediates the storage of phonological information in LTM (Gathercole, Alloway, Willis, & Adams, 2006).

Previous verbal STM studies consistently showed lower performance in a NRT task for individuals with poor reading skills in their native language. This conclusion was mainly based on whole item analyses (correct repetition or not) or phoneme analyses (percentage of phoneme errors) (e.g., Gathercole, 1994; Robertson & Joanisse, 2010). Note that these phoneme analyses were considerably more coarse-grained than the analyses of phoneme identity and serial order retention in the studies reported in this doctoral thesis. Unlike the latter two measures, percentages of phoneme errors shed little light on the underlying factor(s) that are related to poor nonword repetition. Whereas analyses of phoneme errors highlight different error types, they do not characterize them in terms of STM properties.

Nevertheless, in the past decade, there has been an increasing research interest for these STM properties, i.e., the distinction between memory for item information (item STM) and memory for serial order information (serial order STM) in the domains of literacy acquisition and developmental dyslexia. As described in Chapter 1, recent computational models of verbal STM (e.g., Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999, 2006; Page & Norris, 1998) distinguish between these two processes.

Using different types of tasks and different stimuli (verbal and nonverbal), several recent behavioral studies showed a serial order impairment in individuals with dyslexia, both in adults (e.g., Martinez Perez, Majerus, & Poncelet, 2013; Hachmann et al., 2014; Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015; but see Wang, Xuan, & Jarrold, 2016) and in children (e.g., Martinez Perez, Majerus, Mahot, & Poncelet, 2012b; but see Staels & Van den Broeck, 2014). In summary, these studies suggest that both the recognition of serial order (as measured with judging tasks, in which participants have to decide whether the order of two lists with (non)verbal item is the same) and the active recall of serial order (as measured with tasks in which participants have to recall an ordered list) are hampered in individuals with dyslexia. However, these reported studies were not conclusive regarding item STM retention. Whereas some studies did find an impaired item STM for individuals with dyslexia (e.g., Martinez Perez et al. 2012b, 2013; Staels & Van den Broeck, 2014), others did not (e.g., Hachmann et al., 2014; Bogaerts et al., 2015; Wang et al., 2016).

Although these studies are valuable to further our understanding of STM processing and the problems with verbal STM confronting individuals with a reading disability, none of them focused on item and serial order processing *at the phoneme level*. In our view, this level is crucial, given the massive evidence that individuals with dyslexia experience a basic problem at the level of phonemic representations (Boada & Pennington, 2006; Carroll & Snowling, 2004; Elbro, 1996; Elbro, BorstrØm, & Petersen, 1998; Ramus & Szenkovits, 2008). Another

problem is that previous studies examined the two types of STM retention with different sets of items. In other words, differences between item and serial order performance may not only have been affected by the nature of the different memory tasks but also by the materials themselves. To avoid this problem, we studied both types of STM retention with a single item set. Using this STM distinction, we attempted to identify the factor(s) characterizing the overall performance patterns in the studied target groups. The methods and techniques we used to analyze the data are summarized in the next section.

MEASURING ITEM AND SERIAL ORDER STM PROCESSING WITH THE NRT

We decided that the NRT was the ideal instrument for simultaneously studying verbal STM retention for item and serial order information at the level of phonemes. Indeed, accurate memory for both the identity of phonemes and the serial order of these phonemes in the nonword input are key for overall NRT success.

In all studies, we used binary scores (correct vs. error) to encode participants' individual responses at two levels: the entire nonword (nonword level) and the nonword's constituent phonemes (phoneme level). Within the phoneme level a distinction was made between the constituent phonemes themselves (i.e., item STM) and the order of the correctly recalled phonemes (i.e., serial order STM). Thus, three different binary scores were computed: one at the nonword level (overall response score) and two at the phoneme level (phoneme identity and phoneme serial order scores). By calculating the *phoneme level* scores in a binary way, we wanted to know how likely it was that participants could recall *all* phonemes (not necessarily in their correct order) and the serial order of *all* correctly recalled phonemes (as serial order performance only makes sense for correctly recalled phonemes). Note that this particular formulation of the two questions at the phoneme level is already implied by the formulation of the question at the nonword level: how likely is it that participants can recall all phonemes, together with all serial order relations among these phonemes (i.e., the whole nonword)? In other words, both questions at the phoneme level result from splitting up the question that is raised in traditional NRT studies.

In each experimental study, we thus had three dependent variables: nonword level performance, phoneme identity performance and phoneme serial order performance. We used two statistical techniques to find out whether the participants differed from each other with

respect to these three dependent variables: (a) simple Pearson correlations between participants' reading scores and their scores on each of the three dependent variables (exploratory analysis); and (b) generalized linear mixed model analyses in which participants' reading group (study 1) or reading level (study 2) or language group (study 3) was used to predict their performance on each of the three variables (in-depth analysis). As the two phoneme retention measures, i.e. phoneme identity and serial order were not statistically independent (i.e., there was a significant correlation between them in all studies), the mixed model that was built to predict performance on a phoneme-related variable included the other phoneme-related variable as a covariate. This guaranteed that the effect of reading performance on these dependent variables was uncontaminated by their correlation (see also Verbeke & Molenberghs, 2000).

ITEM AND SERIAL ORDER STM PROCESSING IN RELATION TO READING

First, we investigated the effect of literacy on the retention probability of the entire nonword. This was necessary to show that we could replicate previously reported effects, i.e., to validate our method. In line with these studies at the nonword level, we showed problems with verbal STM in adults with dyslexia (study 1) and less skilled young readers (study 2) as compared to typical readers. Similarly, in study 3 we showed inferior overall NRT performance in typically developing sequential bilinguals compared to typically developing monolinguals, i.e., individuals with a larger vocabulary in the language of the NRT than the former group.

Next, and most importantly, we studied the effect of literacy on the retention probability of all phonemes in the nonword (with or without serial order errors) (item STM retention), and on the retention probability of all (relative) serial positions of the correctly recalled phonemes (order STM retention).

STM EFFECTS IN CHILDREN'S AND ADULTS' L1

The consistent co-occurrence between problems at the level of *serial order retention* (order STM retention) and poor reading adults with dyslexia and less skilled readers, discussed in Chapters 2 and 3, extends most studies that focused on item and serial order processing in STM to the level of phonemes (with adults: e.g., Hachmann et al., 2014; Martinez Perez et al., 2013;

Bogaerts et al., 2015; with children: e.g., Martinez Perez et al., 2012a, 2012b; Staels & Van den Broeck, 2014).

In contrast, *phoneme identity retention* (item STM retention) turned out to be unaffected by individuals' literacy skills, both in adults and in young readers. As mentioned above, previous research on item STM performance yielded mixed results. Whereas Hachmann et al. (2014) and Bogaerts et al. (2015) also failed to find item retention problems for individuals with dyslexia, both Martinez Perez et al. (2013) and Staels and Van den Broeck (2014) did show an item retention effect. One plausible explanation for this inconsistent finding is the difference in the item retention tasks that were used. As mentioned in Chapter 3, the Martinez Perez et al. studies and Staels and Van den Broeck study used a single nonword delayed repetition task, in which participants were asked to repeat a one-syllable nonword after a 8-second filled delay. Although these authors claimed to rule out serial order in their item task by using one syllable CVC-nonwords, serial order can never be ruled out when asking participants to repeat a sequence of phonemes that form a nonword, at least when the score is based on whole item performance (see also Hachmann et al., 2014).

Nevertheless, our consistent failure to find a relationship between literacy skills and the *retention of item information* (*phoneme identity*) comes as a surprise, as the items in our studies were individual phonemes. Given the strong emphasis on the low quality of poor readers' phoneme representations in the phonological deficit hypothesis, one would expect such readers to have problems with the retention of the phonemes in a nonword. However, there is a strong reason to believe that our null effects are Type II errors, i.e., a failure to reject a false null hypothesis: the strong correlation between phoneme identity performance and serial order performance in both studies (.69 with children; .76 with adults). As mentioned previously and clearly described in Chapters 2 and 3, in all mixed model analyses, we included serial order performance as a covariate of phoneme identity performance (and vice versa) to remove the variance in the dependent variable resulting from the strong correlation between these two factors. Thus, we enabled the model to calculate an uncontaminated estimate of the effect of literacy on the dependent variable, i.e., phoneme identity performance or serial order performance. Obviously, taking such a strong correlation into account made it much more difficult to detect an effect of literacy on the dependent variable. As a considerable part of the variance was removed from the data by this statistical control of the correlation (the R square values in Chapters 2 and 3 were .48 and .58, respectively), there was less variance to be accounted for by the literacy factor. Hence, it is reasonable that an effect of literacy on phoneme identity performance would have emerged if the correlation had been smaller. The fact that this

effect was associated with relatively small p-values (p = .12 in both experiments) makes this even quite plausible. Note that this correlation obviously hampered the detection of an effect of literacy on both phoneme-related variables to the same extent. The fact that, despite this, such an effect was found on serial order retention and not on identity retention, both in the adult study and in the children study, underscores the importance of serial order retention capacity for reading.

In sum, the additional value of our experiments described in Chapters 2 and 3 involves the evidence that a failure of the retention mechanism for serial order at the phoneme level occurred more often in less skilled monolingual young readers and adults with dyslexia compared to literate young and adult controls.

STM EFFECTS IN CHILDREN'S L2

Our experimental work presented in Chapter 4 provides more insight into sequential bilinguals' lower NRT performance compared to monolinguals. Whereas a failure in serial order retention occurred equally often in both groups, a failure in phoneme identity retention occurred more often in sequential bilinguals compared to monolinguals. This observation not only informs us on the problem facing sequential bilinguals but also sheds light on the nature of verbal STM. It confirms the idea that item information is processed differently than serial order information in STM. More specifically, sequential bilinguals' lower expressive vocabulary skills combined with their lower performance on item retention indicates that item retention in STM depends on access to the linguistic system in LTM whereas serial order is related to a distinct processing system in STM (e.g., Nairne & Kelley, 2004). Previous research already showed that children's ability to imitate unfamiliar strings of phonemes that are phonotactically legal but lack semantics is related to vocabulary size (Gathercole, 2006). Based on our results presented in Chapter 4, we concluded that the presence of fewer consolidated LTM representations for words in typically developing sequential bilinguals' second language (compared to monolingual controls), i.e., as measured by their lower expressive vocabulary skills, influence their item STM processing but not their serial order processing at the phoneme level. As pointed out in Chapter 4, the null effect of participants' language group (monolingual versus sequential bilingual) on serial order retention is compatible with this account. As the NRT items in our sequential bilinguals' second language respected the phonotactic regularities in their native language, typically developing readers in this language would not be expected to encounter problems in serial order performance. Indeed, our experiments with monolinguals (Chapters 2

and 3) strongly suggest that serial order problems in verbal STM should not be worse in sequential bilingual than in monolingual readers if the process of learning to read develops typically in both groups.

INTEGRATIVE ACCOUNTS OF SERIAL ORDER PROCESSING PROBLEMS IN LESS SKILLED READERS

In our experiments in Chapters 2 and 3 we clearly showed that serial order retention is a crucial factor for giving a better description of the widely reported differences in overall NRT performance between less skilled and skilled readers. Apart from this (descriptive) contribution to current literacy research, and, more specifically, NRT research, our findings furthermore enabled us to consider different theoretical accounts regarding problems with STM processing in individuals with a reading impairment.

STORAGE-BASED PERSPECTIVES

THE ROLE OF PHONOTACTIC KNOWLEDGE

Let us first consider the role of phonotactics in nonword repetition and serial order processing. Phonotactic probability refers to the likelihood with which phonemes (or phonemes types, e.g., fricatives and plosives) follow each other in real words of a language (Jusczyk, Luce, & Charles-Luce, 1994). The notion that phonotactic probability is a characteristic of phonological representations was already highlighted by Vitevitch, Luce, Charles-Luce, and Kemmerer (1997) two decades ago: *"It seems certain that information regarding the probability — and not simply the legality or illegality — of a given phonetic sequence is, in one form or another, represented in memory."* (p. 60). Earlier work had indeed shown that phonotactic probability affects NRT performance (Jusczyk et al., 1994). One example is that more accurate repetition performance is found for nonwords with a high phonotactic probability compared to nonwords with a low phonotactic probability (Edwards, Beckman, & Munson, 2004). Since phonotactic information can only have been stored in LTM, this finding also directly underlines the impact of LTM on STM processing, i.e., more support from LTM is available for high probability items than for low probability items.

One line of reasoning is that less skilled readers in their native language, like the ones studied in Chapter 2 and Chapter 3, have poorer knowledge of phonotactic regularities, possibly

as a result of problems with statistical learning[15] (see for instance Kahta & Schiff, 2016). Hence, they could not make use of this support (stored in LTM) as much as typically developing readers. In other words, this lack of LTM support might be the cause of inferior storage in STM and retrieval from it, like phonemes' serial order in a NRT. We acknowledge, however, that we cannot make a decisive statement regarding this issue, as we did not manipulate the nonwords' phonotactic probability (see also further). Note also that the relationship between poor reading and statistical learning is far from clear. In a recent systematic review that aimed to find convincing evidence for a group difference in statistical learning between participants with and without dyslexia, the authors concluded that "the existing evidence does not allow us to draw strong conclusions" (Schmalz, Altoè, & Mulatti, 2017, p. 158). Based on a meta-analysis of the literature on this topic, they go on to say that this is "due to the presence of publication bias and questionable research practices".

The above reasoning raises a question when turning to our bilingual study described in Chapter 4. Is it compatible with the potential role of phonotactic knowledge on these children's STM performance? We have arguments to believe it does. The mental lexicon (i.e., LTM knowledge) of the sequential bilingual Turkish-Dutch speaking children's consisted of Turkish and Dutch lexical representations, although the latter representations were smaller in number compared to the monolingual children (i.e., smaller vocabulary in Dutch). In contrast to the groups of less skilled readers in Chapter 2 and Chapter 3, this group of typically developing sequential bilinguals does not give us reasons to question their ability to extract phonotactic features from LTM. If reading skill correlates with the ability to induce the phonotactic regularities in a language, these readers should experience no problems in learning these regularities. Our results indeed showed that they performed equally well on serial order retention. However, one might object that there is another reason why children learning to read in a second language may have less phonotactic knowledge than a control group of native speakers: vocabulary size. In contrast to the adults with dyslexia and the 10-year old children learning to read in their native language that we studied in Chapters 2 and 3, respectively, typically developing young sequential bilinguals who learn to read in their second language may have a vocabulary that is too small to support sufficient phonotactic knowledge. Hence, despite the fact that these bilinguals were not expected to have problems with the STM retention

[15] Theoretically, poorer phonotactic knowledge might also be due to the typically smaller vocabulary size of poor readers. However, it seems unlikely that our adults with dyslexia and our poor reading children did not know enough words to support adequate phonotactic knowledge of Dutch.

of serial order information as the result of problems with a sequencing process in STM, it might still come as a surprise that their vocabulary size did not cause worse serial order performance compared to the group of age-matched native speakers. However, one property of our NRT items can explain why no difference was observed. The nonwords we used in our experiments were compatible with phonotactic regularities in both Dutch and Turkish. The point we want to make here is that a difference in serial order retention between sequential bilingual and monolingual children might have occurred in case the nonwords had violated the phonotactic regularities in Turkish, i.e., bilinguals' native language. In the latter case, the serial order 'problem' would not have been due to poorer phonotactic knowledge as the result of poor literacy skills, but as the result of less exposure to sequential bilinguals' second language phonotactic regularities.

In sum, the observed serial order effect in adults with dyslexia (study 1) and poor readers (study 2) may be explained in terms of higher-order phonological units, i.e., phonotactic sequences that are retrieved from LTM for encoding and storage in STM.

THE ROLE OF ORTHOGRAPHIC KNOWLEDGE

Apart from the potential role of phonotactics in nonword repetition, more specifically in serial order retention, another storage-based interpretation could be suggested, i.e., the role of orthographic representations in NRT performance. Orthographic representations refer to the abstract mental representations of the written forms of words. Several studies have shown that children and adults with a reading impairment have problems with the formation of orthographic representations (e.g., Suárez-Coalla, Avdyli, & Cuetos, 2014). According to the self-teaching hypothesis (Share, 1995,1999), orthographic representations are developed through repeated accurate reading. If so, it follows that fewer and/or less adequate orthographic representations will be available in disabled readers' LTM. However, the main question remains how these adequate (in good readers) or less adequate (in poor readers) orthographic representations affects phonological processing in STM, i.e., when performing a NRT task. As highlighted in Chapter 2, previous studies showed that performance on the NRT is indeed influenced by orthographic factors (Nation & Hulme, 2011). One potential line of reasoning is that orthographic information is automatically activated when listening to speech, even when it is not needed to conduct a task such as a NRT task (Seidenberg & Tanenhaus, 1979). This automatic activation of orthographic representations could facilitate participants' memory for short-lived phonemic segments in the NRT (Geudens, 2006). If poor readers have less adequate

orthographic representations compared to better readers, it follows that they will benefit less from this knowledge base. As a result, orthographic knowledge will provide weaker support for their NRT performance.

Another idea is that the nature of children's phonological representations is changed as a result of learning to read (see also the discussion whether serial order problems cause reading problems or vice versa, i.e., the cause vs. effect discussion). In this view, children are thought to integrate orthographic and phonological information, which causes a restructuring of their phonological representations (Perre, Midgley, & Ziegler, 2009). Better readers' phonological representations will be restructured more effectively than poor readers' phonological representations. Although the potential role of orthographic knowledge on NRT performance is worth considering, this factor does not explain the discrepancy we observed between phoneme identity and serial order. Given the role of serial order one could suggest that individuals with dyslexia are worse than controls in recalling the order of the letters in the orthographic representation of the auditorily presented nonword.

THE ROLE OF A DOMAIN-GENERAL SEQUENCE MECHANISM

Another possible account of our findings is the idea that there is a specific sequence memory mechanism in STM that is not influenced by other memory systems such as LTM. Several recently published studies focusing on the link between serial order processing and literacy skills subscribe to the idea of a domain-general impairment in individuals with dyslexia (e.g., Bogaerts et al., 2015; Martinez Perez et al., 2012a; 2012b; Hachmann et al., 2014). According to these studies, a domain-general mechanism in STM is responsible for retaining the order of units of any kind (hence, also phonemes). Our consistent finding in Chapter 2 and Chapter 3 of serial order retention problems in impaired or less skilled readers, then, could be explained by a basic difference between these readers' sequence memory mechanism in STM and that of controls. The fact that serial order impairments have also been detected in individuals with dyslexia with tasks in which no support of LTM could play a role (e.g., when using audio-tactile sequences, see Laasonen et al., 2012) could be seen as support for a domain-general deficit in STM. In their study, Laasonen et al. (2012) examined whether poor STM in a group of adults with dyslexia influenced their processing of sensory stimulus sequences in addition to phonological material. They used binary nonverbal stimuli such as light flashes, tone bursts, finger touches, and cross-modal combinations. They showed an underperformance in order

retention in sensory STM compared to controls, indicating a general difficulty in the representation of temporal sequences in STM.

The fact that similar problems were also found in other cognitive areas (e.g., in the domain of dyscalculia, see Attout & Majerus, 2015) could suggest that learning difficulties in multiple cognitive areas might be due to serial order STM problems (Jaroslawska, Gathercole, Logie, & Holmes, 2016), possibly areas in which the retention of serial order is crucial to performance. Given the evidence provided in these behavioral studies, it is tempting to interpret our findings in Chapter 2 and Chapter 3 along the same lines. In this view, our results would extend the evidence in support of this domain-general serial order effect by linking this STM mechanism to phonemes.

At first sight, this view seems to ignore 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, due to their problems with the development of accurate phoneme representations. However, the poor functioning of a domain-general STM mechanism in less skilled readers and individuals with dyslexia obviously does not exclude that these individuals might also have problems with the development of high-quality phonemic representations. Indeed, dyslexia is likely to involve multiple cognitive mechanisms rather than stemming from a deficit in a single cognitive mechanism (Van Bergen, van der Leij, & de Jong, 2014).

A RETRIEVAL-BASED PERSPECTIVE

As extensively discussed in Chapter 2, another view that has received some attention holds that the problem in individuals with a reading disability is not situated at the level of storage but at the level of retrieval (e.g., Ramus & Szenkovits, 2008; Boets et al., 2013). In this view, impaired readers have intact phonological representations in STM. However, the access to these phonological representations is hampered at the retrieval stage. Following this view, our results obtained in Chapter 2 and Chapter 3 would imply that less skilled readers face problems with the order in which they have to retrieve the phonological information from the nonword's representation in STM, not with the temporary storage of the nonword itself. As explained in Chapter 2, in this case, NRT performance does not reflect the underlying representation of a nonword but the process of rebuilding this representation by serially accessing its phonemes or phonotactic units in a temporary STM representation.

Our studies cannot inform us whether the locus of the effects is the storage or the retrieval level. As NRT responses involve the retrieval of a stored representation of a spoken

nonword in STM, they reflect both stages at the same time. Hence, the reader should be aware that the nature of the NRT makes our findings compatible with a problem at either (or both) of these loci.

CONCLUSION

Taken together, our consisting finding of serial order retention problems in individuals with poor reading skills or individuals with dyslexia when performing a NRT task could support a view that emphasizes the importance of phonological/phonotactic and/or orthographic knowledge or the access to it but it could also be explained in terms of a potential domain-general problem. Importantly, this latter domain-general interpretation could also account for the interpretation in terms of phonotactic knowledge we suggested above. Note again that we do not subscribe to the view that a domain-general *deficit* in individuals with dyslexia is a unique cause of dyslexia. Given the extensive body of evidence favoring low quality phonological representations in these groups as well as recent multiple cognitive deficit models for understanding dyslexia (see Pennington, 2006), we doubt the idea that there is a single cause of dyslexia, as proposed by some researchers (e.g., Bogaerts et al., 2015). In the next subsection, we zoom in on the widely accepted phonological deficit hypothesis.

WHAT ABOUT THE PHONOLOGICAL DEFICIT HYPOTHESIS?

According to the phonological deficit hypothesis, the most widely held hypothesis on the basic problem underlying dyslexia, the impairment in letter-sound conversion in persons with dyslexia stems from a deficit in the processing of phonemes, i.e., the elementary constituents of spoken words (Ramus, 2003). Our experiments, however, revealed a null-effect for the retention of phonemes' identity in STM. How can poor readers recall a nonword's phonemes equally well as their better reading peers if the phonological deficit hypothesis is correct? In the empirical chapters of this doctoral thesis, we pointed out that this null effect is likely due to a statistical issue (see also above in this chapter). The two phoneme-related dependent variables – participants' recall of the nonword's phonemes and their serial order – showed relatively strong correlations in all studies. As the effect of literacy skills on each phoneme-related variable should explain variance on this variable that could not be the result of the other variable, this shared variance had to be statistically removed from the data. As a result, the

variance that was left to be explained by a group difference (individuals with dyslexia versus controls) or by differences in reading skill was considerably reduced. The fact that we still found a clear effect of literacy skills on serial order retention in all studies involving poor readers underscores the importance of being able to retain the serial order of phonemes for reading. In contrast, the fact that we failed to find an effect of literacy skills on the retention of phoneme identity does not cast doubt on the phonological deficit hypothesis. After statistically controlling for the correlation between the retention of identity and serial order, there may have been too little variance on this dependent variable that could be uniquely explained by differences in literacy skill. Hence, our results should not be construed as evidence disproving the phonological deficit hypothesis and exclusively favoring the hypothesis that poor and disabled readers have a problem with the storage or retrieval of serial order information. Individuals with poor reading skills or dyslexia might suffer from both problems simultaneously.

CLINICAL IMPLICATIONS AND PRACTICAL RELEVANCE

Diagnostic tools for assessing individuals with dyslexia or less skilled readers usually focus on performance on a range of skills, with the purpose of making a profile of an individual's strengths and weaknesses. Apart from one's achievement on reading (and spelling in the case of dyslexia assessment) and oral language skills, important underlying cognitive skills such as phonological awareness, STM and LTM memory, and speed of phonological processing receive attention too. The findings discussed in this doctoral thesis may therefore have clinical implications for future diagnostic and remediation purposes.

First, we acquired new insights regarding underlying influencing factors of the widely reported underperformance on nonword repetition in less skilled readers and sequential bilingual typically developing children learning to read in a second language. Our outcomes clearly indicate that a more detailed analysis of serial order processing at the phoneme level, e.g., when conducting a NRT task, could be helpful in clinical contexts. The NRT might be used in conjunction with a verbal (but non-phonemically oriented, e.g., Martinez Perez et al., 2012) and a nonverbal sequence memory task (e.g., Hachmann et al., 2014). Thus, one could derive an overall performance profile of the poor reader at the level of STM serial order retention.

Second, if reduced serial order processing abilities are causally involved in the development of dyslexia, training of these serial order processing skills in children at risk for reading impairment may improve their later reading skills. The extent to which focusing on STM serial order processing in the context of remediation can effectively lead to improved reading outcomes is, of course, crucially dependent on whether such a causal link exists between poor serial order processing and reading problems.

First, future fundamental studies are necessary to determine whether the observed serial order effects are (one of) the cause(s) rather than a consequence of reading impairment. Indeed, linear mixed effects models can only reveal correlations between the outcome variable and an independent variable – and a correlation does not inform us on what is the cause and what is the effect (there may even be a 'tertium quid', see Bryant & Bradley, 1985, for an excellent review of this type of interpretation problems in the dyslexia literature). More concretely, reading inevitably requires attention to serial order information (the order of graphemes and, hence, their corresponding phonemes). As reading practice involves an implicit training in the ability to keep track of serial order information, poor readers' worse performance on serial order STM performance may be the consequence of their reading problem rather than its cause.

And even then fundamental research on the causal nature of poor serial order processing is not sufficient. After all, previous research has shown that an important association between a (predictive) cognitive factor and reading (development) not necessarily implies that training this cognitive skill is also effective in young children (see for instance, de Jong & Vrielink, 2004) as described in the next section. Thus, training studies are needed before moving to concrete proposals for the role of serial order processing in the prevention and remediation of reading problems. This study clearly shows the importance of this line of applied research.

METHODOLOGICAL LIMITATIONS

In all our studies, we used a cross-sectional study design enabling us to show an *association* between reading (dis)ability and serial order performance. As mentioned in the previous paragraph, such a design makes it impossible to make out whether poor readers' worse performance on phonemes' serial order is one of the plausible *causes* of their reading problem or its *effect.* In a recent opinion paper, Huettig, Lachmann, Reis and Petersson (2017) pushed this cause or effect discussion forward again. Overall, they argue that the large range of impairments displayed by individuals with dyslexia are also shown in illiterate individuals who

had received no or very little reading instruction. In their view, this might indicate that the extensively described performance differences at the cognitive level (e.g., phonological processing skills) should be seen as secondary consequences of suboptimal or reduced reading experience (see also Vellutino, Fletcher, Snowling, & Scanlon, 2004), rather than a cause of reading disability. Focusing on verbal STM and more specifically performance in a NRT task, Huettig et al. (2017) claim that lower NRT performance in less skilled readers, both individuals with dyslexia and illiterate adults, results from a lack of reading experience which in turn gives rise to less fine-grained phonological representations. They do not deny the evidence showing that NRT accuracy in preliterate children predicts later literacy skills (as for instance shown by Gathercole & Baddeley, 1993). However, they argue that preliterate differences in NRT performance are not necessarily caused by an underlying phonological deficit.

Projecting this cause-effect discussion onto our results would lead to the following reasoning. From an *effect* point of view, learning to read can indeed lead to an increased ability to retain serial order given the importance of the serial order of graphemes in written words. When reading, one continuously 'practices' the retention of serial order information (i.e., sequence of graphemes). The alternative *cause(s)* point of view would be that preliterate children that are more efficient in the retention of serial order information in nonwords become better readers. Given the growing attention in dyslexia research regarding the potential role of serial order processing, it is important to find out whether serial order impairment could be one of the causal factors to reading (dis)ability. This knowledge would also directly affect remedial and intervention approaches as stated previously.

A potential solution to show a causal connection between serial order retention skills and reading performance could be to set up large-scale longitudinal studies. Indeed, this type of study design enables the field to study preliterate children with and without a family or genetic risk of dyslexia and monitor a number of potential child factors (e.g., cognitive processing, …) as well as environmental factors (e.g., reading experience) throughout their school career. This may shed further light on the current cause-effect discussion in dyslexia research. However, just as the sole use of cross-sectional studies creates an interpretive problem with respect to the causal nature of an independent variable, a weakness of longitudinal data is that there may be a third 'confounding' (i.e., unmeasured) factor, which influences both serial order retention and reading performance. In that case, the relation between serial order retention and reading performance would not be a causal one, but would be due to the fact that both are influenced by a third unknown factor. Therefore, some researchers suggest that longitudinal studies and intervention studies should be combined (e.g., Bryant & Bradley, 1985). Support

for a causal connection between serial order retention and reading performance would be supported in case an intervention that successfully improves children's serial order retention skills also predicts their reading performance (Bradley & Bryant, 1983).

Although the inclusion criteria were well-considered in all studies reported in this dissertation, we did not control for participants' IQ. However, there are some indications that it may be important to take this factor into account. Staels and Van den Broeck (2014) found that nonverbal intelligence was more strongly involved in serial order STM than in item STM. Hence, including a general measure for intelligence in our studies had been desirable. Other researchers showed that 4-year old at-risk children who went on to be classified as children with dyslexia were impaired on both verbal and nonverbal IQ relative to both children not at risk and children at-risk but who developed no dyslexia (van Bergen, de Jong, Maassen, Krikhaar, Plakas, & van der Leij, 2014). This seems to suggest an atypical cognitive development[16] for children who will fail to learn to read.

A final remark concerns the within-item manipulation of our retention variable (item retention vs. serial order retention). As argued in the introductions to the empirical studies, this manipulation allowed us to compare the impact of reading skill on these two types of STM retention with a single set of materials. However, this benefit comes with a cost. As the two retention types are highly correlated, the effect of reading skill can only be assessed after the statistical removal of these variables' shared variance on the dependent variable. This makes it more difficult to detect a significant effect, i.e., to reject a false null hypothesis. For instance, the two studies concerning the effect of L1 reading skill on phoneme retention (Chapters 2 and 3) failed to show a relation between reading skill and item retention. This is an unexpected outcome when considering the many demonstrations that poor readers find it difficult to become aware of phonemes. It is likely that the poor statistical power of the linear mixed model (due to the inclusion of the covariate) is responsible for this null effect. Even though these considerations even further underscore the importance of the effect of reading skill on serial order retention, future research is needed to find out whether poor readers have more problems with the retention of a nonword's phonemes as well.

[16] Stanovich (1991, 1994) argues against the idea that there exists a high correlation between IQ and reading as there are numerous studies where researchers find very low correlations between IQ and reading. More particularly, individuals with dyslexia may have a high or low IQ. However, an experiment studying the relationship between reading skill and STM retention performance will be better controlled when IQ is controlled for.

FUTURE DIRECTIONS

Although not included in this doctoral thesis, in a recent study, we focused on sequential bilingual poor and better L2 readers. In line with the results provided in Chapters 2-4, we wanted to examine whether we would find the same effect of literacy on item and order STM as in monolingual poor and better readers (*manuscript in preparation*). Our preliminary results suggest that the distinction between item STM and serial order STM also holds for sequential bilingual children learning to read in a second language and varying along a range of literacy skills in that language.

Another future line of research could focus on item and serial order STM processing at the phoneme level in pre-readers. As pointed out above, all current research is cross-sectional. Thus, it remains unclear whether impaired serial order STM should be seen as a cause of reading problems or as a consequence (see the discussion above). In case adequate serial order is a prerequisite of adequate reading skills, one should be able to detect problems with the recall of serial order information (of phonemes in the NRT, of non-phonemic verbal items, of nonverbal items) in children-at risk for dyslexia, i.e., in the pre-reading stage.

Although this dissertation exclusively focused on behavioral studies, evidence for the existence of two, partially independent retention processes in STM also comes from neuroimaging studies. For instance, Beneventi, Tonnessen, Ersland, and Hugdahl (2010) provide evidence regarding neuronal correlates that may relate to a working memory deficit in dyslexia. Using fMRI, children with and without dyslexia had to solve a verbal working memory task, the so called N-back task. Participants saw a sequence of stimuli (letters on a computer screen), which were presented one by one. For each stimulus, they needed to decide whether the current stimulus was the same as the one presented N trials ago. The N in this experiment was 0 (control condition), 1 trials and 2 trials. The higher the number, the more difficult the task became, as the result of an increased working memory load. The results showed that both groups recruited the same cortical network while performing this task. However, the results also revealed that the control group showed significantly more activation than individuals with dyslexia in two regions that have been associated with working memory processes including *temporal order memory* (see also Wager & Smith, 2003). Interestingly, Majerus et al. (2008) investigated the neural substrates of order STM in relation to bilingual language acquisition using an fMRI design. They compared two groups of German - French bilingual adults who differed in their second language proficiency. Participants performed a

STM item and serial order task. Based on their results, these authors concluded that low proficiency bilinguals activated STM networks for order in a less efficient and differentiated way. Hence, they suggested that their results might explain low proficiency bilinguals' poorer storage and learning capacity for verbal sequences.

All these studies used different tasks to compare item and serial order retention, as was the case in behavioral experiments focusing on these two types of memory (see the introductions to the various empirical chapters). It would be interesting to set up an event-related potentials (ERPs) study design in which item and serial order processing are explored at the phoneme level using a NRT task. This would make it possible to study item and serial order retention within items, as we have done in the studies reported in this doctoral thesis. The hypothesis would be that different ERP components are associated with the failure (or success) to recall a phoneme and the failure (or success) to recall a phoneme's position. This type of study has already been used previously, albeit with another aim, i.e., to determine how the lexicality of a stimulus affected both its immediate recognition and repeated processing (see Cheng, Schafer, & Riddell, 2014).

Finally, it would be worthwhile further investigating one of the suggested accounts regarding the observed serial order effect at the phoneme level, more particularly, the one focusing on the role of phonotactic knowledge. A next step in further understanding the role of phonotactics in phonological processing could be to manipulate the phonotactic probability of items. In case phonotactics is crucial and an underlying factor that guides serial order performance, it would be interesting to set up a study in which the phonotactic probability is manipulated in a NRT. Based on the outcomes in Chapter 2 and 3, the hypothesis would be that the difference between less skilled readers and controls on serial order performance is largest when using high 'phonotactic probability' words, whereas the difference between both groups would be smaller when low 'phonotactic probability' nonwords are included. Such a result would also be compatible with a domain-general account of poor readers' worse performance on the STM retention of phonemes' serial order in a nonword. As these readers would experience more problems with the implicit discovery of phonotactic patterns in the spoken language, due to less effective statistical learning (across cognitive domains), they would have less LTM support for encoding a nonword item in the NRT compared to their better reading peers.

As far as the importance of serial order recall and the role of phonotactics in sequential bilinguals is concerned, we could set up a study in which four groups of bilinguals are compared to monolinguals when performing a NRT task: (1) typically developing sequential bilinguals

with a native language that is completely different from the monolinguals' language in terms of phonotactic regularities, (2) poor second language readers with a native language that is completely different from the monolinguals' language in terms of phonotactic regularities, (3) typically developing sequential bilinguals speaking a language whose phonotactic regularities are highly similar to those that are used in the experimental items, and (4) poor second language readers speaking a language whose phonotactic regularities are highly similar to those that are used in the experimental items. Based on our experimental work in Chapter 4, we hypothesize that both bilingual groups whose language share less or no phonotactic regularities with the target language, i.e., the language in which the NRT performance is administered, will show less adequate serial order performance compared to the other group of bilinguals (whose native language does share phonotactic regularities with the language of the test) and monolinguals. Even though such an outcome would lend support to the idea that participants' knowledge about the phonotactic regularities of a language (stored in LTM) is a necessary ingredient for serial order retention and NRT performance, the basic problem would be different in both groups. Indeed, whereas the serial order retention problem in poor second language readers might stem from an ineffective (domain-general) sequence memory in STM, the same serial order problem in the subgroup of typically developing second language readers could be due to their poor knowledge of the phonotactics of the language of the experimental items (as an effective sequence memory in STM is expected in the latter group).

CONCLUSION

The three empirical studies presented in this doctoral thesis highlight the link between serial order retention in STM at the phoneme level and reading (impairment). We have demonstrated that the distinction between two types of STM retention, i.e., *phoneme identity* retention and *serial order* retention, contribute to a more precise understanding of NRT performance on the one hand and the relation between the two task components and reading proficiency on the other hand. Furthermore, we have shown that less skilled/disabled reading is more often associated with a failure to retain the serial order of phonemes. Overall, our findings contribute to a better understanding of the lower NRT performance that is typically observed in monolingual/bilingual disabled or less skilled readers and sequential bilingual children who need to process non-native phonological information. In sum, our research adds to a growing

body of research regarding the involvement of serial order memory for reading success, albeit without ignoring the potential role of phonotactic knowledge.

REFERENCES LIST

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

Badecker, W. (1996). Representational properties common to phonological and orthographic output systems. *Lingua, 99,* 55–83.

Beneventi, H., Tonnessen, F.E., Ersland, L., & Hugdahl, K. (2010). Working memory deficit in dyslexia: Behavioral and fMRI evidence. *International Journal of Neuroscience, 120,* 51-59.

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

Boets, B., Op de Beeck, H.P., Vandermosten, M., Scott, S.K., Gillebert, C.R., Mantini, D., … Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science, 342, 1*251–1254. doi:10.1126/science.1244333

Bogaerts, L., Szmalec, A., Hachmann, M.W., Page, M.P.A., & Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities, 43–44,* 106–122.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301,* 419-421.

Brown, G.D.A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107,* 127-181.

Bryant, P., & Bradley, L. (1985). *Children's reading problems.* Oxford: Blackwell.

Burgess, N., & Hitch, G.J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review, 106*, 551–581.

Burgess, N. & Hitch, G.J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*, 627–652.

Caramazza, A., Miceli, G., Villa, G., & Romani, C. (1987). The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition, 26,* 59–85.

Carroll, J., & Snowling, M. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry, 45,* 631-640.

Cheng, X., Schafer, G., Riddell, P.M. (2014). Immediate Auditory Repetition of Words and Nonwords: An ERP Study of Lexical and Sublexical Processing. *PLoS ONE, 9*(3)*,* e91988. doi: 10.1371/journal.pone.0091988

de Jong, P., & Vrielink, L. (2004). Rapid automatic naming: Easy to measure, hard to improve (quickly). *Annals of Dyslexia, 54,* 65-88.

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.

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., 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

Gathercole, S.E. (1994). Neuropsychology and working memory: A review. *Neuropsychology, 8*, 494-505. doi: 10.1037/0894-4105.8.4.49

Gathercole. S.E., & Baddeley, A.D. (1993). *Working memory and language.* Hove: Lawrence Erlbaum Associates.

Gathercole, S.E. (2006). Complexities and constraints in nonword repetition and word learning. *Applied Psycholinguistics, 27*, 599–613.

Gathercole, S.E., Alloway, T.P., Willis, C.S., & Adams, A.M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93,* 265–281.

Geudens, A. (2006). Phonological awareness and learning to read a first language: Controversies and new perspectives. *LOT Occasional Series, 6,* 25–43.

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.

Huettig, F., Lachmann, T., Reis, A., & Petersson, K.M. (2017). Distinguishing cause from effect - Many deficits associated with developmental dyslexia may be a consequence of reduced and suboptimal reading experience. *Language, Cognition and Neuroscience.* doi: 10.1080/23273798.2017.1348528.

Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., 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., 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.

Kahta, S., & Schiff, R. (2016). Implicit learning deficits among adults with developmental dyslexia. *Annals of Dyslexia,* 1–16.

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

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A. L., Martinez, T., Weekes, B., & Marquet, P. (2008). Neural networks for short-term memory for order differentiate high and low proficiency bilinguals. *Neuroimage, 49,* 1698-1713.

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

Martinez Perez, T., Majerus, S., & Poncelet, M. (2013). Impaired short-term memory for order in adults with dyslexia. *Research in Developmental Disabilities, 34,* 2211–2233. doi: 10.1016/j.ridd.2013.04.005

McCloskey, M., Badecker, W., Goodman-Schulman, R.A., & Aliminosa, D. (1994). The structure of graphemic representations in spelling: Evidence from a case of acquired dysgraphia. *Cognitive Neuropsychology, 11,* 341–392.

Messer, M., Leseman, P., Boom, J., & Mayo, A. (2010). Phonotactic probability effect in nonword recall and its relation with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology, 105,* 306–323. doi: 10.1016/j.jecp.2009.12.006

Nairne, J.S., & Kelley, M.R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50,* 113-133.

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

Page, M.P.A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761–781.

Pennington B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101,* 385–413. doi: 10.1016/j.cognition.2006.04.008

Perre, L. & Midgley, K., & Ziegler, J. (2009). When beef primes reef more than leaf : Orthographic information affects phonological priming in spoken word recognition. *Psychophysiology, 46,* 739-46. doi: 10.1111/j.1469-8986.2009.00813.x.

Ramus, F. (2003). Dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13,* 212–218.

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

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

Schmalz, X., Altoè, G., & Mulatti, C. (2017). Statistical learning and dyslexia: a systematic review. *Annals of Dyslexia, 67,* 147-162. doi:10.1007/s11881-016-0136-0

Seidenberg, M.S., & Tanenhaus, M.K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 546–554.

Share, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Share, D.L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72,* 95–129.

Snowling, M. (2000). *Dyslexia.* 2nd. Oxford: Blackwell.

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.

Suárez-Coalla, P., Avdyli, R., & Cuetos, F. (2014). Influence of context-sensitive rules on the formation of orthographic representations in Spanish dyslexic children. *Frontiers in Psychology, 5,* 1409.

Summers C., Bohman T.M., Gillam R.B., Peña E.D., Bedore L.M. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language and Communication Disorders, 45,* 480–493. doi: 10.3109/13682820903198058

Van Bergen, E., de Jong, P.F., Maassen, B., Krikhaar, E., Plakas, A., & van der Leij, A. (2014). IQ of four-year-olds who go on to develop dyslexia. *Journal of Learning Disabilities, 47(5),* 475-484. doi: 10.1177/0022219413479673

Van Bergen, E., van der Leij, A., & de Jong, P. F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in Human Neuroscience, 8,* 346. doi: 10.3389/fnhum.2014.00346

Vellutino, F.R., Fletcher, J.M., Snowling, M.J., & Scanlon, D.M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2–40. doi: 10.1046/j.0021-9630.2003.00305.x

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

Vitevitch, M.S., Luce, P.A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech, 40,* 47–62.

Wager, T., & Smith, E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience, 3,* 255-274.

Wang, X., Xuan, Y., & Jarrold, C. (2016). Using a process dissociation approach to assess verbal short-term memory for item and order information in a sample of individuals with a self-reported diagnosis of dyslexia. *Frontiers in Psychology, 7,* 208. doi: 10.3389/fpsyg.2016.00208

APPENDIX A

Target Items for the Nonword Repetition Task (NRT)

	Three-syllable items		**Four-syllable items**		**Five-syllable items**	
	Phonological similar	Phonological distinct	Phonological similar	Phonological distinct	Phonological similar	Phonological distinct
Low word complexity	/pippókket/ /nimmúnnaf/ /savvéffus/	/kummíggar/ /possállin/ /gikkállom/	/ninnellúmmar/ /nossiggéffas/ /peddattíttup/	/lummoggáppes/ /beffottírral/ /hissorréffum/	/bikkottaddúkkep/ /gaggollissúggef/ /nammonniffúnnem/	/sorrammikkéttul/ /lurreffanníppos/ /hussallimmóggep/
High word complexity	/vesrófsif/ /pirdápket/ /nomlúnfam/	/lemróspag/ /pigdúlmek/ /harlónwig/	/kepdarbítpuk/ /zosgefzílvas/ /mifnemlúmzan/	/sirpegwótnal/ /jafkertúmsil/ /kegjolbífmas/	/vugzasgoflífsef/ /piptafbipkétduk/ /munfomlinzémbam/	/gambiskeflúnjor/ /wogtumdisjáfkel/ /bognupsarlíftek/