The role of linguistic diversity in early reading comprehension

Liza J. van den Bosch

Behavioural
Science
Institute

THE ROLE OF LINGUISTIC DIVERSITY IN EARLY READING COMPREHENSION

Liza J. van den Bosch

The research presented in this dissertation was carried out at the
Behavioural Science Institute of the Radboud University, the Netherlands.

ISBN: 978-94-6375-633-4

Cover design: Jeanne | ontwerp & illustratie
Lay-out: Anna Bleeker | www.persoonlijkproefschrift.nl
Printing: Ridderprint | www.ridderprint.nl

© Liza J. van den Bosch, 2020

All rights reserved. No parts of this publication may be reproduced, distributed, stored in a retrieval system, or transmitted in any form or by any means without prior written permission from the author.

THE ROLE OF LINGUISTIC DIVERSITY IN EARLY READING COMPREHENSION

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen
op dinsdag 14 januari 2020
om 14.30 uur precies

door

Liza Jane van den Bosch

geboren op 27 augustus 1991
te Nootdorp

Promotoren
Prof. dr. L.T.W. Verhoeven
Prof. dr. P.C.J. Segers

Manuscriptcommissie
Prof. dr. C.F. Rowland (voorzitter)
Prof. dr. W.B.T. Blom (Universiteit Utrecht)
Prof. dr. A.M. Backus (Tilburg University)

Table of Contents

Chapter 1	General introduction	7
Chapter 2	The role of linguistic diversity in the prediction of early reading comprehension performance: A quantile regression approach	25
Chapter 3	The impact of first and second language vocabulary on early second language reading comprehension development	55
Chapter 4	Online processing of causal relations in beginning readers from diverse linguistic backgrounds	81
Chapter 5	Online processing of coherence relations in beginning readers from diverse linguistic backgrounds	107
Chapter 6	General discussion	145
	Nederlandse samenvatting	163
	Dankwoord	171
	Curriculum Vitae	175

Chapter 1

General introduction

Reading is one of the most crucial skills children acquire at primary school. It enables them to comprehend written text, which is a prerequisite for academic success and for participation in society. Reading comprehension can be defined as the process of simultaneously extracting and constructing meaning from written language (Snow, 2002). This is a complex process that involves a range of cognitive and linguistic factors (Stanovich, 1980). Reading comprehension is considered to be even more challenging for children who learn to read in their second language (L2 readers), such as children from language minorities for whom the language of instruction at school differs from their home language. Previous research has shown that L2 readers generally lag behind their monolingual peers (L1 readers) in terms of reading comprehension performance (Melby-Lervåg & Lervåg, 2014). This L1-L2 performance gap can already be observed in the intermediate (Droop & Verhoeven, 2003; Verhoeven, 2000) and higher primary grades (Farnia & Geva, 2013; Mancilla-Martinez & Lesaux, 2010).

Remarkably, little is known about the role of linguistic diversity in the early stages of reading development. It is by no means clear to what extent L1 readers and L2 readers differ in the transition they make from reading isolated words to comprehending sentences and short texts. Therefore, the aim of the present dissertation was to examine reading comprehension outcomes and reading comprehension processes in beginning L1 readers and L2 readers and to explain between-group variation as well as individual variation. Before going into detail about the present research, this introductory chapter first presents a theoretical framework on reading comprehension and describes how reading comprehension processes can be studied by means of online measures. Then, the role of linguistic diversity is introduced and an overview of previous studies on reading comprehension in L1 readers and L2 readers is provided.

Reading Comprehension Theories

Given the complexity of reading comprehension and the multiple factors involved, several theories have emerged in the field of reading research in the past decades. Among these theories, different reading comprehension components and/or processes have been identified, with some theories being more specific than others.

A prominent, general theory is the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990), which defines reading comprehension as the product of two broad components: word reading and language comprehension. Word reading skills are required for decoding written text and language comprehension skills

are required for understanding written text. If readers do not sufficiently master one or both of these skills, this will result in reading comprehension difficulties. Furthermore, it is assumed that the relative importance of both components changes throughout reading development, which is evidenced by empirical studies showing that the amount of variance explained by decoding skills decreases, whereas the amount of variance explained by listening comprehension increases across grades (Catts, Hogan, & Adlof, 2005; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009). This theory has been broadly supported by empirical studies across a range of languages (Florit & Cain, 2011), but it does not explain what the underlying linguistic and cognitive factors of reading comprehension are and how the process of reading comprehension actually evolves.

The Construction-Integration model (Kintsch & Van Dijk, 1978; Kintsch, 2004) is a more detailed and process-oriented model that distinguishes multiple levels of text comprehension. In the process of reading, the reader first needs to decode written text to access the meaning of the individual words (surface level). Once the meanings have been selected, the reader needs to combine the words into meaningful clauses and sentences (textbase level). Then, ideas presented in successive sentences have to be integrated by making inferences between text elements and the reader's prior knowledge to establish a coherent mental model of the text (situation model level). This model states that text comprehension includes both bottom-up word-reading processes and top-down comprehension processes (Kintsch, 2005).

An overarching theory that addresses the components as well as the processes of reading comprehension from an integrative perspective is the Reading Systems Framework (Perfetti & Stafura, 2014). According to this theory, three knowledge sources (i.e., linguistic, orthographic, and general knowledge) are involved in the reading process, which takes place within a cognitive system with limited processing resources. The reading process involves decoding, word identification, meaning retrieval, sentence parsing, inference making, and comprehension monitoring. This framework can be used to identify reading problems by measurable weaknesses within one or more of the components (knowledge sources and processes) or to identify pressure points in the reading system. The Lexical Quality Hypothesis (Perfetti, 2007) assumes that one of those pressure points is the lexicon, which connects the word identification system (i.e., lower-level processes) and the comprehension system (i.e., higher-level processes). This hypothesis claims that lexical quality, defined as the knowledge of written word forms (i.e., word

recognition) and word meanings (i.e., vocabulary knowledge), affects reading comprehension. Indeed, empirical evidence has been found for the impact of lexical quality, in particular vocabulary knowledge, on children's reading comprehension performance (Swart et al., 2017; Verhoeven & Van Leeuwe, 2008).

Reading Comprehension Processes

As part of the reading process, the reader needs to establish a coherent mental representation of the text during reading by making meaningful connections between different text elements, both at a local and a global level. At the local level, this involves comprehension processes in which the reader integrates word meanings within a sentence or across sentences, also referred to as word-to-text integration processes (Perfetti, Yang, & Schmallhofer, 2008). These integration processes are central to the reading process of beginning readers; once children have acquired accurate decoding skills they move from reading isolated words to making sense out of short stretches of text, such as clauses and sentences. Apart from child-related factors, such as their lower-level and higher-level comprehension skills, there are also textual factors that can support children's word-to-text integration processes. For instance, linguistic markers in the text, such as connectives (e.g., before, because) that signal the semantic relation between clauses and/or sentences can guide the reader in how to construct meaning during reading (Sanders & Noordman, 2000). This can be illustrated by the following sentences:

1A. Tom was crying <u>because</u> he fell on his knee.
1B. Tom was crying <u>although</u> he fell on his knee.
2A. <u>Before</u> Laura jumped in the swimming pool she put on her swimsuit.
2B. <u>After</u> Laura jumped in the swimming pool she put on her swimsuit.

The connectives signal different types of semantic relations (i.e., causal, adversative, and temporal relations) between clauses (Halliday & Hasan, 1976; Sanders, Spooren, & Noordman, 1992) and largely determine the meaning of these sentences. Readers who have accurate knowledge of connectives will notice that the A-sentences make more sense than the B-sentences. For children, however, this may be challenging, since their knowledge of connectives and understanding of the semantic relations they signal is still developing (Cain & Nash, 2011). Nevertheless, connectives frequently occur in written text, even for beginning readers (Tellings & Penning-de Vries, 2019), but it is by no means clear how children deal with these linguistic markers during reading.

Therefore, it is highly important to gain detailed insight into children's online comprehension processes to explain their reading comprehension outcomes. Ultimately, this involves the use of online measures that reveal moment-to-moment comprehension processes during reading (Perfetti & Stafura, 2014). In that respect, eye-tracking technology has proven to be a useful tool for investigating cognitive processes underlying reading; the readers' eye movements during reading are assumed to be good indicators of their ongoing cognitive processes (Rayner, 1998; Rayner, Chace, Slattery, & Ashby, 2006). The major advantage of eye-tracking is that it can tap into real-time comprehension processes in a natural reading setting without interrupting the reading process (Sereno & Rayner, 2003), in contrast to think-aloud or self-paced reading paradigms. Technological advances and increased understanding of eye-movements during reading in the past decades (Radach & Kennedy, 2004), enabled researchers nowadays to use this methodology more often to study online text processing. With respect to connectives, previous studies have successfully used eye-tracking to gain detailed insights into how these linguistic markers facilitate online processing in adults (Canestrelli, Mak, & Sanders, 2013; Cozijn, Noordman & Vonk, 2011) and adolescents (Van Silfhout, Evers-Vermeul, & Sanders, 2015), but such studies with children are scarce. Although researchers have started to use eye-tracking more often to study real-time reading processes in children (Blythe & Joseph, 2011), these studies mainly included more experienced, monolingual readers in the upper grades of primary school (e.g., De Leeuw, Segers, & Verhoeven, 2016; Van der Schoot, Vasbinder, Horsley, & Van Lieshout, 2008).

The Role of Linguistic Diversity

As a result of globalization and migration, many children grow up in linguistically diverse environments nowadays. Worldwide, a considerable number of children can be considered as being bilingual (or multilingual). Although definitions of bilingualism vary and different types of bilingualism have been distinguished depending on the degree of language exposure (balanced vs. unbalanced bilinguals), the age of acquisition (early vs. late bilinguals) and the order of acquisition (simultaneous vs. sequential bilinguals), it is clear that bilinguals differ from monolinguals by the fact that they acquire more than one language. With respect to reading development, being bilingual implies that children learn to read in a multilingual context. For children from a linguistic minority background this often means that they learn to read in the majority language, which differs from the language they speak at home. In many countries, these children are completely immersed in a second language (L2) reading curriculum with minimal attention paid to their native (L1)

language and literacy skills. Children from a linguistic minority background are thus confronted with the challenging task of becoming literate in a language that they have been little exposed to and that they have not fully acquired yet. This affects their performance at school, including their reading comprehension performance.

Variation between L1 readers and L2 readers

Previous studies that compared monolinguals (L1 readers) and bilinguals (L2 readers) have demonstrated that L2 readers generally fall behind their monolingual peers with respect to their reading comprehension performance (Melby-Lervåg & Lervåg, 2014). This L1-L2 performance gap is mostly evidenced in the upper primary grades (Farnia & Geva, 2013; Mancilla-Martinez & Lesaux, 2010), but can also be observed in the intermediate primary grades (Droop & Verhoeven, 2003; Verhoeven, 2000). However, longitudinal studies on reading comprehension that compared L1 readers and L2 readers are scarce and inconclusive about whether and how the performance gap changes over time (Droop & Verhoeven, 2003; Farnia & Geva, 2013; Lervåg & Aukrust, 2010; Limbird, Maluch, Rjosk, Stanat, & Merkens, 2014; Verhoeven, 2000). Yet, previous studies have revealed some relevant insights into L2 reading comprehension in relation to the theories discussed at the beginning of this chapter.

It has been found that the Simple View of Reading and the Lexical Quality Hypothesis also apply to L2 readers (Verhoeven & Van Leeuwe, 2012; Verhoeven, Voeten, & Vermeer, 2019), and that L2 readers' reading comprehension difficulties usually stem from delays in language comprehension rather than poor word reading ability. More specifically, L2 readers often show a delay in their (L2) oral language skills (Melby-Lervåg & Lervåg, 2014), such as their vocabulary knowledge and (morpho)syntactic knowledge (i.e., knowledge about the structure of words and sentences). By contrast, L2 readers' decoding skills are found to be similar, or even more efficient, compared to their monolingual peers (Droop & Verhoeven, 2003; Geva & Zadeh, 2006; Mancilla-Martinez & Lesaux, 2010). With respect to the Construction-Integration model and the Reading Systems Framework, it can thus be stated that L2 readers' difficulties mainly occur at the higher levels of text comprehension (i.e., textbase level and situation model) and relate to the higher-order comprehension processes. There is no reason to assume that L2 readers' cognitive capacity limits their reading comprehension. Although it has been criticized (e.g., Duñabeitia et al., 2014; Paap & Sawi, 2014), recent studies have even suggested a cognitive advantage of bilingualism regarding executive functions, such as working memory (Blom, Küntay, Messer, Verhagen, & Leseman, 2014; Morales,

Calvo, & Bialystok, 2013), as a consequence of dual language management (Grundy & Timmer, 2017). This might enable them to compensate for their delays in oral language skills during comprehension processes (Bialystok & Craik, 2010).

Yet, previous studies comparing L1 readers and L2 readers have primarily focused on the reading comprehension outcomes by using offline measures while much less attention has been devoted to reading comprehension processes. To gain more detailed insights into their comprehension processes during reading, online measures can be used to investigate to what extent L1 readers and L2 readers differ in that respect. Previous studies examining integration processes in adult readers revealed processing differences between L1 readers and L2 readers (Kaan, Kirkham, & Wijnen, 2016; Martin et al., 2013), but there are hardly any studies in which online measures have been used to examine comprehension processes in young L1 readers and L2 readers. For instance, it is by no means clear if and how L2 readers make use of linguistic markers, such as connectives, during reading in order to construct meaning. On the one hand, it can be argued that connectives help L2 readers to identify text structure and to establish a coherent mental model as found for young monolingual readers (Cain & Nash, 2011). On the other hand, it can be predicted that because of their lower linguistic proficiency in the target language, L2 readers are not competent enough to benefit from the coherence marking provided by connectives as found for adult L2 readers (Degand & Sanders, 2002). It has been shown that young L2 readers have lower knowledge of connectives (Crosson, Lesaux, & Martiniello, 2008) and that the impact on reading comprehension outcomes is smaller compared to their monolingual peers (Crosson & Lesaux, 2013), but studies investigating the impact of connectives on L2 readers' actual comprehension processes are currently lacking.

Individual Variation

Although comparing children based on their language background (L1 vs. L2) can reveal fundamental group differences, categorizing them as either monolingual or bilingual may result in an unfair comparison when only one language is considered, given that bilingual children usually receive less exposure in each of their two languages compared to monolingual children (Bialystok, Luk, Peets, & Yang, 2010). Furthermore, such a comparison may not do justice to individual variation across children (Luk & Bialystok, 2013), since considerable variation exists in the quantity and quality of language exposure. Consequently, within-group variance may be similar or even larger than between-group variance. Therefore, it is highly important to disentangle children's language background (L1 vs. L2) from their

actual ability and to examine individual differences among children, and among L2 readers in particular. Ultimately, to obtain a full picture of L2 readers' proficiency, not only the language of instruction (L2), but also their native language (L1) should be considered.

More specifically, it is essential to assess bilingual children's vocabulary knowledge in both their languages as indicators of their mental lexicon, since discrepancies in word knowledge may exist among their two languages (Bialystok et al., 2010). For instance, due to using their languages in different contexts (at home vs. at school), L2 readers may understand a specific word in one language but not the other, or may have a deeper understanding of particular words (i.e., higher lexical quality) in either language. However, previous studies barely assessed children's vocabulary knowledge in their native language, probably due to a heterogeneous L2 group with diverse native languages (e.g., Farnia & Geva, 2013; Lesaux, Lipka, & Siegel, 2006) and/or a lack of reliable measures. The few studies that did consider children's L1 vocabulary knowledge showed mixed results. In line with the notion that competence in a second language is partly based on competence in a first language (Cummins, 1979) some studies found an effect of L1 vocabulary on L2 reading comprehension (Proctor, August, Carlo, & Snow, 2006; Raudszus, Segers, & Verhoeven, 2018), whereas others did not (Gottardo & Mueller, 2009; Lervåg & Aukrust, 2010; Mancilla-Martinez & Lesaux, 2010). To shed light on these discrepant findings, individual variation in terms of proficiency in both languages should be considered in a group of L2 readers with a similar native language, given that cross-linguistic influence may depend on children's language proficiency and the language distance across their languages (Nicoladis, 2012).

The Present Dissertation

Taken together, it can be concluded that previous studies have revealed fundamental group differences between monolingual children (L1 readers) and bilingual children (L2 readers) regarding reading comprehension outcomes and the explanatory factors. Nevertheless, previous studies show several shortcomings. To begin with, they were mostly focused on the intermediate/upper primary grades, whereas a performance gap already seems to exist in the lower primary grades. Second, insofar studies have been conducted, they usually followed a cross-sectional rather than a longitudinal design with an almost one-sided focus on the (offline) reading comprehension outcomes while minimal attention has been devoted to (online) reading comprehension processes. Finally, in previous studies a comparison was made between L1 readers and L2 readers starting from the proposition that linguistic diversity can be seen as a dichotomy thereby neglecting individual variation among children. Consequently, more insight is needed into reading comprehension performance, development and online processes in the early stages of reading acquisition and into how individual variation among children from diverse linguistic backgrounds can be explained.

Aim and Research Questions

The present dissertation investigated the role of linguistic diversity in early reading comprehension. The general aim was to examine reading comprehension performance, development and online processes in beginning L1 readers and L2 readers and to explain between-group variation as well as individual variation. Accordingly, the following research questions were central to the present dissertation:

1. To what extent do L1 readers and L2 readers differ in early reading comprehension performance, development, and online comprehension processes?
2. How can individual variation in reading comprehension performance and development among L1 readers and L2 readers be explained?
3. How can individual variation in online comprehension processes among L1 readers and L2 readers be explained?

To address these research questions, a two-fold research project was set up, focusing on the lower grades of primary school, where reading instruction is a central part of the curriculum.

With respect to reading comprehension performance and development, a cross-sectional study (Chapter 2) and a longitudinal follow-up study (Chapter 3) were conducted to answer research questions 1 and 2. With respect to online comprehension processes, two experimental studies (Chapters 4 and 5) were conducted to answer research questions 1 and 3. Together, these studies shed light on the role of linguistic diversity by not only comparing L1 readers and L2 readers from a binary view, but also by exploring other, more refined ways of explaining individual variation across children with respect to early reading comprehension.

Participants and Design

All four studies are conducted with monolingual Dutch children (L1 readers) and bilingual Turkish-Dutch children (L2 readers). In the Netherlands, one out of eight primary school children have a linguistic minority background (CBS, 2015). Turkish is the largest linguistic minority (CBS, 2016) and the most widely spoken additional language among Dutch primary school children (Extra, 2002). Usually, these Turkish-Dutch children grow up in families with a lower socioeconomic background in a situation of emergent bilingualism; they can be considered as sequential bilinguals. Although there is considerable varation across families (Bezcioglu-Goktolga & Yagmur, 2018), they are mainly exposed to their first language (Turkish) at home, while their second language (Dutch) gradually enters their lives once they go to school (Scheele, Leseman, & Mayo, 2010). From first grade onward they receive formal reading instruction in their second language, while very minimal attention is paid to their native language at school.

For the present dissertation, longitudinal data were collected at three time points across second and third grade of primary education with an interval of approximately six months: T1) start second grade, T2) end second grade, and T3) mid third grade, as shown in Figure 1.1. Given the focus on the early stages of reading comprehension, the data collection started in second grade assuming that after one year of formal reading instruction children sufficiently master the alphabetic principle (Schaars, Segers, & Verhoeven, 2017). The first wave of data collection included a broad assessment of children's cognitive, language, and reading skills for the cross-sectional study (Chapter 2). The second and third waves included follow-up measures of their reading skills for the longitudinal study (Chapter 3). In addition, the two experimental eye-tracking studies were conducted mid third grade, where the focus in reading instruction shifts toward reading comprehension and children's decoding skills have become more efficient, allowing for reliable recording of their eye-movements during reading. The first

experimental study (Chapter 4) was conducted with a separate sample, whereas the second experimental study (Chapter 5) was conducted with a representative subsample of the participants from the longitudinal study (third time point).

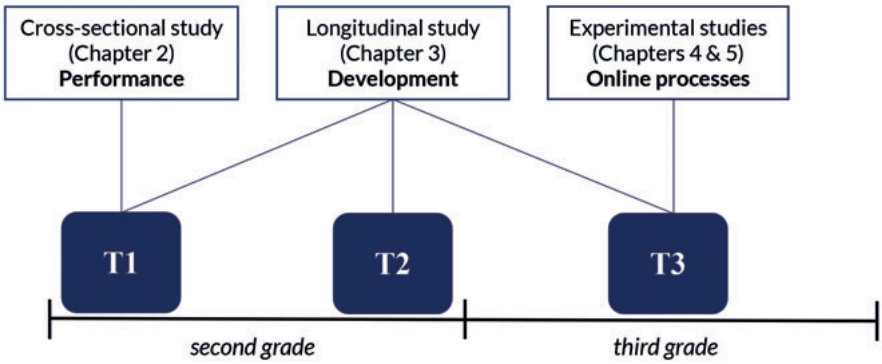

Figure 1.1 Graphical overview of the different types of studies on different aspects of reading comprehension.

Outline of this Dissertation

This dissertation consists of four empirical studies, which are submitted or accepted for publication in international peer-reviewed journals. These are different types of studies in which advanced statistical and/or methodological techniques have been used.

Chapter 2 describes a cross-sectional study on children's reading comprehension performance in second grade. Using classical and quantile regression analyses, this study investigated whether the explanatory factors of early reading comprehension performance differed depending on children's language background (L1 vs. L2) and/or reading comprehension ability level (poor vs. average vs. good).

Chapter 3 represents a longitudinal follow-up study on these children's reading comprehension development across second and third grade. Based on vocabulary knowledge in both their languages, cluster analysis identified four subgroups of L2 readers with varying levels of L1 (Turkish) and L2 (Dutch) vocabulary knowledge and different patterns of growth in reading comprehension.

Chapter 4 and Chapter 5 contain experimental studies using eye-tracking technology to investigate children's online processing of different types of semantic

relations in third grade. First, a small-scale study on causal coherence relations (e.g., 'Tom ate an apple because he was hungry.') was conducted (Chapter 4). To elaborate on that, a more extensive study focusing on both causal coherence relations and temporal coherence relations (e.g., 'Sara watched television after she baked biscuits.') was performed (Chapter 5).

To conclude, Chapter 6 provides a summary and discussion of main findings of the four empirical studies. Furthermore, limitations, directions for future research, and implications for educational practice are described in this final chapter.

References

Bezcioglu-Goktolga, I., & Yagmur, K. (2018). Home language policy of second-generation Turkish families in the Netherlands. *Journal of Multilingual and Multicultural Development*, *39*, 44-59. doi:10.1080/01434632.2017.1310216

Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science*, *19*, *19*-23. doi:10.1177/0963721409358571

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*, 525-531. doi:10.1017/S1366728909990423

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish-Dutch children. *Journal of Experimental Child Psychology*, *128*, 105-119. doi:10.1016/j.jecp.2014.06.007

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye-movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.). *The Oxford Handbook of Eye-movements* (pp. 643–662). Oxford: Oxford University Press.

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*, 429-441. doi:10.1037/a0022824

Canestrelli, A. R., Mak, W. M., & Sanders, T. J. M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, *28*, 1394-1413. doi:10.1080/01690965.2012.685885

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25-40). Mahwah, NJ: Lawrence Erlbaum.

CBS Statistics the Netherlands (2015). Leerlingen in (speciaal) basisonderwijs; migratie-achtergrond [Pupils in primary (special) education; migration background]. Retrieved from https://opendata.cbs.nl/#/CBS/nl/dataset/83295NED/table?ts=1558986759092

CBS Statistics the Netherlands (2016). Bevolking naar migratieachtergrond [Population by immigration background]. Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Cozijn, R., Noordman, L. G. M., & Vonk, W. (2011). Propositional integration and world knowledge inference: Processes in understanding because sentences. *Discourse Processes*, 48, 475-500. doi:10.1080/0163853X.2011.594421

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading*, *36*, 241-260. doi:10.1111/j.1467-9817.2011.01501.x

Crosson, A. C., Lesaux, N. K., & Martiniello, M. (2008). Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds. *Applied Psycholinguistics*, *29*, 603-625. doi:10.1017/S0142716408080260

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*, 222-251. doi:10.3102/00346543049002222

De Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text related characteristics on student's reading behaviour and text comprehension: An eye movement study. *Scientific Studies of Reading*, *20*, 248-263. doi:10.1080/10888438.2016.1146285

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing: An Interdisciplinary Journal*, 15, 739-757. doi:10.1023/A:1020932715838

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, *38*, 78-103. doi:10.1598/RRQ.38.1.4

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: Myth or reality? *Experimental Psychology*, *61*, 234-251. doi:10.1027/1618-3169/a000243

Extra, G. (2002). *De andere talen van Nederland: Thuis en op school. [The other languages of the Netherlands: At home and at school]*. Bussum, the Netherlands: Coutinho.

Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*, 389-4212. doi:10.1111/jrir.12003

Florit, E., & Cain, K. (2011). The Simple View of Reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*, 553-576. doi:10.1007/s10648-011-9175-6

Geva, E., & Zadeh, Z. Y. (2006). Reading efficiency in native English-speaking and English as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading*, *10*, 31-57. doi:10.1207/s1532799xssr1001_3

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, *10*, 330-344. doi:10.1037/a0014320

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, *7*, 6-10. doi:10.1177/074193258600700104

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*, 325-340. doi:10.1177/0267658316678286

Halliday, M. E. K., & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2, 127-160. doi:10.1007/BF00401799

Kaan, E., Kirkham, J., & Wijnen, F. (2016). Prediction and integration in native and second language processing of elliptical structures. *Bilingualism: Language and Cognition*, *19*, 1-18. doi:10.1017/S1366728914000844

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (Fifth edition, pp. 1270-1328). Newark, DE: International Reading Association.

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes*, *39*, 125-128. doi:10.1080/0163853X.2005.9651676

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394. doi:10.1037/0033-295X.85.5.363

Lervåg, A. & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*, 612-620. doi:10.1111/j.1469-7610.2009.02185.x

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, *19*, 99-131. doi:10.1007/s11145-005-4713-6

Limbird, C. K., Maluch, J. T., Rjosk, C. R., Stanat, P., & Merkens, H. (2014). Differential growth patterns in emerging reading skills of Turkish-German bilingual and German monolingual primary school students. *Reading and Writing: An Interdisciplinary Journal*, *27*, 945-968. doi:10.1007/s11145-013-9477-99

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*, 605-621. doi:10.1080/20445911.2013.795574

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority children. *Journal of Educational Psychology*, *102*, 701-711. doi:10.1037/a0019135

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*, 574-588. doi:10.1016/j.jml.2013.08.001

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409-433. doi:10.1037/a0033890

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*, 187-202. doi:10.1016/j.jecp.2012.09.002

Nicoladis, E. (2012). Cross-linguistic influence in French-English bilingual children's Possessive constructions. *Bilingualism: Language and Cognition*, *15*, 320-328. doi:10.1017/S1366728911000101

Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, *5*, 1–15. doi:10.3389/fpsyg.2014.00962

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

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22-37. doi:10.1080/10888438.2013.827687

Perfetti, C., Yang, C., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*, 303-318. doi:10.1002/acp.1419

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology*, *98*, 159-169. doi:10.1037/0022-0663.98.1.159

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, 16, 3-26. doi:10.1080/09541440340000295

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422. doi:10.1037/0033-2909.124.3.372

Rayner, K., Chace, K. H., Slattery, T. J., Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*, 241-255. doi:10.1207/s1532799xssr1003_3

Raudszus, H., Segers, E., & Verhoeven, L. (2018). Lexical quality and executive control predict children's first and second language reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *31*, 405-424. doi:10.1007/s11145-017-9791-8

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1-35. doi:10.1080/01638539209544800

Sanders, T., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*, 37-60. doi:10.1207/S15326950dp2901_3

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2017). Word decoding development in incremental phonics instruction in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal*, *30*, 1529-1550. doi:10.1007/s11145-017-9735-3

Scheele, A. F., Leseman, P. P. M., & Mayo, A. F. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*, 117-140. doi:10.1017/S0142716409990191

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences*, *7*, 489-493. doi:10.1016/j.tics.2003.09.010

Snow, C. E. (2002). *Reading for understanding.* Santa Monica, CA: RAND Education and the Science and Technology Policy Institute.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*, 32-71. doi:10.2307/747348

Swart, N. M., Muijselaar, M. M. L., Steenbeek-Planting, E. G., Droop, M., De Jong, P. F., &
Verhoeven, L. (2017). Cognitive precursors of the developmental relation between lexical quality and reading comprehension in the intermediate elementary grades. *Learning and Individual Differences*, *59*, 43-54. doi:10.1016/j.lindif.2017.08.009

Tellings, A., & Penning-de Vries, B. (2019). Connective frequency in child-directed texts: Dutch corpus analysis. Manuscript submitted for publication.

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, *32*, 383-401. doi:10.1111/j.1467-9817.2009.01401.x

Van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. D. M. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading*, *31*, 203–223. doi:10.1111/j.1467-9817.2007.00354.x

Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, 52, 47-76. doi:10.1080/0163853X.2014.905237

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*, 313-330. doi:10.1207/S1532799XSSR0404_4

Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, *22*, 407-423. doi:10.1002/acp.1414

Verhoeven, L., & Van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing*, *25*, 1805-1818. doi:10.1007/s11145-011-9346-3

Verhoeven, L., Voeten, M., & Vermeer, A. (2019). Beyond the simple view of early first and second language reading: The impact of lexical quality. *Journal of Neurolinguistics*, *50*, 28-36, doi:10.1016/j.jneuroling.2018.03.002

Chapter 2

The role of linguistic diversity in the prediction of early reading comprehension performance: A quantile regression approach

This chapter is based on:
Van den Bosch, L. J., Segers, E., & Verhoeven, L. (2019). The role of linguistic diversity in the prediction of early reading comprehension: A quantile regression approach. *Scientific Studies of Reading*, *23*, 203-219. doi:10.1080/10888438.2018.1509864

Abstract

In this cross-sectional study, classical and quantile regression analyses were performed to investigate whether predictor variables for early reading comprehension performance differed depending on language background and/or ability level in a mixed group of 161 monolingual (L1) and bilingual (L2) children in second grade (6–8 years). Although L2 readers scored lower on oral language skills and reading comprehension, the prediction of reading comprehension was similar for L1 readers and L2 readers. Classical regression identified decoding, vocabulary, and morphosyntactic knowledge as unique predictors with no interaction with language background. Quantile regression demonstrated that the prediction of reading comprehension differed across ability levels. Decoding and morphosyntactic knowledge were uniquely related to reading comprehension across the ability range, vocabulary was uniquely related for poor comprehenders only and working memory was uniquely related for good comprehenders only. In both types of analyses, language background did not explain unique variance, indicating that individual differences in the predictor variables explained the L1–L2 performance gap in early reading comprehension across the ability range.

Introduction

Given its utmost societal relevance, it is important to gain insight into the factors that predict reading comprehension performance. Reading comprehension has been associated with linguistic diversity, with second language (L2) readers generally falling behind their native language (L1) peers (Melby-Lervåg & Lervåg, 2014), although there may be L2 readers who do not struggle with reading comprehension and L1 readers who do. In studies investigating the factors accounting for variation in reading comprehension, children usually have been split up into L1 and L2 readers. Mixed results have been reported; some studies have shown that oral language skills are more important for L2 readers compared to L1 readers (Babayiğit, 2014; Droop & Verhoeven, 2003; Geva & Farnia, 2012; Lervåg & Aukrust, 2010), whereas other studies have found no differences (Lesaux, Lipka & Siegel, 2006; Verhoeven & Van Leeuwe, 2012). These studies, however, did not attempt to disentangle language background from ability level when investigating the explanatory factors of reading comprehension. Furthermore, most of the studies were focused on upper primary grades, whereas comprehension difficulties can already be observed in the early stages of reading acquisition (Verhoeven, 2000; Verhoeven, Voeten & Vermeer, 2019).

In the present study, we therefore combined classical and quantile regression analyses to gain insight into individual variation in *early* reading comprehension performance in a mixed group of L1 and L2 readers. First, we used classical regression analysis to investigate whether key predictors of reading comprehension (determined on the basis of prominent theories) differed depending on children's language background (L1 vs. L2). As a next step, we used quantile regression analysis to investigate whether the key predictors differed depending on children's ability level (poor vs. average vs. good). Together, these two complementary approaches shed light on the role of linguistic diversity in the prediction of early reading comprehension across ability levels.

Individual Variation in Reading Comprehension

Reading comprehension is a complex activity involving a range of cognitive and linguistic factors (Stanovich, 1980). According to the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990), there are two central components of reading comprehension: word reading skills and language comprehension. More recently, Perfetti and Stafura (2014) aimed to capture the complexity of reading comprehension in the Reading Systems Framework. According to this theory,

three knowledge sources (i.e., linguistic, orthographic, and general knowledge) are involved in the reading process, which takes place within a cognitive system with limited processing resources. The Reading Systems Framework can be used to identify pressure points (i.e., candidates for a cause of reading comprehension difficulty) in the reading system.

The Lexical Quality Hypothesis (Perfetti, 2007) assumes that one of those pressure points is the lexicon, which can be defined as knowledge of written word forms and word meanings, often referred to as vocabulary knowledge. The lexicon connects the word identification system with the comprehension system. First, linguistic and orthographic information is needed for rapid word identification (i.e., word decoding). Then, the comprehension system requires input from the word identification system to build meaning units. For integrating information at the sentence level, morphology (i.e., rules about the structure of words) and syntax (i.e., rules about the structure of sentences) play an important role (Perfetti & Stafura, 2014), also called morphosyntactic knowledge. Furthermore, the cognitive capacities that underlie the process of reading are crucial. Working memory capacity (i.e., temporary storage and manipulation of information; Baddeley, 2003) is required for higher-level comprehension processes, such as integrating information, inference making and comprehension monitoring (Kendeou, Van den Broek, Helder, & Karlsson, 2014; Oakhill, Hartt, & Samols, 2005).

Based on these three central theories, we have defined four factors as key predictors of reading comprehension for the present study: decoding skills, vocabulary knowledge, morphosyntactic knowledge, and working memory skills. Empirical studies have shown that word decoding skills and vocabulary knowledge predict reading comprehension, particularly in the early stages (Cain, 2010; Catts, Hogan, & Adlof, 2005). Furthermore, morphosyntactic knowledge (Muter, Hulme, Snowling, & Stevenson, 2004) and working memory also explain unique variance in young children's reading comprehension (Cain, Oakhill, & Bryant, 2004; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000). However, these studies included only monolingual children; little is known about *early* reading comprehension in L2 readers, despite the detection of a performance gap between L1 and L2 readers in the early stages of reading development (Verhoeven, 2000; Verhoeven et al., 2018).

Reading Comprehension in Second Language Readers

At the start of primary education, L2 readers are confronted with the challenging task of learning to read in a language that is not their native language, which often

results in falling behind in reading comprehension (Melby-Lervåg & Lervåg, 2014). In relation to the four key predictors, previous studies have provided some relevant insights into reading comprehension in L2 readers. Their comprehension difficulties are usually not a consequence of poor decoding skills; L2 readers' decoding skills are found to be in the average range (Geva & Zadeh, 2006; Mancilla-Martinez & Lesaux, 2010), and they develop similar to L1 readers' decoding skills (Verhoeven, 2000). The main source of their comprehension difficulties seems to be related to their oral language skills; L2 readers have lower vocabulary knowledge than L1 readers and their morphosyntactic knowledge also falls behind (Melby-Lervåg & Lervåg, 2014). Some studies have shown that oral language skills are even more important factors for L2 reading comprehension (Babayiğit, 2014; Droop & Verhoeven, 2003; Geva & Farnia, 2012; Lervåg & Aukrust, 2010), whereas other studies report that those skills are equally important for L1 and L2 readers (Lesaux et al., 2006; Verhoeven & Van Leeuwe, 2012). Regarding their cognitive capacities, there is no reason to assume that L2 readers' working memory capacity limits their reading comprehension. Although the results are mixed, recent findings have even suggested that bilinguals have larger working memory capacity than monolinguals because of experience with dual language management (Grundy & Timmer, 2017). This advantage has mostly been observed in bilingual children (Blom, Küntay, Messer, Verhagen, & Leseman, 2014; Morales, Calvo, & Bialystok, 2013) and might enable them to compensate for their limits in oral language skills during comprehension processes.

Reading Comprehension Across the Distribution

To gain more insight into children's reading comprehension difficulties, comparisons between good and poor comprehenders have often been made (e.g., Cain & Oakhill, 2006; Lesaux et al., 2006; Nation, Cocksey, Taylor, & Bishop, 2010). Those studies included primarily monolingual children who were split up into two subgroups. However, by dichotomizing a continuous variable, such as reading comprehension, information about the distribution of scores gets lost, resulting in a loss of power. The one study that did include children from diverse linguistic backgrounds (Lesaux et al., 2006) used percentile scores to classify children as poor or good comprehenders, resulting in a limited sample size and large discrepancies. Quantile regression is an analysis technique recently applied to reading research that does not require the sample to be split up, but still allows for an investigation of (the strength of) the relation between predictor variables and the outcome variable across different points in the distribution of scores, so-called quantiles (Petcher & Logan, 2014).

By using a quantile regression approach, recent studies have revealed that explanatory factors can vary across the range of word reading ability (McIlraith & LRRC, 2016) and reading comprehension ability (Hua & Keenan; 2017, LRRC & Logan, 2017; Mancilla-Martinez & Lesaux, 2017; Tighe & Schatschneider, 2016). Hua and Keenan (2017) showed that for several reading comprehension tests, the contribution of word decoding skills and listening comprehension varied as a function of reading comprehension ability in younger and older readers (8-18 years). Furthermore, a study with third-grade children and a broader range of predictor variables (LRRC & Logan, 2017) showed that decoding skills were related to reading comprehension for poor and average comprehenders but not for good comprehenders, whereas vocabulary and morphosyntactic knowledge showed a similar relation across ability levels. It was concluded that quantile regression is a useful approach for the study of reading in both typical and atypical readers. With respect to the latter, there is one quantile regression study with only L2 readers showing that the relation between early language skills and later reading comprehension outcomes was stronger for L2 readers with higher ability levels (Mancilla-Martinez & Lesaux, 2017).

To our knowledge, there are no studies to date in which quantile regression analysis is used with a mixed group of L1 and L2 readers. Previous research has shown that L1 and L2 readers differ not only in reading comprehension ability, but also in oral language skills (Melby-Lervåg & Lervåg, 2014), and mixed results have been reported about whether the relation between oral language skills and reading comprehension differs for L1 and L2 readers (Babayiğit, 2014; Droop & Verhoeven, 2003; Geva & Farnia, 2012; Lervåg & Aukrust, 2010; Lesaux et al., 2006; Verhoeven & Van Leeuwe, 2012). Instead of investigating whether the influence of key predictors differs because of children's language background, quantile regression can be used to investigate whether the relation between key predictors and reading comprehension differs depending on children's ability level. Children's scores are placed along a continuum of reading comprehension ability regardless of their language background to do justice to individual variation in performance. Of particular interest are children at the lower end of the distribution, who should be considered for extra support. With quantile regression we can identify the key predictors for this particular group, which is essential for the identification and subsequent remediation of reading comprehension problems.

The Present Study

In the present study, we investigated the factors that explain individual variation in early reading comprehension performance in a mixed group of L1 and L2 readers of Dutch. Based on a variety of prominent theories, we defined decoding skills, vocabulary knowledge, morphosyntactic knowledge, and working memory skills as key predictors of reading comprehension. In contrast to most previous studies, we focused on beginning readers (second grade) based on the conviction that early intervention in reading comprehension problems is crucial for later school success. We applied two complementary approaches to answer our research questions. First, with classical regression, we investigated whether key predictors of reading comprehension differed depending on children's language background (L1 vs. L2). Second, with quantile regression, we investigated whether key predictors of reading comprehension differed depending on children's ability level (poor vs. average vs. good). In both types of analysis, we investigated whether language background explained unique variance over and above the key predictors. This allowed us to examine the role of linguistic diversity in the prediction of early reading comprehension performance across ability levels.

Regarding the first approach, we expected similar key predictors of reading comprehension for L1 and L2 readers, although we expected L2 readers to fall behind their L1 peers in terms of oral language skills and reading comprehension performance. We expected our second approach to be more relevant by showing that the relations between key predictors and reading comprehension vary across ability levels, as found in previous studies with monolingual or bilingual children. The relation with oral language skills was expected to be stronger for children with lower ability levels; delays in vocabulary and morphosyntactic knowledge will hinder their comprehension processes. If differences in oral language skills are not sufficient to explain the L1-L2 performance gap in reading comprehension, this is most likely to be true at the lower ability level.

Method

Participants

The participants were 161 second grade students (*M* age = 7.5 years, range = 81–105 months) from 13 primary schools in the Netherlands located in lower-middle class neighborhoods in urban areas. Prior to the study, ethical approval was obtained from our institution's ethics committee and parental (informed) consent was sought. Based on language background information provided by the teachers,

children were selected to participate. Because of our specific selection criteria and the large variation in the children's language background only a limited number of children per class were eligible for participation. Children with only Dutch as their home language were classified as L1 readers (n = 83, 51% boys) and children with Turkish as their home language were classified as L2 readers (n = 78, 64% boys). Since Turkish is the largest linguistic minority (CBS, 2016) and the most widely spoken additional language among Dutch primary school children (Extra, 2002), only Turkish-Dutch children were selected as L2 readers to maximize homogeneity. From a typological perspective, there is a large distance between Dutch and Turkish and Turkish has a more transparent orthography than Dutch (Extra & Verhoeven, 2011). The instructional language at all schools was Dutch and from first grade onwards all children received formal reading instruction in Dutch only with similar reading methods used across schools.

A questionnaire was used to determine the children's and their parents' native country and their language usage at home (see Table 2.1). All L2 readers indicated that they spoke Turkish at home, but there was some variation reflective of the population: For the majority (59%) Turkish was their main home language whereas others spoke Dutch to a larger extent (33%) or reported equal use of both languages (8%). Most children indicated that they spoke primarily Turkish with at least one parent, while they heard their parents speaking mainly Turkish to each other (82%). The L2 readers' Turkish language proficiency was confirmed by the fact that all L2 readers scored above chance level on a Turkish receptive vocabulary test (Blom, 2019) consisting of 96 multiple-choice items (A-D) (percentage correct: M = 61.7, SD = 10.6, *range* = 35.4–83.3, Cronbach's α = .86 in the current study). L2 readers who had Turkish as their dominant home language scored higher on Turkish vocabulary than L2 readers who had Dutch as their dominant home language ($t(70) = -2.60$, $p = .011$, $d = 0.62$).

The two language groups did not differ in age ($t(159) = 0.41$, $p = .686$, $d = 0.06$) and non-verbal reasoning (Raven, 1965; ($t(159) = 0.99$, $p = .325$ $d = 0.16$). Considering their parents' average level of education measured on a four-point scale (1 = primary school, 4 = higher education/university), the L1 readers had a significantly higher socio-economic status (SES) than the L2 readers ($U = 2360.00$, $z = -3.04$, $p = .002$, $r = -.24$) despite coming from the same classrooms. This corresponds to other studies with Dutch monolingual and Turkish-Dutch bilingual children (e.g., Blom et al., 2014; Droop & Verhoeven, 2003)

Table 2.1 Demographics for L1 Readers and L2 Readers: Country of Birth and Home Language

	L1 readers (*n* = 83)		L2 readers (*n* = 78)	
	Frequency	Percentage	Frequency	Percentage
Country of birth:				
- The Netherlands (NL)	82	98.8%	72	92.3%
- Turkey	0	0.0%	4	5.1%
- Other	1	1.2%	2	2.6%
Parents' county of birth:				
- Both parents born in NL	59	71.1%	6	7.7%
- One parent born in NL	16	19.3%	30	38.4%
- Neither parent born in NL	7	8.4%	39	50.0%
Home language:				
- Only Dutch	83	100.0%	0	0.0%
- Turkish and Dutch	0	0.0%	76	97.4%
- Only Turkish	0	0.0%	2	2.6%

Measures

Decoding skills. The children's decoding skills were assessed with a word reading test and a pseudoword reading test (Verhoeven & Keuning, 2017), each consisting of three lists with an increasing number of syllables (e.g., *boek*, *blaas*, *dromer / laas*, *stoef*, *gluifel)*. For each list, the children were asked to read aloud the (pseudo)words as quickly and accurately as possible within one minute. The children's score was the total number of correctly read (pseudo)words across the three lists (Cronbach's $\alpha = .95–.97$; Verhoeven & Keuning, 2017). On average, the children read more words than pseudo words ($t(160) = 22.32$, $p < .001$, $d = 1.76$) but their scores on both tests were highly correlated ($r = .95$, $p < .001$).

Vocabulary knowledge. The children completed two standardized tests for receptive and expressive vocabulary ($r = .68$, $p < .001$ in the current study). The 64-item receptive vocabulary test (T-TOS; Verhoeven, Keuning, Horsels, & Van Boxtel, 2013) was presented on a laptop. For each item, a word (noun, verb, or adjective) was presented auditorily along with four pictures on the screen (e.g., (in Dutch) *juggler*, *to sketch*, *muscular*). The children had to select the picture that matched the spoken word. The test was terminated after eight consecutive mistakes. A child's score was the number of items answered correctly (Cronbach's $\alpha = .88$, Verhoeven et al., 2013).

Expressive vocabulary was measured with the 35-item word knowledge task of the WISC-III-NL (Kort et al., 2005). Each item consisted of a word that the children were asked to describe orally (e.g., (in Dutch)*'What is a fairy tale?'*). The items increase in difficulty. The children's answers were scored as either 2 (correct), 1 (partly correct), or 0 points (incorrect or no answer). The test was terminated after four consecutive mistakes (0 points). A child's score was the total number of points (Cronbach's α = .83, Kort et al., 2005).

Morphosyntactic knowledge. The children completed two subtests (r = .58, p < .001 in the current study) from the T-TOS (Verhoeven et al., 2013). The morphology subtest includes 36 items assessing children's ability to apply four morphological rules in Dutch: inflexion of plural form of nouns, comparative/superlative of adjectives, inflexion of past tense, and past participle. The items were orally presented supported by images (e.g., (in Dutch) *'This bag is heavy, but that bag is even … (heavier), and that one is the … (heaviest)'*).

The receptive syntax subtest assesses children's knowledge of syntactic patterns and elements that determine the meaning of a sentence: word order, prepositions, grammatical roles, pronouns, quantifiers, passive, post-modified subject, ellipsis, and relative clause. The task, presented on a laptop, consists of 33 items. For each item, four pictures were presented, and the children had to select the picture that corresponded with a sentence they heard (e.g., (in Dutch) *'The dog is bitten by the cat.')*. For both subtests, the score was determined by the number of items answered correctly (Cronbach's α = .87/.82, Verhoeven et al., 2013).

Working memory. The children completed the Digit Span Backwards (WISC-NL-III; Kort et al., 2005). In this standardized task, children first listen to a sequence of digits and then repeat the sequence backwards (e.g., *3-8-6; 6-8-3*). Each item consists of two sequences of the same length, starting with two-digit sequences. When a child answered at least one sequence of an item correctly, the next item was presented with a longer sequence (one extra digit). The task stopped when a child gave an incorrect answer to both sequences of one item. The score was determined by the number of sequences repeated correctly (Cronbach's α = .64, Kort et al., 2005).

Reading comprehension. The children completed three standardized paper-and-pencil tests for reading comprehension (r = .46–.58, p < .001 in the current study). For children who completed only two out of three tests (n = 12; 7.5%), the

composite score was based on two tests instead of three. Two tests specifically aimed at comprehension of semantic relations and anaphoric relations (Verhoeven, 1992) and consisted of 27 multiple-choice items (A-D) and 24 multiple-choice items (A-D), respectively. The children's scores corresponded to the number of items answered correctly for each test (Cronbach's α = .75 for both tests in the current study). The third test was the CITO test (Cito B.V., 2014), which is a more general measure of reading comprehension. This standardized test is part of the Dutch national student monitoring system in primary education. This test consisted of 40 multiple-choice questions (A-D), including a variety of text and question types (Cronbach's α = .86; Jolink, Tomesen, Hilte, Weekers, & Engelen, 2015). A child's score was a norm-based score determined from the number of items answered correctly.

Procedure

The data were collected during the first half of the second-grade school year by the first author and three trained research assistants. The children performed several tasks in a fixed order during two individual sessions of 30 minutes each. The sessions occured at school during school hours. The reading comprehension tests were group administered in class by the teacher. The children completed these tests silently without a time limit.

Data Analyses

Except for working memory, all factors were assessed by at least two measures. Therefore, the average z-score on both tests was calculated for all factors (based on the present sample), and these composite scores were used in the analyses. Z-scores were calculated for working memory. First, we performed independent samples t-tests to check for between-group differences with the linear step-up procedure (Benjamini & Hochberg, 1995) to control for multiple testing. To investigate whether key predictors of reading comprehension differed depending on the children's language background, we performed an ordinary least squares (OLS) multiple hierarchical regression analysis, also referred to as classical regression. The variables were included stepwise: 1) four key predictors; 2) language background; and 3) interaction terms between the key predictors and language background. Although there was a relatively strong correlation between vocabulary and morphosyntactic knowledge (r = .78, p < .001), the assumption of no multicollinearity was not violated (largest VIF = 3.1, average VIF = 1.9, tolerance statistics > 0.3).

To investigate whether key predictors of reading comprehension differed depending on children's ability, we performed quantile regression analyses across multiple points in the distribution. In addition to the four key predictors, language background was included as an extra predictor variable to investigate whether it was uniquely related to reading comprehension across the distribution. The relations between predictor variables and the outcome variable were estimated at specific points (i.e., quantiles) along the distribution of the outcome variable. Given that the scores for reading comprehension were not completely normally distributed but skewed to the left (-.66), quantile regression can provide more accurate statistical estimates than OLS regression. There was no loss of power; all data points contributed to the estimation of these quantiles based on the proximity of each data point by means of a weighting matrix (Petscher & Logan, 2014). Based on our sample size, the number of parameters and the distribution of the data (Cade & Noon, 2003) we selected three quantiles, 0.25, 0.50, and 0.75, as an indication of three ability levels, poor (P), average (A) and good (G) comprehenders, respectively. First, we conducted individual quantile regression analyses to investigate the relation across the three quantiles for each predictor variable separately. Second, we performed a multiple quantile regression analysis including all predictor variables simultaneously to investigate unique relations across the distribution. Quantile regression analyses were performed in the R computing environment (version 3.4.2; R Core Team, 2017) using the quantreg package (version 5.34; Koenker, 2017).

Results

Group Differences for Predictor Variables and Reading Comprehension

The means and standard deviations for all measures (raw scores) and composite scores are shown in Table 2.2. There were no group differences for decoding skills or working memory, but the L2 readers scored significantly lower on vocabulary, morphosyntactic knowledge, and reading comprehension with medium to large effect sizes (Cohen, 1988). The L1-L2 performance gap in reading comprehension still existed after controlling for differences in SES ($F(1, 158) = 22.58$, $p < .001$, $\eta_p^2 = .13$).

Table 2.2 Descriptive Statistics and Mean Differences for L1 Readers and L2 Readers

	L1 readers (*n* = 83)		L2 readers (*n* = 78)		Mean differences		
Measures/min-max	*M (SD)*	*range*	*M (SD)*	*range*	*t*	*p*	*d*
DEC							
Words[a] / 0–420	136.9 (54.8)	37–309	130.5 (47.9)	40–245	0.80	.428	0.12
Pseudo words[a] / 0–420	100.6 (43.0)	28–266	101.9 (37.8)	22–199	-0.20	.840	0.03
Composite score[b]	0.0 (1.1)		0.0 (0.9)		0.30	.765	0.04
VOCAB							
Receptive[a] / 0–64	44.1 (5.5)	27–56	33.3 (7.1)	10–48	10.84	< .001*	1.70
Productive[a] / 0–70	23.0 (6.1)	11–36	16.6 (4.5)	6–28	7.67	< .001*	1.19
Composite score[b]	0.6 (0.7)		-0.6 (0.7)		10.43	< .001*	1.64
MSK							
Morphology[a] / 0–36	25.7 (4.5)	14–35	18.9 (5.2)	3–32	8.86	< .001*	1.40
Syntax[a] / 0–33	29.1 (2.7)	19–33	26.3 (3.5)	15–31	5.65	< .001*	0.90
Composite score[b]	0.5 (0.7)		-0.5 (0.8)		8.37	< .001*	1.32
WM[a] / 0–14	3.5 (1.1)	1–6	3.4 (1.1)	1–6	0.92	.358	0.15
RC							
Semantic rel.[a] / 0–27	23.3 (2.6)	14–27	21.0 (4.1)	9–27	4.24	< .001*	0.67
Anaphoric rel.[a] / 0–24	20.2 (3.1)	6–24	18.6 (3.9)	9–24	2.74	.007*	0.45
CITO test[a] / 0–252	137.7 (25.8)	84–210	114.7 (21.3)	70–190	6.13	< .001*	0.97
Composite score[b]	0.3 (0.7)		-0.4 (0.9)		5.51	< .001*	0.87

Note. DEC = decoding skills, VOCAB = vocabulary knowledge, MSK = morphosyntactic knowledge, WM = working memory, RC = reading comprehension.
* mean difference is significant based on critical p-values from linear step-up procedure (Benjamini & Hochberg, 1995) to correct for multiple testing.
[a] Values are based on raw scores.
[b] Values are based on composite scores, calculated by the average z-score on both tests.

Table 2.3 Pearson Correlations across Variables based on Composite Scores for Total Sample (*N*=161)

	DEC	VOCAB	MSK	WM	RC
DEC	–				
VOCAB	.05	–			
MSK	.05	.78***	–		
WM	.19*	.17*	.19*	–	
RC	.23**	.62***	.67***	.24**	–

Note. DEC = decoding skills, VOCAB = vocabulary knowledge, MSK = morphosyntactic knowledge, WM = working memory, RC = reading comprehension.
*p < .05, **p < .01, ***p < .001

Table 2.4 OLS Multiple Regression Analysis for Early Reading Comprehension Performance

	R^2	*B*	*SE*	β
Step 1	.513***	.16	.05	.18**
Decoding skills		.21	.08	.23*
Vocabulary knowledge		.44	.09	.46***
Morphosyntactic knowledge		.06	.05	.08
Working memory				
Step 2	.000			
Decoding skills		.16	.05	.18**
Vocabulary knowledge		.22	.09	.24*
Morphosyntactic knowledge		.44	.09	.47***
Working memory		.06	.05	.08
Language background		.03	.12	.02
Step 3	.019			
Decoding skills		.09	.06	.11
Vocabulary knowledge		.36	.11	.39**
Morphosyntactic knowledge		.26	.13	.27*
Working memory		.11	.07	.13
Language background (LB)		-.01	.12	-.01
Decoding skills x LB		.12	.10	.09
Vocabulary knowledge x LB		-.33	.19	-.22
Morphosyntactic knowledge x LB		.35	.18	.26
Working memory x LB		-.10	.10	-.08

*$p < .05$, **$p < .01$, ***$p < .001$

Prediction of Reading Comprehension for L1 readers and L2 readers

The correlations among the five factors are shown in Table 2.3. Reading comprehension was positively correlated with the other factors; there was a weak correlation with decoding and working memory, and a moderate to strong correlation with vocabulary and morphosyntactic knowledge.

Next, we performed an OLS multiple hierarchical regression analysis to investigate the unique predictors of reading comprehension and to ascertain whether the influence of those predictors differed for L1 and L2 readers. The results (see Table 2.4) showed that decoding skills, vocabulary, and morphosyntactic knowledge were uniquely related to reading comprehension, but working memory was not ($F(4, 156) = 41.15$, $p < .001$, $R^2 = .513$). Language background (Step 2: $R^2 = .000$, $p = .799$) and the interaction terms (Step 3: $R^2 = .019$, $p = .186$) did not explain unique variance, indicating that the relation between key predictors and reading comprehension did not differ for the two groups. Adding SES as a control variable did not alter the

results and SES did not explain unique variance in reading comprehension (β = .04, p = .579).

Prediction of Reading Comprehension Across the Distribution

Figure 2.1 shows the distribution of reading comprehension scores split by language background. On average, L1 and L2 readers differed in reading comprehension performance (see Table 2.2), but there was considerable variation within both groups (see Figure 2.1). Both L1 readers and L2 readers were represented in all four quartiles, with an overrepresentation of L2 readers in lower quartiles and an overrepresentation of L1 readers in higher quartiles.

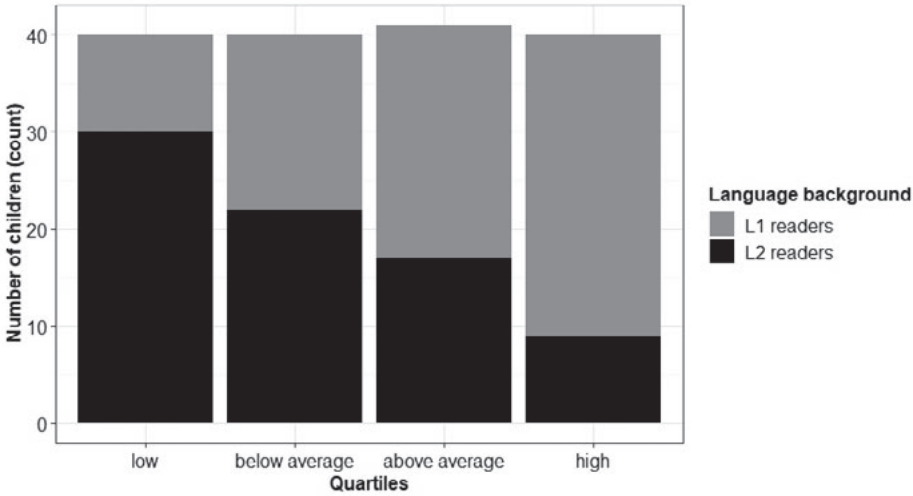

Figure 2.1 Distribution of reading comprehension performance across quartiles (low = 0-25%, below average = 25-50%, above average = 50-75%, high = 75-100%) split by language background (L1 readers vs. L2 readers).

Individual quantile regression analyses. The results of the individual quantile regression analyses are shown in Table 2.5 (column 1-3). All estimates were statistically different from zero (i.e., the confidence intervals did not include zero), apart from decoding skills for the higher quantile (0.75). This indicates that vocabulary, morphosyntactic knowledge, working memory and language background were related to reading comprehension across the distribution. By contrast, decoding was related to reading comprehension for poor and average comprehenders, but not for good comprehenders.

To ascertain whether the relation of each predictor was stronger at one point in the distribution than another, we calculated statistical significance tests of between-quantile slopes (see Table 2.5 column 4-6). Again, we used the linear step-up procedure (Benjamini & Hochberg, 1995) to correct for multiple testing across quantiles. The comparisons showed that decoding was significantly stronger related to reading comprehension at the lower and mediate quantiles compared to the higher quantile. Furthermore, vocabulary knowledge was significantly stronger related to reading comprehension at the lower quantile compared to the mediate and higher quantiles.

A similar pattern was found for language background; however the interpretation of a dichotomous predictor is substantially different from a continuous predictor (Petcher & Logan, 2014). The quantile regression coefficients (i.e., slope coefficients) are indicative of the performance gap in reading comprehension between L2 readers (dummy code = 1) as compared to L1 readers (dummy code = 0). The L1-L2 performance gap in reading comprehension performance was not uniform across the distribution: At the lower quantile, the performance gap was 0.99, and decreased to 0.56 at the mediate quantile and 0.38 at the higher quantile. The results of the between-quantile slope differences showed that the L1-L2 performance gap was significantly larger at the lower quantile compared to the mediate and higher quantiles. For morphosyntactic knowledge and working memory, no significant differences in slope coefficients were found between the three quantiles.

Table 2.5 Individual Quantile Regression Coefficients and Comparisons of Coefficients for Poor (P), Average (A), and Good (G) Comprehenders

	Quantile Regression Coefficients [95% CI]			*p*-value of comparisons		
	0.25 (P)	0.50 (A)	0.75 (G)	P vs. A	P vs. G	A vs. G
DEC	0.36* [0.07, 0.42]	0.22* [0.03, 0.30]	0.06 [-0.03, 0.21]	.189	**.010**[a]	**.001**[a]
VOC	0.68* [0.49 0.77]	0.51* [0.37, 0.64]	0.45* [0.34, 0.58]	**.029**[a]	**.013**[a]	.444
MSK	0.72* [0.60, 0.77]	0.63* [0.56, 0.72]	0.57* [0.44, 0.76]	.256	.107	.322
WM	0.22* [0.13, 0.40]	0.20* [0.05, 0.39]	0.19* [0.09, 0.32]	.811	.761	.857
LB	-0.99* [-1.34, -0.62]	-0.56* [-0.84, -0.42]	-0.38* [-0.55, -0.26]	**.024**[a]	**.007**[a]	.188

Note. DEC = decoding skills, VOC = vocabulary knowledge, MSK = morphosyntactic knowledge, WM = working memory, LB = language background.
* Coefficient is significant ($p < .05$) because the confidence interval does not include zero.
[a] between-quantile difference is significant based on critical p-values from linear step-up procedure (Benjamini & Hochberg, 1995) to correct for multiple testing.

Multiple quantile regression analysis. Finally, we conducted multiple quantile regression analysis, because we were especially interested in how each variable was uniquely related to reading comprehension across the three quantiles after controlling for the other variables. The results of the multiple quantile regression analysis are shown in Table 2.6 (column 2-4), which can be compared to the regression coefficients of the OLS multiple regression analysis (column 1) retrieved from Table 2.4 Step 2. Decoding and morpho-syntactic knowledge were uniquely related to reading comprehension across the distribution, indicating that these two constructs are important for reading comprehension regardless of ability level. By contrast, differential relations were found for vocabulary and working memory. Vocabulary knowledge was significant only at the lower quantile, while working memory was significant only at the higher quantile. Language background was not related to reading comprehension across the distribution after controlling for the other predictor variables.

Again, adding SES showed no unique effect at any of the quantiles, and the results hardly changed; the confidence intervals for working memory just cross the zero boundary [95% CI = -0.01, 0.21] only at the higher quantile (0.75). This is probably due to limited power in the extreme of the distribution as working memory was still significant when the higher quantile was set to 0.70 [95% CI = 0.01, 0.18].

Table 2.6 Multiple OLS Regression and Multiple Quantile Regression Analyses for Poor (P), Average (A), and Good (G) Comprehenders

	OLS [95% CI]	Quantile regression [95% CI]		
		0.25 (P)	0.50 (A)	0.75 (G)
Intercept	-0.03 [-0.18, 0.12]	-0.34 [-0.57, -0.22]	-0.07 [-0.18, 0.15]	0.30 [0.19, 0.56]
DEC	0.16* [0.06, 0.25]	0.16* [0.03, 0.23]	0.16* [0.05, 0.26]	0.11* [0.02, 0.21]
VOC	0.22* [0.04, 0.40]	0.23* [0.08, 0.39]	0.19 [-0.11, 0.35]	0.09 [-0.12, 0.36]
MSK	0.44* [0.27, 0.61]	0.51* [0.33, 0.70]	0.50* [0.30, 0.66]	0.55* [0.27, 0.72]
WM	0.06 [-0.03, 0.16]	0.08 [-0.02, 0.19]	0.10 [-0.01, 0.20]	0.09* [0.02, 0.18]
LB	0.03 [-0.21, 0.27]	-0.02 [-0.22, 0.26]	0.12 [-0.17, 0.34]	0.09 [-0.26, 0.44]

Note. DEC = decoding skills, VOC = vocabulary knowledge, MSK = morphosyntactic knowledge, WM = working memory, LB = language background.
* Coefficient is significant ($p < .05$) because the confidence interval does not include zero.

To examine the explained variance across quantiles, we calculated *pseudo* R^2 (Petscher, Logan, & Zhou, 2013). The amount of explained variance was the largest at the lower quantile (37.2%) and somewhat lower at the mediate quantile (31.1%) and higher quantile (26.2%).

Discussion

We aimed to gain insight into the factors that explain individual variation in early reading comprehension. We used two complementary approaches to investigate whether the key predictors of reading comprehension differed for L1 readers and L2 readers (considering the distribution of scores as a whole) and/or across ability levels (considering specific points across the distribution of scores). The main findings of the present study are discussed in relation to previous studies, and the theoretical and practical implications are described.

Early Reading Comprehension in L1 readers and L2 Readers

Our findings showed that even at the very start of reading acquisition, there was a performance gap in reading comprehension between L1 and L2 readers (Melby-Lervåg & Lervåg, 2014; Verhoeven, 2000; Verhoeven et al., 2018), and that L2 readers scored considerably lower on oral language skills (i.e., vocabulary and morphosyntactic knowledge), whereas their decoding skills and working memory capacity were comparable with L1 readers. A closer look at the distribution of the children's reading comprehension scores revealed that there was considerable variation within both groups, and there was overlap between L1 and L2 readers' score distribution; L1 and L2 readers were both represented in the lower and higher ends of the distribution. This finding corresponds with Lesaux et al. (2006), who showed similar percentiles for L1 and L2 readers when classifying children as poor or good comprehenders. Moreover, it stresses the relevance of investigating the role of linguistic diversity *across* the distribution of reading comprehension performance.

Regarding the first approach, classical regression showed that the predictor variables for early reading comprehension performance were similar for L1 and L2 readers. Decoding skills, vocabulary, and morphosyntactic knowledge were identified as unique predictors of early reading comprehension performance, irrespective of language background. Considering the beta values, vocabulary and morphosyntactic knowledge were the most important factors for children's reading comprehension, whereas decoding skills were less important. The latter

is inconsistent with previous studies showing that early reading comprehension is strongly related to decoding skills for young monolingual readers of English (Cain, 2010; Catts et al., 2005; Lonigan & Burgess, 2017).

One explanation for this discrepancy relates to differences in language transparency across studies. The children in our study read in Dutch, which is more transparent than English and may result in earlier mastery of decoding accuracy (Ellis et al., 2004; Verhoeven & Van Leeuwe, 2009). Consequently, children's decoding skills may play a less prominent role in early reading comprehension for transparent orthographies, as found by Florit and Cain's meta-analysis (2011) on the Simple View of Reading in different types of orthographies. Another explanation relates to our mixed sample. Half of our sample consisted of L2 readers, who often have reading comprehension difficulties that probably stem from delays in oral language skills rather than poor decoding skills (Melby-Lervåg & Lervåg, 2014). In that respect, some studies have even shown that there is a stronger relation between oral language and reading comprehension for L2 readers as compared to L1 readers (Babayiğit, 2014; Droop & Verhoeven, 2003; Geva & Farnia, 2012; Lervåg & Aukrust, 2010). However, the current study is consistent with a more balanced view in which the importance of the underlying constructs is independent of children's language background (Lesaux et al., 2006; Verhoeven & Van Leeuwe, 2012).

Early Reading Comprehension across Ability Levels

Regarding the second approach, quantile regression demonstrated that the relation between key predictors and reading comprehension differed depending on the children's ability level as expected based on other recent studies using quantile regression (Hua & Keenan; 2017, LRRC & Logan, 2017; Mancilla-Martinez & Lesaux, 2017; Tighe & Schatschneider, 2016). Our findings showed that decoding skills and morphosyntactic knowledge were uniquely related to reading comprehension across the ability range, but vocabulary knowledge was related to reading comprehension for poor comprehenders only, whereas working memory was related to reading comprehension for good comprehenders only. This suggests that poor comprehenders may have particular difficulty with lower-level comprehension skills (i.e., vocabulary knowledge), which hinders their comprehension processes, whereas good comprehenders rely less on lower-level skills and thus have more cognitive capacity available for higher-level comprehension skills (i.e., skills requiring working memory) which benefits their comprehension processes.

In terms of age, it seems most relevant to compare our findings to the study of LRRC and Logan (2017) who also used quantile regression to investigate pressure points in reading comprehension in young readers. LRRC and Logan (2017) found similar results for (morpho)syntactic knowledge (i.e., the strongest unique predictor across the distribution), but they showed reversed results for decoding skills and vocabulary; vocabulary knowledge was a consistently significant predictor across ability levels, whereas decoding was only a significant predictor at the lower level.

Two explanations can be proposed related to our sample. First, the children in our sample were, on average, one year younger and less experienced in reading. This may have led to a more prominent role for decoding skills across the distribution as shown by the multiple quantile regression analysis, although the individual quantile regression analysis showed that this relation was significantly weaker (and non-significant) for good comprehenders. This discrepancy between analyses may be due to the complex nature of unique and shared effects in multiple regression, where the magnitude of unique effects is influenced by the extraneous variance of the other predictor variables included (Lindenberger & Pötter, 1998). Since we were mainly interested in the *unique* effects of the key predictors, greater value was attached to multiple quantile regression for the interpretation of the results. Second, our sample included L1 readers as well as L2 readers, with an overrepresentation of L2 readers at the lower end of the distribution. Given that L2 readers often show delays in terms of vocabulary knowledge in the target language (Melby-Lervåg & Lervåg, 2014), this may explain the more prominent role of vocabulary knowledge at the lower ability level. In other words, our findings suggest that vocabulary knowledge is the main obstacle, or pressure point, for beginning readers with lower reading comprehension performance.

Apart from vocabulary knowledge, the relation between working memory and reading comprehension also differed across ability levels. In contrast to Cain et al. (2004), who studied 8-to-11-year-olds, our classical regression analysis did not show a unique relation between working memory and reading comprehension. By contrast, our quantile regression analysis revealed that this relation seems to exist for higher performing children in our sample. We thus tentatively suggest that in the early stages of reading comprehension, particularly good comprehenders engage in higher-level comprehension processes, as these require working memory (Kendeou et al., 2014). This also relates to the lower amount of explained variance at the higher level; possibly other higher-order skills not included in the present

study, such as inference making and comprehension monitoring, may be more important for good comprehenders (Cain et al., 2004).

The Role of Linguistic Diversity in Early Reading Comprehension

In both approaches, we also investigated whether children's language background (L1 vs. L2) adds to the prediction of reading comprehension over and above the key predictors. Although language background was related to reading comprehension across the ability range (as shown by the individual quantile regression), in both the classical and quantile multiple regression analyses, language background did not explain additional variance over and above the variance accounted for by the key predictors. This was even the case at the lower ability level, where the performance gap between L1 and L2 readers was largest. Hence, the current study demonstrated that the prediction of reading comprehension is similar for beginning L1 and L2 readers because the key predictors (i.e., decoding skills, vocabulary and morphosyntactic knowledge) were equally important for L1 and L2 readers and individual differences in these skills were sufficient to explain the L1-L2 performance gap in early reading comprehension across the distribution.

This contrasts with Lesaux et al. (2006), who found that after controlling for several linguistic and cognitive predictors, children's language status (ESL vs. EL1) still explained unique variance in fourth-grade reading comprehension performance. However, Lesaux et al. (2006) did not control for decoding skills and vocabulary knowledge, which are two central components of reading comprehension. Not including vocabulary knowledge, which is one of the most prominent differences between monolingual and bilingual children (Bialystok, Luk, Peets & Yang, 2010), is likely to account for the additional variance explained by language background in Lesaux et al. (2006). In line with our findings, Babayiğit (2014) also found that individual differences in oral language skills (i.e., vocabulary and morphosyntactic skills) explained the differences in fifth-grade reading comprehension between L1 and L2 readers of English.

Limitations and Future Directions

The present study has some limitations that require careful interpretation of the results. First, we only included L2 readers from the largest linguistic minority to maximize the homogeneity of the L2 group. Second, we were not able to include L2 readers' native language skills as was done by Mancilla-Martinez and Lesaux (2017), because we conducted our analyses with a mixed group of L1 and L2 readers. Nevertheless, in our study, L2 reading comprehension was not related to

L1 vocabulary (r = .18, p = .118), and having either Turkish or Dutch as the dominant home language did not affect L2 reading comprehension scores ($t(70)$ = 0.18, p = .861, d = 0.04). Third, we used cross-sectional data rather than longitudinal data, which means that we cannot draw conclusions about the prediction of reading comprehension over time.

Nevertheless, the current study demonstrates the relevance of investigating key predictors of reading comprehension beyond the average relation and across the distribution of reading comprehension performance. In future research the use of a quantile regression approach should be extended in several ways: 1) by including another or more heterogeneous group of L2 readers; 2) by including a broader range of explanatory factors, such as socio-linguistic (e.g., quantity/quality of language usage) or meta-cognitive factors (e.g., comprehension monitoring), and 3) by including a larger sample to allow for latent variable modeling and reliable detection of interaction effects.

Conclusions and Implications

This study has provided new insights into the factors that explain individual variation in early reading comprehension in a mixed group of L1 and L2 readers. By combining classical and quantile regression analyses, we demonstrated that the explanatory factors of reading comprehension are more likely to vary as a function of children's ability level than their language background. By extending the use of quantile regression to a linguistically diverse group of beginning readers, this study adds to previous research showing that relations between key predictors and reading comprehension can vary across the distribution, even at the very beginning of reading acquisition, where intervention should ultimately start. Our findings revealed that children with lower vocabulary knowledge are particularly at risk for reading comprehension problems rather than L2 readers as a group. Furthermore, our study indicated that in the early stages of reading development, the key predictors of reading comprehension are similar for L1 and L2 readers, although, on average, L2 readers showed a delay with respect to reading comprehension and some of the key predictors, such as vocabulary and morphosyntactic knowledge.

The conclusion that the prediction of reading comprehension is more universal than expected and that the role of linguistic diversity is less prominent than assumed has implications for educational practice. Educational professionals and teachers should be cautious about classifying or selecting children on the basis of their language background, because categorizing children as either monolingual

or bilingual does not do justice to the individual variation across children (Luk & Bialystok, 2013) and may be off target when it comes to the differentiation or allocation of extra support in education. Instead, children's language and literacy skills should guide teachers (and policy makers) to meet individual needs in the classroom.

References

Babayiğit, S. (2014). The role of oral language skills in reading and listening comprehension of a text: A comparison of monolingual (L1) and bilingual (L2) speakers of English Language. *Journal of Research in Reading*, *37*, 22-47. doi:10.1111/j.1467-9817.2012.01538.x

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189-208. doi:10.1016/S0021-9924(03)00019-4

Benjamini, Y., & Hochberg, Y. (1995). Controlling for the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistics Society*, *57*, 289-300.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. Bilingualism: Language and Cognition, 13, 525-531. doi:10.1017/S1366728909990423

Blom, E. (2019). Domain-general cognitive ability predicts bilingual children's receptive vocabulary in the majority language. *Language Learning*, *69*, 292-322. doi:10.1111/lang.12333

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish-Dutch children. *Journal of Experimental Child Psychology*, *128*, 105-119. doi:10.1016/j.jecp.2014.06.007

Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, *1*, 412–420. doi:10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2

Cain, K. (2010). *Reading development and difficulties*. Chichester, UK: BPS Blackwell.

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, *76*, 683-696. doi:10.1348/000709905X67610

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31-42. doi:10.1037/0022-0663.96.1.31

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25-40). Mahwah, NJ: Lawrence Erlbaum Associates.

CBS Statistics the Netherlands (2016). Bevolking naar migratieachtergrond [Population by immigration background]. Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Cito B.V. (2014). *Begrijpend lezen 3.0 voor groep 4: M4-toets [Reading Comprehension 3.0 for second grade: M4 test]*. Arnhem, the Netherlands: Cito B.V.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, *38*, 78-103. doi:10.1598/RRQ.38.1.4

Dunn, D. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition, manual*. Minneapolis, MN: NCS Pearson, Inc.

Ellis, N. C., Natsume, I., Stavropoulou, K., Hoxhallari, L., Van Daal, V. H. P., Polyzoe, N., Petalas, M. (2004). The effects of the orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. *Reading Research Quarterly*, *39*, 438–468. doi:10.1598/RRQ.39.4.5

Extra, G. (2002). *De andere talen van Nederland: Thuis en op school. [The Netherlands'other languages: At home and at school]*. Bussum, the Netherlands: Coutinho.

Extra, G., & Verhoeven, L. (2011). Processes of language change in a migration context. In G. Extra and L. Verhoeven (Eds.), *Bilingualism and Migration* (pp.29-57). Berlin, Germany: Walter de Gruyter.

Florit, E., & Cain, K. (2011). The Simple View of Reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*, 553-576. doi:10.1007/s10648-011-9175-6

Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading and Writing: An Interdisciplinary Journal*, *25*, 1819-1845. doi:10.1007/s11145-011-9333-8

Geva, E., & Zadeh, Z. Y. (2006). Reading efficiency in native English-speaking and English as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading*, *10*, 31-57. doi:10.1207/s1532799xssr1001_3

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, *7*, 6-10. doi:10.1177/074193258600700104

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*, 325-340. doi:10.1177/0267658316678286

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, *2*, 127-160. doi:10.1007/BF00401799

Hua, A. N., & Keenan, J. M. (2017). Interpreting reading comprehension test results: Quantile regression shows that explanatory factors can vary with performance level. *Scientific Studies of Reading*, *21*, 225-238. doi:10.1080/10888438.2017.1280675

Jolink A., Tomesen M., Hilte M., Weekers A., & Engelen R. (2015). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 4 [Scientific justification for reading comprehension test 3.0 for second grade]*. Arnhem, the Netherlands: Cito B.V.

Kendeou, P., Van den Broek, P., Helder, A, & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, *29*, 10-16. doi:10.1111/ldrp.12025

Koenker, R. (2017). Quantreg: Quantile regression. R package version 5.34. Retrieved from https://CRAN.R-project.org/package=quantreg

Kort, W., Schittekatte, M., Dekker, P. H., Verhaeghe, P., Compaan, E. L., Bosmans, M., & Vermeir, G. (2005). *WISC-III NL. Handleiding en Verantwoording* [Manual and Justification]. In D. Wechsler (Ed.), London, UK: The Psychological Corporation.

Language and Reading Research Consortium (LRRC), & Logan, J. (2017). Pressure points in reading comprehension: A quantile multiple regression analysis. *Journal of Educational Psychology*, *109*, 451-464. doi:10.1037/edu0000150

Lervåg, A. & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*, 612-620. doi:10.1111/j.1469-7610.2009.02185.x

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, *19*, 99-131. doi:10.1007/s11145-005-4713-6

Lindenberger, U., & Pötter, U. (1998). The complex nature of unique and shared effects in hierarchical linear regression: Implications for developmental psychology. *Psychological Methods*, *3*, 218-230.

Lonigan, C. J., & Burgess, S. R. (2017). Dimensionality of reading skills with elementary-school-aged children. *Scientific Studies of Reading*, *21*, 239-253. doi:10.1080/10888438.2017.1285918

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*, 605-621. doi:10.1080/20445911.2013.795574

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority children. *Journal of Educational Psychology*, *102*, 701-711. doi:10.1037/a0019135

Mancilla-Martinez, J., & Lesaux, N. K. (2017). Early indicators of later English reading comprehension outcomes among children from Spanish-speaking homes. *Scientific Studies of Reading*, *21*, 428-448. doi:10.1080/10888438.2017.1320402

McIlraith, A. L., & Language and Reading Research Consortium (LRRC) (2016). Predicting word reading ability: A quantile regression study. *Journal of Research in Reading*. doi:10.1111/1467-9817.12089

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409-433. doi:10.1037/a0033890

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*, 187-202. doi:10.1016/j.jecp.2012.09.002

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as founding of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*, 663-681. doi:10.1037/0012-1649.40.5.665

Nation, K., Cocksey, J., Taylor, J., & Bishop, D. (2010). A longitudinal investigation of the early language and reading skills in children with reading comprehension impairment. *Journal of Child Psychology & Psychiatry*, *51*, 1031–1039. doi:10.1111/j.1469-7610.2010.02254.x

Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing: An Interdisciplinary Journal*, *18*, 657-686. doi:10.1007/s11145-005-3355-z

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

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22-37. doi:10.1080/10888438.2013.827687

Petscher, Y., & Logan, J. A. R. (2014). Quantile regression in the study of developmental sciences. *Child Development*, *85*, 861-881. doi:10.1111/cdev.12190

Petscher, Y., Logan, J. A. R., & Zhou, C. (2013). Extending conditional means modeling: An introduction to quantile regression. In Y. Petscher, C. Schatschneider, & D. L. Compton (Eds.). *Applied quantitative analysis in education and social sciences* (pp. 3–33). New York, NY: Routledge.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Raven, J. C. (1965). *Guide to using coloured progressive matrices, Sets A, AB and B*. London, UK: BHK Levis & Co.

Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory and children's reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *13*, 81-103. doi:10.1023/A:1008088230941

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*, 32-71. doi:10.2307/747348

Tighe, E. L., & Schatschneider, C. (2016). A quantile regression approach to understanding the relations among morphological awareness, vocabulary and reading comprehension in adult basic education students. *Journal of Learning Disabilities*, *49*, 424-436. doi:10.1177/0022219414556771

Verhoeven, L. (1992). *Lezen met Begrip 1* [Reading for Understanding 1]. In L. Verhoeven (Ed.). Arnhem, the Netherlands: Cito.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*, 313-330. doi:10.1207/S1532799XSSR0404_4

Verhoeven, L., & Keuning, J. (2017). The nature of developmental dyslexia in a transparent orthography. *Scientific Studies of Reading*, 1-17. doi:10.1080/10888438.2017.1317780

Verhoeven, L. T. W., Keuning, J., Horsels, L., Van Boxtel, H. (2013). *Testinstrumentarium Taalontwikkelingsstoornissen T-TOS* [Set of Instruments Developmental Language Disorder]. Arnhem, the Netherlands: Cito.

Verhoeven, L. & Van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from Dutch. *Scientific Studies of Reading*, *13*, 205-223. doi:10.1080/10888430902851356

Verhoeven, L., & Van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing: An Interdisciplinary Journal*, *25*, 1805-1818. doi:10.1007/s11145-011-9346-3

Verhoeven, L., Voeten, M., & Vermeer, A. (2019). Beyond the simple view of early first and second language reading: The impact of lexical quality. *Journal of Neurolinguistics*, *50*, 28-36, doi:10.1016/j.jneuroling.2018.03.002

Chapter 3

The impact of first and second language vocabulary on early second language reading comprehension development

This chapter is based on:
Van den Bosch, L. J., Segers, E., & Verhoeven, L. (submitted). *First and second language vocabulary affect early second language reading comprehension development.* Manuscript submitted for publication.

Abstract

This longitudinal study examined variation in early reading comprehension development for second language (L2) readers compared to first language (L1) readers and investigated the impact of L2 readers' first and second language vocabulary knowlege. The participants were 75 Dutch monolingual children (L1 readers) and 71 Turkish-Dutch bilingual children (L2 readers), aged between six and eight years old at the start of the study. In a longitudinal design, three waves of data were collected across second and third grade. The results showed that, on average, the L2 readers had lower reading comprehension scores than the L1 readers, but this performance gap narrowed over time. To further investigate variation among the L2 readers, four categorical subgroups of L2 readers were identified with varying levels of L1 (Turkish) and L2 (Dutch) vocabulary knowledge by means of cluster analysis. Group membership was related to reading comprehension and showed an interaction with time, indicating that reading comprehension performance of the two L2 subgroups with high L1 vocabulary increased more over time as compared to the L1 readers. The L2 subgroup with high vocabulary in both languages even caught up with their monolingual peers in third grade. Together, these findings demonstrate how individual differences in L1 and L2 vocabulary knowledge explain variation in early L2 reading comprehension development and highlight the importance of considering L2 readers' first language in research as well as in education.

Introduction

Previous research has shown that bilingual children who learn to read in their second language often fall behind their monolingual peers in terms of reading comprehension (Melby-Lervåg & Lervåg, 2014). There is broad empirical evidence for the notion that vocabulary knowledge is highly important for reading comprehension performance (Perfetti, 2007), both for monolinguals and bilinguals (e.g., Droop & Verhoeven, 2003). However, most studies with second language (L2) readers did not include their first language (L1) vocabulary and consequently do not provide a complete picture (Bialystok, Luk, Peets, & Yang, 2010). Although L2 vocabulary knowledge is often found to be the strongest determinant of L2 reading comprehension (Lervåg & Aukrust, 2010; Gottordo & Mueller, 2009), also L1 vocabulary knowledge may be of influence depending on children's L1 proficiency (Cummins, 1979; Nicoladis, 2012). To investigate the impact of both L1 and L2 vocabulary knowledge on L2 reading comprehension, studies including L2 readers with varying levels of L1 and L2 proficiency are needed. Also, despite large variation in reading comprehension among beginning readers and L2 readers in particular (Verhoeven, Perfetti, & Pugh, 2019), there are relatively few longitudinal studies on the *early* stages of reading comprehension. The present study aimed to identify subgroups of L2 readers based on individual differences in L1 and L2 vocabulary knowledge to explain variation in early reading comprehension development in bilingual children (L2 readers) as compared to monolingual children (L1 readers).

Reading Comprehension

The importance of reading comprehension is widely acknowledged and several leading theories exist. The Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) identifies word reading skills and language comprehension as two broad, central components of reading comprehension required for decoding and understanding written text, respectively. If readers do not sufficiently master one or both of these skills, this will result in reading comprehension difficulties. The relative importance of the two components is found to differ depending on the orthographic transparency of the language (Florit & Cain, 2011) and the stage of reading development (Tilstra et al., 2009).

The Lexical Quality Hypothesis (Perfetti, 2007) strongly emphasizes the role of vocabulary knowledge and states that variation in the quality of word representations affects reading comprehension. The lexical quality of a specific word is determined by the reader's knowledge of the word's form, meaning, and

use. Consequently, the reader's lexicon of a given language includes words of widely varying lexical quality, ranging from never encountered to well-known words. The crucial role of the lexicon as a connection between the word identification system and the comprehension system is further emphasized in the Reading Systems Framework (Perfetti & Stafura, 2014), which aims to capture the complexity of reading comprehension in an overarching theorie.

Linguistic Diversity

Due to migration and internationalization, a growing number of children have a linguistic minority background and speak another language at home. At primary school, these children learn to read in a language that is not their native language; a situation that can be regarded as emergent bilingualism. With respect to reading comprehension, previous research has shown that L2 readers often fall behind their monolingual peers as a consequence of lower oral language skills, such as vocabulary knowledge, in the language of instruction, despite accurate decoding skills (Melby-Lervåg & Lervåg, 2014). This L1-L2 performance gap already exists in the early stages of reading acquisition (Verhoeven, 2000).

To promote bilingual children's reading comprehension, it is important to gain insight into reading comprehension development and the underlying factors. However, comparative longitudinal studies with monolingual and bilingual children are relatively scarce, especially in the lower primary grades. Furthermore, mixed results have been reported about whether the performance gap in reading comprehension between L1 and L2 readers increases, decreases, or remains stable across the primary grades (Droop & Verhoeven, 2003; Farnia & Geva, 2013; Lervåg & Aukrust, 2010; Limbird et al., 2014; Verhoeven, 2000). Nevertheless, some relevant insights are provided into the factors that explain reading comprehension growth among monolingual and bilingual children. In the upper primary grades, Farnia and Geva (2013) revealed different reading comprehension growth trajectories for L2 readers of English compared to English monolinguals which were best predicted by syntactic knowledge. Silverman et al. (2015) showed that growth in reading comprehension for Spanish-English bilingual children and monolingual English children in grade 2-5 was related to English (L2) vocabulary knowledge. In the intermediate grades, Droop and Verhoeven (2003) also found a strong relation between L2 vocabulary knowledge and reading comprehension development for Dutch monolingual and bilingual children. Lervåg and Aukrust (2010), who focused on the lower primary grades, also showed that L2 vocabulary was a critical

determinant of the difference in reading comprehension growth between young L1 and L2 readers of Norwegian.

L1 and L2 Vocabulary Knowledge

Taken together, previous studies highlight the importance of vocabulary knowledge for reading comprehension performance and development in monolingual as well as bilingual children. However, most studies investigating the role of vocabulary knowledge in bilingual children's reading comprehension only took into account vocabulary knowledge in the language of instruction (L2), and not their home language (L1). Not taking L1 vocabulary knowledge into account can be problematic. First, from a conceptual point of view, this does not provide a complete picture of children's vocabulary knowledge in terms of lexical quality resulting in an unfair comparison with their monolingual peers (Bialystok et al., 2010). Apart from overlap in their lexicon across languages, bilingual children may also know more/other words or have a deeper understanding of words (higher lexical quality) in one language than the other. Second, from a theoretical point of view, first language proficiency is considered to be important for second language acquisition according to the interdependence hypothesis (Cummins, 1979), which states that competence in L2 is partly based on competence in L1. For cross-language transfer to take place a certain L1 proficiency level is required as explained by the threshold hypothesis (Cummins, 1979).

Previous studies that did investigate the impact of L1 vocabulary knowledge on L2 reading comprehension showed mixed results. Some studies have found a unique contribution of children's vocabulary in their home language (L1) (Proctor, August, Carlo, & Snow, 2006; Raudszus, Segers, & Verhoeven, 2018), whereas other studies showed that L1 vocabulary did not predict later reading comprehension performance or development over and above L2 vocabulary (Gottardo & Mueller, 2009; Lervåg & Aukrust, 2010; Mancilla-Martinez & Lesaux, 2010). An explanation for these mixed results may relate to the fact that previous studies overlooked individual differences among L2 readers by considering them as one group. According to the recent view that bilingualism is not dichotomous but a continuum resulting from variation in bilingual usage and proficiency (Luk & Bialystok, 2013), it is likely that bilingual children differ in their levels of L1 and L2 vocabulary knowledge and therefore the impact on L2 reading comprehension may not be the same for all children. Also, from a methodological point of view, when investigating the effect of L1 vocabulary on L2 reading comprehension after accounting for L2 vocabulary, only a small amount of variance remains left to be explained and

fades away any effect that children's L1 vocabulary may actually contribute. To investigate the impact of children's L1 and L2 vocabulary knowledge in a more nuanced and balanced way, studies are needed that distinguish L2 readers with varying levels of L1 and L2 vocabulary taking a profile-oriented approach rather than a variable-oriented approach.

The Present Study

In the present study, we distinguished subgroups of L2 readers based on individual differences in L1 and L2 vocabulary knowledge to explain variation in early reading comprehension development. This approach allowed us to investigate the L1-L2 reading comprehension performance gap and the impact of L1 and L2 vocabulary knowledge in a more nuanced way by capturing individual variation among L2 readers and compare them to L1 readers. For the present study, we collected longitudinal data for monolingual Dutch children (L1 readers) and Turkish-Dutch bilingual children (L2 readers) across the second and third grade of primary school.

The following research questions were addressed:

1. How does reading comprehension performance develop in the early stages of reading acquisition in L2 readers as compared to L1 readers?
2. To what extent can we explain variation in early reading comprehension development among L2 readers (as compared to L1 readers) on the basis of individual differences in L1 and L2 vocabulary knowledge?

We expected a performance gap in reading comprehension across the second and third grade; the L2 readers would have lower reading comprehension performance than the L1 readers (research question 1). However, the size of the performance gap was expected to differ depending on L2 readers' level of L1 and L2 vocabulary knowledge (research question 2). More specifically, we expected that several subgroups of L2 readers could be distinguished with high vs. low levels of L1 and L2 vocabulary knowledge, that differed in reading comprehension development. If L2 vocabulary is related to L2 reading comprehension development, L2 readers with higher L2 vocabulary would show larger growth in reading comprehension than L2 readers with lower L2 vocabulary. If L1 vocabulary is related to L2 reading comprehension development, L2 readers with higher L1 vocabulary would show larger growth in reading comprehension than L2 readers with lower L1 vocabulary.

Method

Participants

Participants in the current study were recruited from thirteen primary schools located in lower-middle class neighbourhoods in urban areas in the Netherlands. Prior to the start, ethical approval was obtained from our institutions' ethics committee and informed consent was obtained from the parents. Based on language background information provided by the teachers, children with Dutch as their first language (L1) and children with Dutch as their second language (L2) were selected to participate. Regarding the L2 readers, only Turkish-Dutch children were selected to maximize the homogeneity of the L2 group, since Turkish is the largest linguistic minority (CBS, 2016) and the most widely spoken additional language among Dutch primary school children (Extra, 2002). The participating schools were all monolingual schools with a mixed student population, meaning that the L2 readers were immersed in a Dutch language context at school and received formal reading instruction in Dutch only. In the Netherlands, formal reading instruction starts in first grade focusing on sound-letter correspondences and decoding, whereas the focus shifts toward reading comprehension from second grade onwards.

At the start of second grade, 167 children (L1 readers: n = 86, L2 readers: n = 81) met the inclusion criteria and participated (see also Chapter 2). Throughout the longitudinal study, 21 children (12.6%) dropped out due to moving to another school or repeating a class. For the remaining children (87.4%), we had complete data across all three time points. The final sample of the current study consisted of 146 children (mean age = 7.5 years, range = 81-105 months at start of the study) of which 75 L1 readers (53.3% boys) and 71 L2 readers (60.6% boys). Except for one child who was adopted at a very young age, all L1 readers were born in the Netherlands and were monolingual speakers of Dutch. The majority of L2 readers was also born in the Netherlands (91.5%), the other L2 readers were born in Turkey (5.6%) or another country (2.8%). The L2 readers can be referred to as immigrant children growing up in a situation of emerging bilingualism; for the majority (87.3%) either one or both parents were born in Turkey. All L2 readers indicated to have Turkish as their home language, but there was variation in the amount of Turkish exposure at home, which is reflective of the population. For the majority (62.0%), Turkish was their main home language, whereas others spoke Dutch to a larger extent (32.4%) or reported equal use of both languages (5.6%). Most children watched Turkish television at home (88.7%), but did not read in Turkish (77.5%).

As background variables, we compared the two language groups on non-verbal reasoning skills and socioeconomic status (SES) (see Table 3.1). The L1 and L2 readers did not differ in non-verbal reasoning skills assessed by Raven Coloured Progressive Matrices (Raven, 1965). There was a significant difference in SES as determined by their parents' average level of education measured on a four-point scale (1 = primary education, 2 = lower secondary/vocational education, 3 = higher secondary education/intermediate vocational education, 4 = higher education/ university). Despite coming from the same classrooms, the L2 readers had lower SES than the L1 readers, which corresponds to other studies with young L2 learners from Turkish families in the Netherlands (e.g., Droop & Verhoeven, 2003; Scheele, Leseman, & Mayo, 2010).

Design

This study has a longitudinal design with three time points of data collection throughout second and third grade with an interval of approximately six months: T1) start second grade, T2) end second grade, and T3) mid third grade. Time (T1, T2, T3) was a within-subject factor, language group (L1 readers vs. L2 readers/L2 subgroups) was a between-subject factor, and reading comprehension performance (in Dutch) was the outcome variable. Other variables of interest were decoding skills (T1, T2, T3), Dutch vocabulary (T1), Turkish vocabulary (T1), and language exposure at home (T1).

Measures

Decoding skills. We administered a word reading test and a pseudoword reading test (Verhoeven & Keuning, 2017), each consisting of three lists with an increasing number of syllables (e.g., *boek*, *blaas*, *dromer / laas*, *stoef*, *gluifel)*. For each list, children were instructed to read aloud the (pseudo)words as quickly and accurately as possible within one minute. Children's score was the total number of correctly read (pseudo)words across the three lists (Cronbach's α = .95–.97; Verhoeven & Keuning, 2017). At each time point, children's scores on both tests were highly correlated (all *r*'s > .90).

Vocabulary knowledge Dutch (L2). We administered a Dutch receptive vocabulary test from the T-TOS (Verhoeven, Keuning, Horsels, & Van Boxtel. 2013) presented on a laptop. For each of the 64 items, a word (noun/verb/adjective) was presented auditorily along with four pictures on the screen. Children had to select the picture that matched the spoken word. The test discontinued after eight consecutive

mistakes. A child's score was the number of items answered correctly (Cronbach's α = .88, Verhoeven et al., 2013).

Vocabulary knowledge Turkish (L1). The Turkish receptive vocabulary was measured with a similar test procedure as the Dutch receptive vocabulary task (i.e., selecting one out of four pictures that corresponds with the spoken word). However, to avoid a learning effect by using an identical test in another language, a translation of the Dutch version of the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007) was used (Blom, 2019). All 96 items were administered and a child's score was the number of items answered correctly (Cronbach's α = .86, in the current study).

Language exposure at home. At the start of the study, we administered a questionnaire to gain insight into L2 readers' language exposure at home. L2 readers were asked about their main home language and their language of preference (Dutch or Turkish). In addition, they had to indicate to what extent they spoke Dutch or Turkish with their family members (i.e., father, mother, siblings, parents together) on a five-point scale (0 = only Dutch / 1 = mostly Dutch / 2 = equal / 3 = mostly Turkish / 4 = only Turkish). We calculated an aggregated score by dividing children's sumscore for the four items by the maximum score, so that a higher score indicated a larger amount of Turkish exposure at home. This aggregated score ranged from 0 (no Turkish exposure) to 1 (only Turkish exposure) and showed a moderate, positive correlation with Turkish receptive vocabulary (r = .48, p < .001).

Reading comprehension. The CITO reading comprehension test was used to assess children's reading comprehension performance (Cito B.V., 2014). This standardized test is part of the Dutch national student monitoring system in primary education (first to sixth grade), which is administered in class by the teacher twice a year. For the current study, we collected children's scores on the three tests administered across second and third grade (CITO M4; CITO E4; CITO M5). These are paper-and-pencil tests that include a variety of texts and question types, which increase in difficulty across grades. These tests consist of 40 multiple-choice items (A-D) in second grade and 50 multiple choice items (A-D) in third grade. Rather than raw scores, scaled scores were used (based on the number of items correct) that are appropriate to measure growth across time points, since they form a vertical scale across the primary grades (maximum score in third grade = 280).

Procedure

Data were collected by the first author and several trained research assistants at school during school hours. At each time point, children participated in individual testing sessions of 30 min each, including a variety of language and literacy related tasks administered in a fixed order. The reading comprehension tests were group administered in class by the teacher, as part of the Dutch national student monitoring system.

Data Analyses

To answer the first research question, a mixed-design ANOVA was performed with Time (T1, T2, T3) as within-subjects factor, Language background (L1 vs. L2) as between-subjects factor, and reading comprehension performance as outcome variable. In addition, SES was added as a covariate to make sure that between-group differences in terms of language background could not (completely) be attributed to differences in SES. We checked whether assumptions of equality of error variances, equality of covariance matrices, and sphericity were met. If the assumption of sphericity was violated, degrees of freedom were corrected using Huyn-Feldt estimates of sphericity (Field, 2009). Furthermore, we specifically looked at the distribution of reading comprehension scores. At T1 and T2, the scores were normally distributed (i.e., skewness and kurtosis values below 1.00). At T3, however, the kurtosis value was considerably high (4.70) due to an extreme value (i.e., more than 3SD above the mean) of one child who obtained the maximum score (280). From a statistical point of view (i.e., this extreme value largely influenced the data) and a conceptual point of view (i.e., maximum scale scores are not reliable according to the guidelines of the reading comprehension test), we decided to remove this case from the analyses (Field, 2009) resulting in a normal distribution (skewness: 0.24, kurtosis: -0.22). Apart from this one missing value for reading comprehension at T3, complete data were available for all other measures across time points.

To answer the second research question, we first conducted K-means Cluster Analysis using the kmeans() function in R (version 3.4.4; R Core Team, 2018) to identify subgroups among the L2 readers. Compared to median split, which is an often used top-down procedure where individuals are categorized on the variable's arbitrary cut-off values (i.e., variable-oriented approach), cluster analysis can be seen as a data-driven, bottom-up approach where the most similar participants on variables of interest are clustered in subgroups (i.e., pattern-oriented approach) (Garcia, MacDonald, & Archer, 2015). Data-driven methods, such as K-means

cluster analysis, better capture the dynamic nature of within and between group variability, whereas the median split procedure is static in nature. As input variables for the cluster analysis, we used L2 readers' z-scores for L1 (Turkish) and L2 (Dutch) vocabulary knowledge (measured at T1) to make them comparable in scale. Next, we again performed a mixed-design ANOVA, but instead of Language background (L2 readers vs. L1 readers) we now selected Language group as a between-subjects factor to compare the L2 subgroups as determined by the cluster analysis to the L1 readers.

Results

Descriptive Statistics

The descriptive statistics for the L1 readers and L2 readers are displayed in Table 3.1. To examine between-group differences, independent samples t-tests were conducted with the linear step-up procedure (Benjamini & Hochberg, 1995) to control for multiple testing. As expected, the two groups did not differ in terms of decoding skills at any time point, but the L2 readers scored considerably lower on Dutch (L2) vocabulary and reading comprehension at each time point, with medium to large effect sizes (Cohen, 1988).

The correlations for the total sample are shown in Table 3.2. Reading comprehension performance across time points showed a weak, positive correlation with decoding skills and a strong, positive correlation with Dutch vocabulary. For the L2 readers, Turkish (L1) vocabulary was not correlated with Dutch (L2) vocabulary, nor any of the other measures.

Table 3.1 Descriptive Statistics and Mean Differences for L1 readers and L2 readers at T1 (start G2), T2 (end G2), and T3 (mid G3)

	L1 readers (*n* = 75)		L2 readers (*n* = 71)		Mean differences		
Measures	*M (SD)*	*range*	*M (SD)*	*range*	*t*	*p*	*d*
DEC WD							
T1	139 (55)	29-309	128 (48)	40-245	1.37	.171	0.21
T2	175 (57)	55-323	159 (48)	70-262	1.71	.089	0.30
T3	212 (55)	95-387	206 (48)	92-285	0.75	.456	0.12
DEC PS							
T1	102 (43)	23-266	100 (38)	22-199	0.34	.735	0.05
T2	121 (44)	38-264	125 (42)	44-209	-0.68	.496	0.09
T3	143 (47)	40-276	151 (41)	59-246	-1.06	.290	0.18
VOC NL							
T1*	43.9 (5.6)	27-56	33.8 (6.7)	15-48	9.91	< .001	1.64
VOC TU							
T1			59.0 (10.2)	34-77			
RC							
T1*	137.9 (26.8)	84-210	115.3 (21.5)	70-190	5.68	< .001	0.93
T2*	141.9 (23.0)	97-201	121.2 (20.6)	78-176	5.71	< .001	0.95
T3*	149.0 (23.2)	99-220	137.2 (20.8)	86-196	3.22	.002	0.54
Background variables							
NVR	28.1 (4.1)	16-35	27.6 (4.3)	18-35	0.64	.525	0.12
SES*	2.8 (0.7)	1.5-4.0	2.3 (0.9)	1.0-4.0	3.41	.001	0.62

Note. DEC WD = decoding words, DEC PS = decoding pseudo words, VOC NL = Dutch vocabulary, VOC TU = Turkish vocabulary, RC = reading comprehension, NVR = non-verbal reasoning, SES = socioeconomic status.
* mean difference is significant based on critical p-values from linear step-up procedure (Benjamini & Hochberg, 1995) to correct for multiple testing.

Table 3.2 Pearson Correlations across Variables for Total Sample (*N* = 146)

	DEC WD	DEC PS	VOC NL	VOC TU	RC T1	RC T2	RC T3
DEC WD	–						
DEC PS	.94***	–					
VOC NL	.07	-.00	–				
VOC TU	.14	.16	.04	–			
RC T1	.20*	.17*	.58***	.07	–		
RC T2	.30***	.25**	.59***	.06	.69***	–	
RC T3	.24**	.23**	.46***	.18	.71***	.73***	–

Note. DEC WD = decoding words, DEC PS = decoding pseudo words, VOC NL = Dutch vocabulary, VOC TU = Turkish vocabulary, RC = reading comprehension.
*$p < .05$, **$p < .01$, ***$p < .001$

Early Reading Comprehension Development: L2 readers vs. L1 readers

The results of the mixed-design ANOVA showed that there was a significant main effect of Language background ($F(1, 142) = 21.02$, $p < .001$, $\eta_p^2 = .129$) and a significant main effect of Time ($F(1.97, 279.14) = 5.58$, $p = .005$, $\eta_p^2 = .038$) on reading comprehension. In addition, there was a significant Time x Language background interaction effect ($F(1.97, 279.14) = 6.52$, $p = .002$, $\eta_p^2 = .044$). Together, the main effects of language background and time indicated that, in general, the L2 readers scored significantly lower than the L1 readers and that children's reading comprehension improved over time, but the interaction effect indicated that the degree of growth in reading comprehension differed by language background. Within-subjects contrast testing showed that the interaction with Language background existed for T1 vs. T3 ($F(1, 142) = 11.97$, $p = .001$, $\eta_p^2 = .078$), and for T2 vs. T3 ($F(1, 142) = 8.60$, $p = .004$, $\eta_p^2 = .057$), but not for T1 vs. T2 ($F(1, 142) = 0.31$, $p = .470$, $\eta_p^2 = .004$). Although the L2 readers scored lower at all time points (see Table 3.1), the growth in reading comprehension from start second grade (T1) and end second grade (T2) to mid third grade (T3) was larger for the L2 readers than the L1 readers (see Figure 3.1).

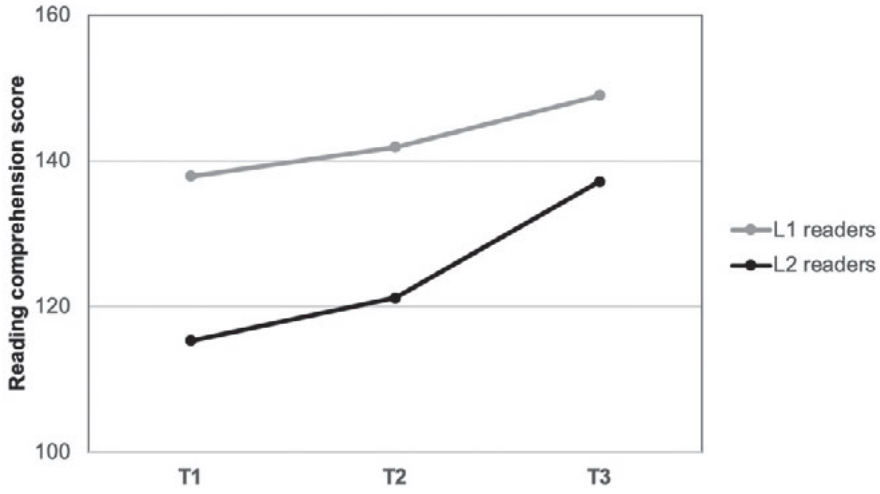

Figure 3.1 Growth in reading comprehension from T1 (start G2) to T2 (end G2) to T3 (mid G3) split by language background (L1 readers vs. L2 readers).

Capturing Individual Differences: Identifying L2 Subgroups

As a next step, we performed K-means cluster analysis to identify L2 subgroups based on individual differences in L1 (Turkish) and L2 (Dutch) vocabulary knowledge. As shown in Table 3.2, there was no significant correlation between

the two vocabulary measures (r = .043, p = .721). This is also reflected in the scatterplot (see Figure 3.2). Based on the screeplot and the amount of explained variance, we found support for a four-cluster solution which explained 70.7% of the variance in Dutch vocabulary (L2) and 65.1% of the variance in Turkish vocabulary (L1). Based on the scatterplot (see Figure 3.2) and the mean z-scores for Dutch and Turkish vocabulary (see Table 3.3), we interpreted the four-cluster solution as follows: 1) one subgroup scored low on both vocabulary measures (LowL2-LowL1, n = 5, 7.0%), 2) one subgroup scored low on L2 vocabulary (Dutch) and high on L1 vocabulary (Turkish) (LowL2-HighL1, n = 21, 29.6%), 3) one subgroup scored high on L2 vocabulary (Dutch) and low on L1 vocabulary (Turkish) (HighL2-LowL1, n = 21, 29.6%), and 4) one subgroup scored high on both vocabulary measures (HighL2-HighL1, n = 24, 33.8%).

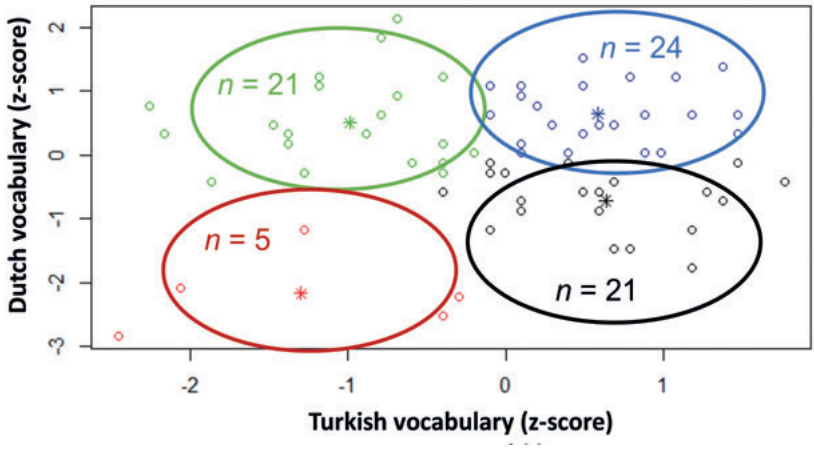

Figure 3.2 Scatterplot for the four-cluster solution: 1) LowL2-LowL1 (red), 2) LowL2-HighL1 (black), 3) HighL2-LowL1 (green), 4) HighL2-HighL1 (blue).

Regarding these subgroups, two remarks should be made. First, schools (n = 13) were equally distributed across subgroups; the five LowL2-LowL1 children came from four different schools, the 21 LowL2-HighL1 children came from nine different schools, the 21 HighL2-LowL1 children came from eight different schools, and the 24 HighL2-HighL1 children came from eleven different schools. Second, the LowL2-LowL1 subgroup included only five children. These were extreme cases in terms of vocabulary; they scored more than 2SDs below the mean on L2 vocabulary (n = 2) or they scored more than 1SD (n = 1) or even 2SDs (n = 2) below the mean on both L2 and L1 vocabulary (see Figure 3.2). Furthermore, their SES was considerably lower compared to the other three L2 subgroups and they had the lowest mean

score for non-verbal reasoning (see Table 3.3). Given the small sample size and their deviant profile, these five children were excluded from the analyses, but for descriptive purposes they are still included in Table 3.3, Table 3.4, Table 3.5 and Figure 3.3.

Table 3.3 Descriptive Statistics for L2 subgroups at T1 (start G2), T2 (end G2), and T3 (mid G3)

	LowL2-LowL1 (*n* = 5)	LowL2-HighL1 (*n* = 21)	HighL2-LowL1 (*n* = 21)	HighL2-HighL1 (*n* = 24)
Measures	*M (SD*	*M (SD)*	*M (SD)*	*M (SD)*
DEC WD				
T1	131 (13)	136 (44)	120 (47)	127 (58)
T2	192 (21)	164 (45)	156 (44)	152 (57)
T3	229 (18)	216 (42)	208 (42)	190 (58)
DEC PS				
T1	106 (12)	105 (33)	92 (35)	101 (47)
T2	155 (20)	136 (39)	115 (36)	119 (48)
T3	174 (20)	160 (38)	149 (39)	141 (47)
VOC NL***				
T1 (raw scores)	19.4 (4.2)	29.1 (3.1)	37.2 (4.6)	38.1 (3.1)
T1 (z-scores)	-2.4 (0.6)	-0.7 (0.5)	0.5 (0.7)	0.6 (0.5)
VOC TU				
T1 (raw scores)	45.8 (9.9)	65.5 (6.4)	49.0 (6.1)	65.0 (5.0)
T1 (z-scores)	-1.3 (1.0)	0.6 (0.6)	-1.0 (0.6)	0.6 (5.0)
RC**				
T1	94.0 (23.5)	104.5 (19.1)	120.5 (15.6)	124.6 (21.4)
T2	101.0 (14.9)	113.6 (18.2)	124.9 (22.5)	128.8 (17.5)
T3	117 (9.5)	130.3 (19.2)	138.6 (17.4)	146.1 (22.4)
Background variables				
EXP TU***	0.6 (0.2)	0.7 (0.2)	0.5 (0.2)	0.6 (0.2)
NVR	25.2 (4.6)	27.0 (4.9)	27.1 (4.3)	29.1 (3.5)
SES	1.2 (0.3)	2.2 (0.7)	2.6 (0.7)	2.4 (1.0)

Note. DEC WD = decoding words, DEC PS = decoding pseudo words, VOC NL = Dutch vocabulary, VOC TU = Turkish vocabulary, RC = reading comprehension, EXP TU = Turkish exposure at home, NVR = non-verbal reasoning, SES = socioeconomic status.
** $p < .01$, *** $p < .001$

Table 3.3 shows the descriptive statistics for the L2 subgroups. The three remaining L2 subgroups did not significantly differ in terms of decoding skills (words: $F(2, 63) = 0.60$, $p = .550$, $\eta_p^2 = .019$, pseudo words: $F(2, 63) = 0.85$, $p = .434$, $\eta_p^2 = .026$), SES ($F(2, 63) = 1.15$, $p = .324$, $\eta_p^2 = .035$) and non-verbal reasoning ($F(2, 63) = 1.94$, $p = .152$, $\eta_p^2 = .058$). By contrast, significant differences between the three L2 subgroups were found for reading comprehension ($F(2, 63) = 5.84$, $p = .005$,

η_p^2 = .156), vocabulary knowledge (L1: $F(2, 63) = 55.44$, $p < .001$, η_p^2 = .638, L2: $F(2, 63) = 40.32$, $p < .001$, η_p^2 = .561) and language exposure ($F(2, 63) = 10.74$,

$p < .001$, η_p^2 = .254). Regarding reading comprehension, children in the LowL2-HighL1 subgroup performed significantly lower than the children in the HighL2-HighL1 subgroup ($p = .004$). Regarding vocabulary knowledge, the subgroups differed as expected based on the cluster analysis (L2 vocabulary: LowL2-HighL1 < HighL2-LowL1 & HighL2-HighL1, L1 vocabulary: HighL2-LowL1 < LowL2-HighL1 & HighL2-HighL1). Regarding language exposure at home, children in the LowL2-HighL1 subgroup reported higher Turkish (L1) exposure at home compared to the children in the HighL2-LowL1 subgroup ($p < .001$). This particular difference was also reflected in their main home language and language preference (see Table 3.4); most children in the LowL2-HighL1 subgroup selected Turkish (L1) as their main home language and language of preference, whereas most children in the HighL2-LowL1 subgroup selected Dutch (L2) as their main home language and language of preference.

Table 3.4 Frequencies (Percentages) for Main Home Language and Language Preference for the L2 Subgroups

	LowL2-LowL1 (*n* = 5)	LowL2-HighL1 (*n* = 21)	HighL2-LowL1 (*n* = 21)	HighL2-HighL1 (*n* = 24)
Main home language:				
- Dutch	2 (40.0%)	2 (9.5%)	12 (57.1%)	7 (29.2%)
- Turkish	3 (60.0%)	18 (85.7%)	8 (38.1%)	15 (62.5%)
- Same extent	0 (0.0%)	1 (4.8%)	1 (4.8%	2 (8.3%)
Language preference:				
- Dutch	1 (20.0%)	8 (38.1%)	11 (52.4%)	11 (45.8%)
- Turkish	2 (40.0%)	11 (52.4%)	8 (38.1%)	12 (50.0%)
- No preference	2 (40.0%)	2 (9.5%)	2 (9.5%)	1 (4.2%)

Early Reading Comprehension Development: L2 subgroups vs. L1 readers

The results of the mixed-design ANOVA showed that there was a significant main effect of Language group ($F(3, 135) = 9.03$, $p < .001$, η_p^2 = .171) and a significant main effect of Time ($F(1.99, 268.99) = 6.15$, $p = .002$, η_p^2 = .044) on reading comprehension. In addition, there was a significant Time x Language group interaction effect ($F(5.98, 268.99) = 2.45$, $p = .025$, η_p^2 = .052) indicating that the degree of growth in reading comprehension differed across language groups. Within-subjects contrast showed that the interaction existed for T1 vs. T3 ($F(3, 135) = 4.57$, $p = .004$, η_p^2 = .092), and for T2 vs. T3 ($F(3, 135) = 2.91$, $p = .037$, η_p^2 = .061), but not for T1 vs. T2 ($F(3, 135) = 0.44$, $p = .724$, η_p^2 = .010). When comparing the three L2 subgroups to the

L1 readers one-by-one, it is shown that the interaction effect only existed for the LowL2-HighL1 subgroup ($F(1.91, 175.72) = 4.94$, $p = .009$, $\eta_p^2 = .051$) and the HighL2-HighL1 subgroup ($F(1.96, 186.02) = 3.32$, $p = .039$, $\eta_p^2 = .034$), but not for the HighL2-LowL1 subgroup ($F(2, 184) = 1.24$, $p = .292$, $\eta_p^2 = .013$). Together, this indicates that the two L2 subgroups with high L1 vocabulary showed larger growth in reading comprehension from start second grade (T1) and end second grade (T2) to mid third grade (T3) compared to the L1 readers (see Figure 3.3).

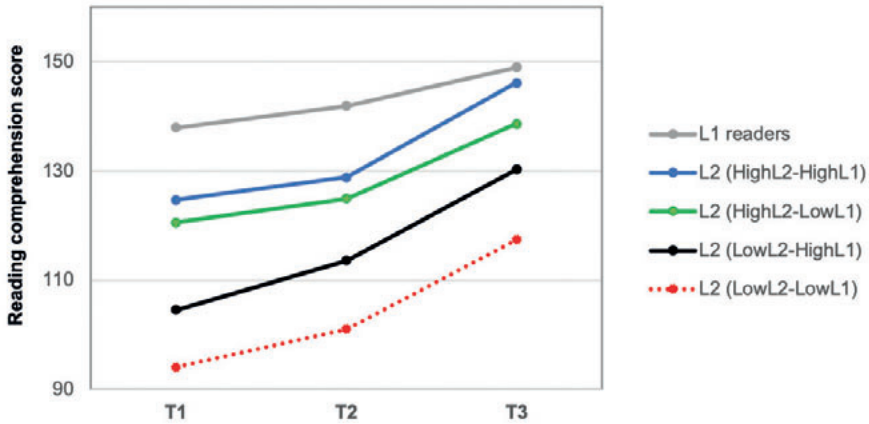

Figure 3.3 Growth in reading comprehension from T1 (start G2) to T2 (end G2) to T3 (mid G3) split by language group (L1 readers vs. L2 subgroups).

As an effect size partial eta squared (η_p^2) was reported above, which may have been influenced by the unequal sample sizes of the groups (L1 readers: $n = 74$ vs. L2 readers: $n = 21$ / $n = 21$ / $n = 24$). Therefore, we also calculated Cohen's *d* effect sizes to further examine and quantify group differences between the L2 subgroups and the L1 group for each time point (see Table 3.5). For the three L2 subgroups included in the analyses, the effect sizes become smaller over time, suggesting that the performance gap (relative to the L1 readers) decreased over time. In accordance with the results reported earlier, the two L2 subgroups with high L1 vocabulary showed the largest decrease in effect sizes over time, particularly from T2 to T3. The two subgroups with high L2 vocabulary showed the smallest effect sizes in general. At T3, a large effect size was observed for LowL2-HighL1 children, a medium effect size for HighL1-LowL2, but only a very small effect size for HighL1-High L2 children (Cohen, 1988), suggesting that the L2 readers with higher L1 *and* L2 vocabulary knowledge caught up with the L1 readers halfway third grade.

Table 3.5 Quantifying Mean Differences (Cohen's d) Between L2 Subgroups and L1 Readers for Reading Comprehension at T1 (start G2), T2 (end G2), and T3 (mid G3)

	LowL1-LowL2 vs. L1 readers	LowL1-HighL2 vs. L1 readers	HighL1-LowL2 vs. L1 readers	HighL1-HighL2 vs. L1 readers
RC T1	1.76	1.45	0.79	0.55
RC T2	2.11	1.36	0.74	0.64
RC T3	1.78	0.88	0.51	0.13

Discussion

In the present study, we aimed to gain more insight into variation in early reading comprehension development of L2 readers as compared to L1 readers. Based on individual differences in L1 and L2 vocabulary knowledge, we identified subgroups of L2 readers to explain variation in early second language reading comprehension among L2 readers.

Early Reading Comprehension Development in L1 and L2 readers

Regarding the first research question, we compared L2 readers' reading comprehension performance across second and third grade to their monolingual peers. First, the results showed that the L2 readers scored considerably lower than the L1 readers on reading comprehension over time, whereas they were even efficient in terms of decoding skills. This finding is in line with the meta-analysis of Melby-Lervåg and Lervåg (2014) and corresponds with the L1-L2 performance gap found in previous studies conducted in the lower (Lervåg & Aukrust, 2010; Verhoeven, 2000), intermediate (Droop & Verhoeven, 2003) and higher primary grades (Farnia & Geva, 2013).

Second, the results showed that both L1 readers' and L2 readers' performance improved over time, with a larger growth evidenced for the L2 readers. This finding indicates that the L1-L2 performance gap in reading comprehension decreased across the lower primary grades, which contrasts with previous studies reporting an increase or no change (Droop & Verhoeven, 2003; Farnia & Geva, 2013; Lervåg & Aukrust, 2010; Limbird et al., 2014). This discrepancy could be explained by the fact that in the lower primary grades most texts require a very basic level of vocabulary and therefore L2 readers may be able to catch up as found in the present study. However, as texts are becoming more complex in the higher primary grades, L2 readers may fall behind their monolingual peers on average, as found in the study of Farnia and Geva (2013) across fourth to sixth grade.

Variation in L1 and L2 Vocabulary Knowledge Among L2 readers

Regarding the second research question on variation among the L2 readers, we first investigated whether subgroups of L2 readers could be identified based on individual differences in L1 (Turkish) and L2 (Dutch) vocabulary. As found in other studies with young L2 readers (Gottardo & Mueller, 2009), these vocabulary scores were not correlated and varied considerably among the L2 readers. We used these two dimensions to group L2 readers based on patterns discovered in the data by means of cluster analysis. The results showed that four subgroups of children with varying levels of L1 and L2 vocabulary knowledge could be distinguished: 1) children with low vocabulary in both their languages, 2) children with low L2 vocabulary and high L1 vocabulary, 3) children with high L1 vocabulary and low L2 vocabulary, and 4) children with high vocabulary in both their languages.

This categorization fits with the recent view that bilingualism should not be seen as a dichotomous variable and shows that there is variation in L1 and L2 vocabulary knowledge among L2 readers, that is likely to be associated with the quantity and quality of language exposure at home (Luk & Bialystok, 2013). That is to say, the L2 readers with higher knowledge in one of their two languages also indicated to have more exposure in that particular language at home and to prefer that particular language. Furthermore, this categorization relates to the different types of bilingualism as distinguished by Cummins (1979) in light of the threshold hypothesis: semilingualism (low levels in both languages), dominant bilingualism (native like level in one of the languages) and additive bilingualism (high levels in both languages).

Variation in Early Reading Comprehension Development

Next, we compared these L2 subgroups to the L1 readers to gain more detailed insight into the performance gap. The findings revealed differences in reading comprehension development among the subgroups that could not be attributed to differences in decoding skills, SES or non-verbal reasoning, but were most likely related to vocabulary knowledge in the language of instruction (L2) and vocabulary in the home language (L1). Regarding L2 vocabulary (Dutch), the results showed that there was a strong, positive correlation with L2 reading comprehension across time, which is in accordance with the meta-analysis of Jeon and Yamashita (2014) showing that L2 vocabulary knowledge is one of the strongest correlates of L2 reading comprehension. Also, the overall performance gap in reading comprehension was smaller (in terms of Cohen's *d* effect sizes) for the L2 subgroups with higher L2 vocabulary than for L2 subgroups with lower L2

vocabulary. Regarding L1 vocabulary (Turkish), the results showed that the two L2 subgroups with higher L1 vocabulary showed an increase in growth compared to the L1 readers; the performance gap decreased. This finding suggests that higher levels of L1 vocabulary knowledge may be beneficial to L2 reading comprehension development, which is in line with Cummins interdependency hypothesis (1979) and more recent meta-analyses showing that apart from within-language associations, also positive cross-language associations between L1 oral proficiency and L2 early literacy and reading exist (Prevoo et al., 2016).

However, some caution is required when interpreting the latter finding as cross-language transfer, since vocabulary in the home language (L1) is likely associated with a wide variety of skills and home language activities/resources that may also (partly) explain the findings of the present study. For instance, the L2 readers with higher L1 vocabulary may have had better meta-linguistic awareness (Bialystok & Barac, 2012) which promotes their reading comprehension (Zipke, Ehri, & Smith Cairns, 2009). Also, the L2 readers with higher L1 vocabulary may have been read to more often in Turkish by their parents (Scheele et al., 2010) resulting in more advanced knowledge of concepts described in narrative texts and a better understanding of text structure which stimulates their reading comprehension (Cain, Oakhill, & Bryant, 2004). Despite these confounding factors, it can be concluded that knowing more words in the home language, which is probably associated with an enriched home language environment, may be valuable for L2 readers' literacy outcomes at school. This is further supported by the finding that particularly the L2 subgroup with high L1 and L2 vocabulary (rather than the L2 subgroup with high L2 vocabulary and low L1 vocabulary) caught up with their monolingual peers in third grade. Apparently, having higher vocabulary knowledge in *both* languages is associated with higher lexical quality in general, which is important for beginning reading achievement (Verhoeven, Voeten, & Vermeer, 2019).

Limitations and Suggestions for Future Research

The present study has some limitations that ask for careful interpretation of the results and for future research to expand on. Children's L1 and L2 vocabulary knowledge were each measured by one single receptive measure indicative of their vocabulary size. However, in terms of lexical quality as a construct that is not only determined by the number of word representations, but also the quality of word representations (Perfetti, 2007), future studies should incorporate a broader range of L1 and L2 measures to assess breadth as well as depth of vocabulary

knowledge (Vermeer, 2001) and to reveal unique and shared domains of L1 and L2 vocabulary (Wolter, 2006). Furthermore, to explain individual differences in vocabulary knowledge and to better understand the circumstances under which L2 readers can catch up with their monolingual peers at school (Paradis & Jia, 2016) other related factors should be investigated in more detail, such as children's home language environment in terms of quantity and quality of language exposure (Unsworth, 2016) and reading and oral language activities (Scheele et al., 2010).

Conclusions and Implications

The present study can be seen as an important step forward in explaining variation in early reading comprehension development of L1 readers and L2 readers. Among L2 readers, subgroups were identified with varying levels of L1 and L2 vocabulary knowledge and different patterns of growth in reading comprehension, indicating that both L1 and L2 vocabulary affect L2 reading comprehension. Building on our previous study presented in Chapter 2 (Van den Bosch et al., 2019), the current findings also demonstrate that categorizing children as either monolingual or bilingual does not do justice to individual variation among children (Luk & Bialystok, 2013) and stresses the importance of considering individual differences among bilingual children in research as well as in education. With respect to educational practice, the present study demonstrates that vocabulary knowledge of bilingual children in both their first and second language can be considered the key factor to close the performance gap in reading comprehension between L1 readers and L2 readers.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling for the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistics Society*, *57*, 289-300.

Bialystok, E., & Barac, R. Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*, 67-73. doi:10.1016/j.cognition.2011.08.003

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*, 525-531. doi:10.1017/S1366728909990423

Blom, E. (2019). Domain-general cognitive ability predicts bilingual children's receptive vocabulary in the majority language. *Language Learning*, *69*, 292-322. doi:10.1111/lang.12333

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31–42. doi:0.1037/0022-0663.96.1.31

CBS Statistics the Netherlands (2016). Bevolking naar migratieachtergrond [Population by immigration background]. Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Cito B.V. (2014). *Begrijpend lezen 3.0 voor groep 4 en 5 [Reading Comprehension 3.0 for second and third grade]*. Arnhem, the Netherlands: Cito B.V.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*, 222-251. doi:10.3102/00346543049002222

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, *38*, 78-103. doi:10.1598/RRQ.38.1.4

Dunn, D. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition, manual*. Minneapolis, MN: NCS Pearson, Inc.

Extra, G. (2002). *De andere talen van Nederland: Thuis en op school. [The other languages of the Netherlands: At home and at school]*. Bussum, the Netherlands: Coutinho.

Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*, 389-4212. doi:10.1111/jrir.12003

Field, A. (2009). *Discovering statistics using SPSS: third edition*. London, UK: Sage.

Florit, E., & Cain, K. (2011). The Simple View of Reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*, 553-576. doi:10.1007/s10648-011-9175-6

Garcia, D., MacDonald, S., & Archer, T. Two different approaches to the affective profiles model: median splits (variable-oriented) and cluster analysis (person-oriented). *PeerJ*, *3*, e1380. doi:10.7717/peerj.1380

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, *10*, 330-344. doi:10.1037/a0014320

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, *7*, 6-10. doi:10.1177/074193258600700104

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160. doi:10.1007/BF00401799

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*, 160-212. doi:10.1111/lang.12034

Lervåg, A. & Aukrust, V.G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*, 612-620. doi:10.1111/j.1469-7610.2009.02185.x

Limbird, C. K., Maluch, J. T., Rjosk, C. R., Stanat, P., & Merkens, H. (2014). Differential growth patterns in emerging reading skills of Turkish-German bilingual and German monolingual primary school students. *Reading and Writing: An Interdisciplinary Journal*, *27*, 945-968. doi:10.1007/s11145-013-9477-99

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*, 605-621. doi:10.1080/20445911.2013.795574

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority children. *Journal of Educational Psychology*, *102*, 701-711. doi:10.1037/a0019135

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409-433. doi:10.1037/a0033890

Nicoladis, E. (2012). Cross-linguistic influence in French-English bilingual children's Possessive constructions. *Bilingualism: Language and Cognition*, *15*, 320-328. doi:10.1017/S1366728911000101

Paradis, J., & Jia, R. (2017). Bilingual children's long-term outcomes in English as a second language: Language environment factors shape individual differences in catching up with monolinguals. *Developmental Science*, *20*, 1-15. doi:10.1111/desc.12433

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

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22-37. doi:10.1080/10888438.2013.827687

Prevoo, M. J. L., Malda, M., Mesman, J., Van IJzendoorn, M. H. (2016). Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research*. *86*, 237-276. doi:10.3102/0034654315584685

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology*, *98*, 159-169. doi:10.1037/0022-0663.98.1.159

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Raudszus, H., Segers, E., & Verhoeven, L. (2018). Lexical quality and executive control predict children's first and second language reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *31*, 405-424. doi:10.1007/s11145-017-9791-8

Raven, J. C. (1965). *Guide to using coloured progressive matrices, Sets A, AB and B*. London, UK: BHK Levis & Co.

Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*, 117-140. doi:10.1017/S0142716409990191

Silverman, R. D., Proctor, C. P., Harring, J. R., Hartranft, A. M., Doyle, B., & Zelinke, S. B. (2015). Language skills and reading comprehension in English monolingual and Spanish-English bilingual children in grades 2-5. *Reading and Writing: An Interdisciplinary Journal*, *28*, 1381-1405. doi:10.1007/s11145-015-9575-y

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, *32*, 383-401. doi:10.1111/j.1467-9817.2009.01401.x

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Lifespan perspectives on bilingualism* (pp. 136-196). Mouton de Gruyter/APA.

Van den Bosch, L. J. Segers, E., & Verhoeven, L. (2019). The role of linguistic diversity in the prediction of early reading comprehension: A quantile regression approach. *Scientific Studies of Reading*, *23*, 203-219. doi:10.1080/10888438.2018.1509864

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*, 217-234.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*, 313-330. doi:10.1207/S1532799XSSR0404_4

Verhoeven, L., & Keuning, J. (2017). The nature of developmental dyslexia in a transparent orthography. *Scientific Studies of Reading*, 1-17. doi:10.1080/10888438.2017.1317780

Verhoeven, L., Keuning, J., Horsels, L., Van Boxtel, H. (2013). *Testinstrumentarium Taalontwikkelingsstoornissen T-TOS* [Set of Instruments Developmental Language Disorder]. Arnhem, the Netherlands: Cito.

Verhoeven, L., Perfetti, C., & Pugh, K. (2019). Cross-linguistic perspectives on second language learning. *Journal of Neurolinguistics*, *50*, 1-6. doi:10.1016/j.jneuroling.2019.02.001

Verhoeven, L., Voeten, M., & Vermeer, A. (2019). Beyond the simple view of early first and second language reading: The impact of lexical quality. *Journal of Neurolinguistics*, *50*, 28-36, doi:10.1016/j.jneuroling.2018.03.002

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics*, *27*, 741-747. doi:10.1093/applin/aml036

Zipke, M., Ehri, L. C., & Smith Cairns, H. (2009). Using semantic ambiguity instruction to improve third graders' metalinguistic awareness and reading comprehension: An experimental study. *Reading Research Quarterly*, *44*, 300-321. doi:10.1598/RRQ.44.3.4

Chapter 4

Online processing of causal relations in beginning readers from diverse linguistic backgrounds

This chapter is based on:
Van den Bosch, L. J., Segers, E., & Verhoeven, L. (2018). Online processing of causal relations in beginning first and second language readers. *Learning and Individual Differences*, *61*, 59-67. doi:10.1016/j.lindif.2017.11.007

Abstract

This experimental study investigated online processing of causal relations in beginning first (L1) and second language (L2) readers (8-10 years old). Children's processing times of two-clause sentences including a causal relation (e.g., 'Tom ate an apple because he was hungry.') were measured by means of eye-tracking. Two text-related factors were investigated: coherence marking (i.e., presence vs. absence of the Dutch connective *omdat* 'because') and linear order of clauses (i.e., cause-effect vs. effect-cause). In addition, syntactic knowledge was included as a child-related factor of interest. The results showed that coherence marking and individual differences in syntactic knowledge influenced children's online sentence processing. In contrast to L1 readers, the absence of a connective led to longer sentence processing times for L2 readers with lower syntactic knowledge; they experienced more difficulty with processing sentences in which no connective was present. Apparently, L2 readers with limited syntactic knowledge benefit from coherence marking provided by a connective, which allows them to establish the causal coherence relation between clauses in a more efficient way. Reversing the linear order of clauses did not affect children's online sentence processing. This study provides an initial step toward the use of online measures to examine sentence processing in beginning L1 and L2 readers aimed at gaining more insight into L2 reading comprehension difficulties.

Introduction

Being able to comprehend written text is a highly important skill. However, children who are confronted with the challenging task of learning to read in their second language, such as children from language minorities for whom the language of instruction at school differs from their home language, often experience reading comprehension difficulties (Droop & Verhoeven, 2003; Lesaux, Lipka, & Siegel, 2006; Mancilla-Martinez & Lesaux, 2010). These children can be referred to as second language readers (L2 readers), since they acquire their second language mainly from their school environment and usually this is the only language in which they learn to read at school. Previous studies have shown that L2 readers' reading comprehension difficulties are probably not a consequence of poor decoding skills. Their decoding skills are found to be in the average range (Geva & Zadeh, 2006; Mancilla-Martinez & Lesaux, 2010) and develop at a more or less equal rate compared to L1 readers (Lesaux & Siegel, 2003; Verhoeven, 2000). By contrast, there is a large discrepancy in the linguistic proficiency of L1 readers and L2 readers, for instance in terms of their vocabulary and syntactic knowledge in the target language (Melby-Lervåg & Lervåg, 2014). Previous research has shown that oral language skills are important for reading comprehension performance and development, even more so for L2 readers (Droop & Verhoeven, 2003; Lervåg & Aukrust, 2010). L2 readers' reading comprehension difficulties can already be observed early in the process of learning to read (Verhoeven, 2000). In these early stages, beginning readers move from reading isolated words to reading sentences and short texts in which they have to integrate word meanings in the context of a sentence or a text, also referred to as word-to-text integration (WTI; Perfetti et al., 2008). For WTI processes to take place, both text and reader characteristics play an important role.

There are textual factors that can help the reader to construct a coherent mental representation of a text. One textual factor in that respect is the use of linguistic markers, such as connectives. Connectives (i.e., conjunctions such as *because*, *therefore*) are cohesive devices that signal the relation between clauses; they are critical to the construction of a coherent text representation (Cain & Nash, 2011). Connectives can be classified according to the type of relationship they signal (i.e., additive, adversative, causal, and temporal), as put forward by Halliday and Hasan (1976) based on their extensive analysis of connective devices. Connectives can guide the reader in how to construct meaning by coding coherence relations, for instance when processing sentences with causal relations (Canestrelli, Mak, &

Sanders, 2013; Sanders & Noordman, 2000). In particular, less-skilled adolescent readers appear to benefit from the presence of (causal) connectives (Land, 2009; Van Silfhout, Evers-Vermeul, Mak, & Sanders, 2014). Furthermore, Cain and Nash (2011) showed that connectives even support online text processing in 8- and 10-year-old less-experienced monolingual readers whose knowledge and comprehension of connectives are still developing.

The question is whether linguistic markers, such as connectives, also support L2 readers with limited proficiency in their second language. On the one hand, based on studies showing that less-skilled readers particularly benefit from linguistic markers in the text, one could argue that this would also hold for L2 readers with poor reading comprehension performance. Connectives may help these readers to identify text structure and to establish a coherent mental model (Degand & Sanders, 2002). On the other hand, one may predict that because of their limited linguistic proficiency in the target language, L2 readers are not competent enough to benefit from the coherence marking provided by connectives. That is to say, connectives are relatively infrequent in texts and may pose extra challenges in terms of both vocabulary knowledge and the type of inter-clausal relationship they signal (Crosson, Lesaux, & Martiniello, 2008). In a study with L2 readers of English, Crosson and Lesaux (2013) showed that bilingual fifth-graders lagged behind their monolingual peers in knowledge of several types of connectives, and that the influence of connectives on reading comprehension varied by readers' linguistic background (i.e., L1 vs. L2). An explanation put forward in line with the findings of Degand and Sanders (2002) is that L2 readers need to be proficient enough in their second language in order to benefit from the coherence marking provided by connectives.

Another relevant textual factor with respect to processing causal relations is word order, or more specifically, linear order. To illustrate, the two following sentences differ in linear order: 1) *Because Tom worked hard, he passed the exam*, 2) *Tom passed the exam, because he worked hard*. The first sentence has an iconic order (i.e., cause-effect), whereas the second sentence has a non-iconic order (i.e., effect-cause). Previous studies have shown contrasting findings with respect to the effect of linear order on online processing. On the one hand, it is assumed that an iconic order facilitates processing: if the order of clauses corresponds to the order of causality in the world, it is easier to understand the sentence (Noordman & De Blijzer, 2000). This is confirmed by a study of Noordman and Vonk (1998), who used several experimental texts containing a causal relation expressed in two

different orders. On the other hand, there are also studies showing that non-iconic sentences are easier to process, rather than iconic sentences (e.g., Graesser, Singer, & Trabasso, 1994; Magliano, Baggett, Johnson, & Graesser, 1993). These studies suggest that readers are more likely to make knowledge-based inferences about causal antecedents (i.e., *Tom passed the exam, so he must have worked hard*), than about effects or consequences (i.e, *Tom worked hard, so he must have passed the exam*).

One explanation put forward for these incongruent findings relates to the differences in predictability of the causal relations that are investigated (Noordman & Vonk, 1998), which can be either a consequence of the specific relations expressed in the sentence (some relations are more predictable than others; Mak & Sanders, 2013) or the cognitive/linguistic capacities of the reader (some readers make better predictions than others; Martin et al., 2013). With respect to the latter, it should be noted that in previous studies university students were involved for whom the predictability of causal relations is relatively high in general. It might well be the case that for beginning readers, and especially beginning L2 readers with limited linguistic proficiency in the target language, a beneficial effect of non-iconic sentences over iconic sentences only prevails when the causal relation is highly predictable in itself. Another explanation lies in the amount of word order flexibility languages permit (Kaiser & Trueswell, 2004). In the case of L2 readers, it has indeed been demonstrated that facilitating effects may occur if the language structure or syntax of their second language is congruous with their first language (Bialystok, Luk, & Kwan, 2005; Nagy, Mcclure, & Mir, 1997).

Apart from the text, characteristics of the reader play a crucial role in online text processing. Children's syntactic knowledge is often mentioned as an important predictor for reading comprehension performance (e.g., Oakhill, Cain, & Bryant, 2003). Moreover, the meta-analysis of Jeon and Yamashita (2014) showed that for L2 reading comprehension, L2 syntactic knowledge was the strongest correlate. That is to say, the reader needs to have accurate syntactic knowledge for understanding the syntactic structure of a sentence. These parsing processes are required for comprehension (Perfetti & Stafura, 2014). Of particular relevance is the case of L2 readers, since their L2 oral language skills, including syntactic knowledge, are shown to stay behind compared to L1 readers (Droop & Verhoeven, 2003; Lesaux et al., 2006), which may in turn influence their online sentence processing.

In order to gain more insight into children's WTI processes at the sentence level, online measures can be recommended that expose reading comprehension while it happens (Perfetti & Stafura, 2014). Previous studies using event-related potentials (ERPs) as an online measure to examine the integration processes in adult readers revealed processing differences between L1 and L2 readers (Kaan, Kirkham, & Wijnen, 2016; Martin et al., 2013). However, there are hardly any studies in which online measures have been used to compare online sentence processing in young L1 and L2 readers. Although eye-tracking is used more often in recent research on children's language and literacy skills (Blythe & Joseph, 2011), its use is limited to older, more experienced readers in upper primary school (e.g., De Leeuw, Segers, & Verhoeven, 2016; Van der Schoot, Vasbinder, Horsley, & Van Lieshout, 2008) and secondary education (Van Silfhout et al., 2014). The few studies on online sentence processing that focused on beginning readers (8-10 years old) were also restricted to monolingual children and focused on age differences rather than variation in linguistic proficiency (Cain & Nash, 2011; Wannacott, Joseph, Adelman, & Nation, 2015). In order to provide more insight into L2 readers' comprehension problems, it is essential to investigate the effect of textual and child-related factors on online sentence processing in beginning L1 and L2 readers.

The Present Study

In the present study, we used an online sentence reading task including two-clause Dutch sentences with a causal relation (e.g., 'Because Tom was hungry, he ate an apple.'). We investigated two textual factors related to syntactic structure (i.e., coherence marking and linear order of clauses) in two groups of beginning readers: Dutch monolingual children with Dutch as their first language (L1 readers) and Turkish-Dutch bilingual children with Dutch as their second language (L2 readers). In addition to children's language background (i.e., L1 vs. L2), we investigated the role of individual differences in syntactic knowledge as a child-related factor of interest for online sentence processing. Eye-tracking was used to measure children's sentence processing times. Taking a closer look at the native Turkish-speaking children as L2 readers of Dutch, two points should be noted. First, they were expected to have lower L2 syntactic knowledge than the L1 readers (cf. Droop & Verhoeven, 2003; Lipka & Siegel, 2007), which may in turn influence the extent to which they make use of connectives during sentence processing and comprehending. Second, their first language (i.e., Turkish) has a more flexible word order (Nilsson, 1991), which may influence the extent to which alternating word order (or linear order of clauses) affects their online text processing and comprehension in L2.

The following research question was central to the present study: To what extent are beginning L1 and L2 readers' online processing times of sentences with causal relations influenced by (a) coherence marking, (b) linear order of clauses, and (c) individual differences in syntactic knowledge? For coherence marking, we expected that beginning readers would benefit from the presence of connectives, and that particularly L2 readers with lower syntactic knowledge would have more difficulty with processing sentences in which no connective is present. For linear order, we expected that non-iconic sentences would be easier to process than iconic sentences, when the predictability of the causal relations is high. In that respect, two contrasting expectations can be formulated for the L2 readers. On the one hand, the facilitating effect of non-iconic sentences may be larger for the L2 readers given their experience with flexibility in word order in their first language. On the other hand, the facilitating effect of non-iconic sentences may be less profound in L2 readers, because their limited linguistic proficiency in their second language, such as lower syntactic knowledge, limits them in making predictions about causal antecedents.

Method

Participants

The participants were third-grade children from eight primary schools in the Netherlands. We included third-graders because they can be categorized as beginning readers (i.e., they have completed only two years of formal reading instruction), nevertheless their word reading skills are sufficiently developed to allow for reliable eye-movement recording during reading. Based on information retrieved from the teachers, monolingual children with Dutch as their home language (L1 readers) and bilingual children with Turkish as their home language (L2 readers) with accurate word decoding skills were selected to participate in this study. The total sample included 46 third-grade children, aged between eight and ten years old (*M* = 9;0, *SD* = 0;6): 27 Dutch monolingual children (age: *M* = 8;11, *SD* = 0;6, 48% boys) and 19 Turkish-Dutch bilingual children (age: *M* = 9;0, *SD* = 0;6, 47% boys). All children had normal or corrected-to-normal vision and none of them were diagnosed with dyslexia. One child was excluded from the study due to calibration problems. Participation was voluntary and written parental consent and children's oral assent were obtained. Prior to the start of the study, ethical approval was obtained from our institution's ethics committee.

All children came from monolingual Dutch schools, meaning that all L2 readers were immersed in a Dutch language context at school and received formal reading instruction in Dutch only. In the Netherlands, Turkish is the largest cultural minority (CBS, 2016). Children from this cultural minority grow up in a situation of emerging bilingualism in families with lower socioeconomic status (SES) in general. A parental questionnaire and interviews with the children showed that the Turkish-Dutch children in our sample were all born in the Netherlands and for most of these children either both their parents (42.9%) or one of their parents (47.8%) was born in Turkey. The Turkish-Dutch children as well as their parents indicated that at least one of the parents spoke primarily Turkish with the child. This is in line with other studies including young L2 learners from Turkish families in the Netherlands (Blom, Küntay, Messer, Verhagen, & Leseman, 2014; De Zeeuw, Schreuder, & Verhoeven, 2013; Scheele, Leseman, & Mayo, 2010; Janssen, Segers, McQueen, & Verhoeven, 2015). Although Turkish is their first language, by the end of the second grade Dutch usually becomes their dominant language (Verhoeven, 2000). In our sample this is confirmed by the fact that the majority of the L2 learners (62.3%) selected Dutch as their preferred language over Turkish. Moreover, all children indicated that they read in Dutch at home; only four of them were literate in Turkish.

We calculated children's SES by averaging their parents' level of education (see Table 4.1), measured on a four-point scale ranging from 1 (primary school) to 4 (higher education or university). As expected based on studies with a comparable sample (e.g., Blom et al., 2014; De Zeeuw et al., 2013, Janssen et al., 2014), the L2 readers had lower SES than the L1 readers on average. In addition, we compared the L1 readers and L2 readers on several other background variables that were considered relevant to the current study (see Table 4.1). To measure children's word decoding skills, we administered the One Minute Test, form A (Brus & Voeten, 1979), which is a measure of decoding efficiency. As expected, no differences were found in decoding efficiency between the L1 readers and L2 readers. To measure children's verbal working memory and short-term memory skills, the Dutch version of the forward and backward digit span from the Wechsler Intelligence Scale for Children (WISC-III; Kort et al., 2005) was used. The L1 readers and L2 readers did not differ with respect to verbal short-term memory and verbal working memory. Furthermore, we collected children's scores on two standardized tests for vocabulary knowledge and reading comprehension, which are paper-and-pencil tasks including seventy and fifty multiple-choice question, respectively. These tests are administered yearly in class by the teacher as part of the Dutch national student monitoring system in primary schools developed by the Dutch National Institute for

Educational Measurement (CITO). As expected, the L2 readers scored significantly lower than the L1 readers on vocabulary knowledge and reading comprehension (see Table 4.1).

Table 4.1 Descriptive Statistics and Mean Differences for L1 readers and L2 readers

	L1 readers (*n* = 27)		L2 readers (*n* = 19)		Mean differences		
Background variables	*M (SD)*	*range*	*M (SD)*	*range*	*t*	*p*	*d*
SES	3.0 (0.7)	1.5-4	2.2 (0.9)	1-4	3.55	.001	0.99
Decoding skills	58.5 (12.3)	34-87	57.6 (7.4)	45-73	0.27	.790	0.09
Verbal short-term memory	7.3 (1.4)	4-10	6.6 (1.1)	5-9	0.79	.094	0.56
Verbal working memory	4.1 (1.2)	2-7	4.5 (1.1)	3-6	-1.18	.247	0.35
Vocabulary knowledge	48.3 (8.8)	32-63	33.7 (11.1)	17-56	5.01	.000	1.46
Reading comprehension[a]	23.9 (11.4)	2-48	15.1 (13.6)	-9-48	2.38	.022	0.70

Note. [a]Scores for the CITO Reading comprehension test can be negative, since this is a norm-based (standardized) score that is used throughout primary school (first grade: *M* = -2.4, *SD* = 15.6); fourth grade: *M* = 33.1, *SD* = 13.2). The mean score for the monolingual children in the present study was comparable to the standardized score in third grade (*M* = 22.3, *SD* = 13.9).

Materials

Online reading task. To investigate sentence processing, we created an online reading task that was based on a design used by Pearson (1974). The task included 16 items, which were two-clause Dutch sentences written in past tense, consisting of seven to eleven words. In order to minimize the influence of vocabulary knowledge, the items only included words that were assumed to be already familiar to the third-grade children based on a Dutch wordlist for six-year olds (Schaerleakens, Kohnstamm, & Lejaegere, 1999). Furthermore, all items included a causal content relation (i.e., the clauses are related because of their propositional content reflecting a real-world causality between two events/states in the world) rather than an epistemic causal relation (i.e., the clauses are related by the speaker's reasoning) (Noordman & De Blijzer, 2000). As a first factor of interest, we investigated coherence marking; half of the items included the Dutch connective *omdat* 'because' to explicitly signal the causal relation between the two clauses, whereas in the remaining items no connective was present and therefore the relation between clauses was more implicit. As a second factor of interest, we looked at the linear order of clauses; half of the sentences had a (iconic) cause-effect order and in the remaining sentences the order was reversed to a (non-iconic) effect-cause order.

Crossing these two textual factors (2 x 2 design) resulted in four within-subjects conditions (see Table 4.2). Each participant was presented with four items in each condition, so 16 items in total (see Appendix A for an overview of all items). The items were presented in pseudorandomized order. The children were instructed to read aloud the sentences that were successively presented on the eye-tracking screen. Comprehension questions were used to check whether children understood the causal relation between the two clauses (*'Why did Tom eat an apple?'*). According to our expectations based on the high predictability of the causal relations, children's performance on these comprehension questions was at ceiling; hardly any response was incorrect. Therefore, these off-line comprehension data were not subjected to statistical analyses.

Table 4.2 Overview of the Four Within-Subject Conditions in the Online Reading Task

Condition	Item	Connective	Linear order
1	Tom at een appel, omdat hij honger had. [Tom ate an apple, because he was hungry.]	present	effect-cause
2	Omdat Tom honger had, at hij een appel. [Because Tom was hungry, he ate an apple.]	present	cause-effect
3	Tom at een appel. Hij had honger. [Tom ate an apple. He was hungry.]	absent	effect-cause
4	Tom had honger. Hij at een appel. [Tom was hungry. He ate an apple.]	absent	cause effect

Syntactic knowledge. A grammaticality judgment task was used to assess children's syntactic knowledge (Verhoeven, Keuning, Horsels, & van Boxtel, 2013). This digital task consists of 55 sentences that were presented auditorily. After listening to each sentence, the children had to indicate whether the sentence was grammatically correct or incorrect by clicking on a green or red button respectively. In total, 17 sentences were grammatically correct and the remaining 38 sentences were grammatically incorrect because of incorrect inflection of nouns (10 sentences), incorrect verb conjugation (10 sentences), incorrect order of words (10 sentences), or incorrect subject-verb agreement (8 sentences). The total score is determined by the number of items answered correctly with a maximum score of 55 (Cronbach's $\alpha = .79$ in the present study). As expected, the L2 readers ($M = 38.2$, $SD = 6.0$) scored significantly lower on this task than the L1 readers ($M = 41.9$, $SD = 5.3$) ($t(44) = 2.22$, $p = .032$, $d = 0.65$).

Apparatus

During the online reading task, children's eye movements were recorded with a Tobii PRO TX300 eye-tracking system (Tobii AB, Danderyd, Sweden). This is a portable, screen-based eye-tracker that captures gaze data at the sampling rate of 300 Hz by using a binocular tracking technique. Children were seated in front of the eye-tracker with a distance of approximately 65 cm between their eyes and the monitor (23 inch with a resolution of 1920 x 1080). Single-line sentences were successively presented in the monospaced font Courier New (bold, 16 point) in black against a light grey background. A nine-point calibration procedure was used to locate the position of children's eyes and the quality was assessed by visual inspection. In case of poor quality in terms of accuracy and/or precision, recalibration was conducted. During the task, the microphone was used to record children's reading and answering.

As an eye-tracking measure, we calculated the total fixation duration in milliseconds (i.e., total time spent reading; Juhasz & Pollatsek, 2011) for each sentence, based on children's eye movements. Total fixation duration is a measure of later processing rather than early processing (e.g., first fixation duration), therefore total fixation duration is assumed to be a more relevant measure for investigating WTI processes in the current study, since we aimed to capture all integration processes at the sentence level. Apart from this aim, also the design of the study makes it more relevant to focus on online processing at the sentence level rather than the clause level, given that clauses were not directly comparable in terms of content, position and length as a consequence of necessary manipulations of both coherence marking and linear order. In comparison to a self-paced reading paradigm, eye-tracking technology is preferable in terms of ecological validity, because it allows for natural reading patterns and does not interfere with the reading process.

Procedure

Children were individually tested during two sessions of approximately twenty minutes in a quiet room at their school. The two sessions were at the same day or two consecutive days. In the first session, the children completed the online reading task while their eye movements were recorded. The children were informed that the study was aimed to investigate reading comprehension in the third grade. They were instructed to read aloud at their normal rate, while comprehending what they were reading, because each sentence would be followed by a comprehension question. The task was preceded by a short practice trial and a calibration procedure to locate the position of the eyes. As soon as the child completed reading

the sentence, the experimenter pressed a button and then the comprehension question appeared. After answering this question, a neutral screen with the item number in the center of the screen was presented for 2000 milliseconds and then the next sentence appeared. In order to prevent these beginning readers from a cognitive overload and to stay focused on their reading comprehension, we decided to let the children read aloud (which is very common for children at this stage of reading development) and let the experimenter press the button after each item. For the purpose of maximizing ecological validity, we decided not to use a chinrest. Voice recording allowed for checking children's comprehension performance afterwards. In the second session, children's syntactic knowledge, decoding skills and cognitive skills were measured by means of standardized tests.

Data Analysis

We conducted linear mixed effects regression analysis (Baayen, Davidson & Bates, 2008) with crossed random effects for subjects and items using the lme4 package (version 1.1.13; Bates, Maechler, Bolker, & Walker, 2015) for the R computing environment (version 3.4.0 Revised; R Core Team, 2017). The outcome variable was the average processing time per word. That is to say, we looked at the total fixation duration (in milliseconds) for each sentence and calculated the average processing time per word in order to correct for differences in sentence length (due to our manipulation of connective presence). All values two standard deviations above or below item *and* subject mean were excluded from the analysis (cf. Van Silfhout et al., 2014). According to these criteria we excluded 2.3% of the data (i.e., 17 cases: 12 cases for L1 readers and 5 cases for L2 readers).

As fixed effects we included connective (0 = connective present, 1 = connective absent), order (0 = cause-effect, 1 = effect-cause), and language background (0 = L1 reader, 1 = L2 reader) as dummy coded variables, and syntactic knowledge as a continuous variable. Random effects were fit using a maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013). This included random intercepts for subjects and items, by-subject random slopes for connective, order and their interaction, and by-item random slopes for language background, syntactic knowledge and their interaction. Log-transformations were performed for processing time as the outcome variable in order to meet the normality requirements for linear modeling (Baayen, 2008). Models were fit using a restricted maximum likelihood technique. A fixed effect was considered significant if the absolute value of the *t*-statistic was larger than or equal to 2.0 (Gelman & Hill,

2007). In addition, *p*-values were estimated from the *t*-distribution based on a formula put forward by Linck and Cunnings (2015) (see also Baayen, 2008, p. 248).

Results

The results of the linear mixed effects analysis are shown in Table 4.3.

Table 4.3 Results of the Linear Mixed Effects Regression Analysis for Processing Time

	Fixed effects			*Random effects*	
				By Subject	By Item
	β	*SE*	*t*	*SD*	*SD*
(Intercept)	-6.781	0.218	31.17	0.112	0.128
Connective (absent vs. present)	-0.005	0.195	-0.02	0.042	-
Order (effect-cause vs. cause-effect)	-0.276	0.186	-1.49	0.020	-
LB (L2 vs. L1)	-0.333	0.296	-1.13	-	0.060
SK	-0.010	0.005	-1.94	-	0.001
Connective x Order	0.206	0.278	0.74	0.063	-
Connective x LB	0.513	0.252	2.04*	-	-
Order x LB	0.456	0.235	1.94	-	-
Connective x SK	0.002	0.004	0.44	-	-
Order x SK	0.005	0.004	1.38	-	-
LB x SK	0.007	0.007	1.01	-	0.001
Connective x Order x LB	-0.554	0.355	-1.56	-	-
Connective x Order x SK	-0.005	0.006	-0.81	-	-
Connective x LB x SK	-0.013	0.006	-2.03*	-	-
Order x LB x SK	-0.011	0.006	-1.88	-	-
Connective x Order x LB x SK	0.014	0.009	1.56	-	-

Note. Connective = connective presence/absence, Order = linear order of clauses, LB = language background, SK = syntactic knowledge. Three factors were dummy coded as follows: Connective (0 = present, 1 = absent), Order (0 = cause-effect, 1 = effect-cause), and Language background (0 = L1 reader, 1 = L2 reader). Model formula: log(RT) ~ Connective*Order*LB*SK + (1+ Order*Connective | subject) + (1+ LB*SK | item).
* $p < .05$.

The results indicated that there were no significant main effects for connective, order, language background, and syntactic knowledge on children's processing time (all *t*'s < 1.94). However, for the first text-related factor of interest (coherence marking), there was a significant two-way interaction effect for Connective x

Language background (see Figure 4.1A) and also a significant three-way interaction effect for Connective x Language background x Syntactic knowledge (see Figure 4.1B). For the second text-related factor of interest (linear order of clauses), no significant effects were found (all *t*'s < 1.94).

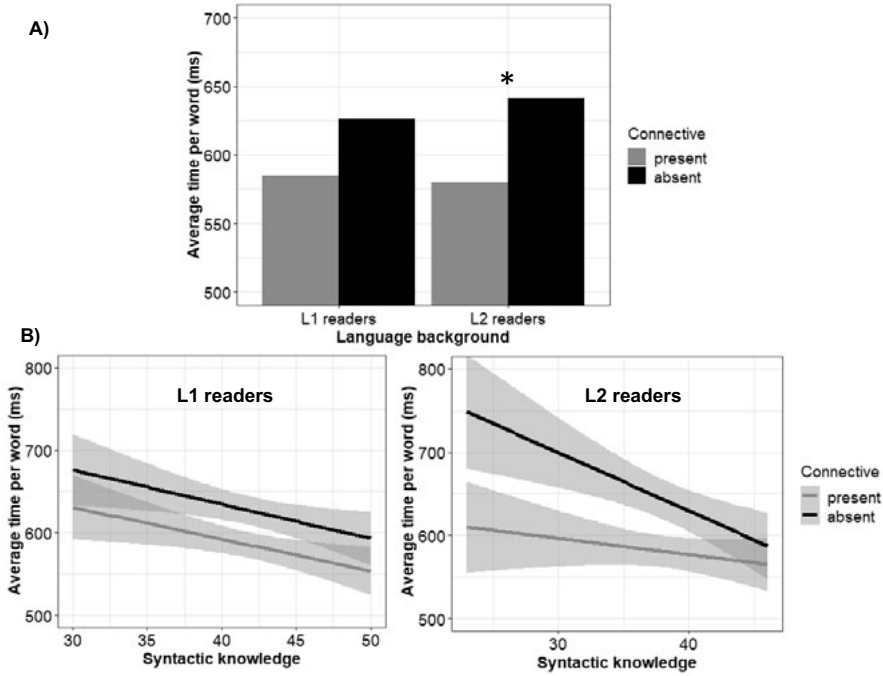

Figure 4.1 Visualization of two-way interaction effect Connective x Language background (means and standard errors in upper chart A) and three-way interaction effect Connective x Language background x Syntactic knowledge (regression lines in lower charts B) split by language background (L1 readers vs. L2 readers).

In order to clarify and determine the direction of the two interaction effects for coherence marking, the linear mixed effects analysis was split up into sub-analyses for L1 readers and L2 readers. For the L1 readers, there were no significant main effects for connective (β = -0.007, *SE* = 0.19, *t* = -0.04, *p* = .971) and syntactic knowledge (β = -0.010, *SE* = 0.01, *t* = -1.73, *p* = .085) and no interaction effect for Connective x Syntactic knowledge (β = 0.002, *SE* < 0.01, *t* = 0.45, *p* = .652). For the L2 readers, however, there was a significant main effect for connective (β = 0.550, *SE* = 0.23, *t* = 2.39, *p* = .017), but not for syntactic knowledge (β = -0.001, *SE* < 0.01, *t* = -0.26, *p* = .796), and a significant interaction effect for Connective x Syntactic knowledge (β = -0.012, *SE* = 0.01, *t* = -2.25, *p* = .025). The interaction effect (see

Figure 4.1B) showed that the lower L2 readers' syntactic knowledge, the larger the differences were in processing times of sentences with and without a connective. In other words, L2 readers with lower syntactic knowledge had longer processing times for sentences without a connective than for sentences with a connective.

Discussion

In the present study, we examined online sentence processing in third-grade monolingual Dutch children (L1 readers) and bilingual Turkish-Dutch children (L2 readers). We investigated to what extent their online processing times of sentences with causal relations were influenced by (a) coherence marking, (b) linear order of clauses, and (c) individual differences in syntactic knowledge. The results showed that both textual and child-related factors influenced beginning readers' online sentence processing. That is, individual differences in children's syntactic knowledge accounted for the effect of coherence marking on online sentence processing for L2 readers, but not for L1 readers.

The Role of Textual and Child-Related Factors

Regarding the first text-related factor (a) coherence marking, we expected that beginning readers would benefit from the presence of a connective and that this would be particularly true for L2 readers with lower syntactic knowledge. However, in contrast to previous studies showing that connectives support young, monolingual readers' text processing and comprehension (Cain & Nash, 2011; Land, 2009, Van Silfhout et al., 2014), we did not find a main effect of coherence marking in the present study. One explanation for this discrepancy concerns the context in which the connectives were presented. In previous studies the connectives were presented in short stories (Cain & Nash, 2011) or expository texts (Van Silfhout et al., 2014), which are more challenging compared to the relatively short and simple sentences as used in the present study. In order to minimize the effect of vocabulary knowledge, the sentences contained all familiar words, there was little variation in sentence structure, and the predictability of the causal relations was high. Possibly, most children were able to establish the causal relations in these relatively easy sentences in an efficient way without further textual support. Other explanations relate to the level of analysis and the reading mode. It may be the case that our sentence-level measure was not sensitive enough to detect a main effect of coherence marking. Whereas previous eye-tracking studies with advanced readers found speed-up effects in the region directly after the connective (Van Silfhout et al., 2014; Cozijn, Noordman, & Vonk, 2011), we looked at a more

global sentence-level measure of processing speed (see also Cain & Nash, 2011) to capture all integration processes across the sentence. This may have led to a lack in sensitivity. To illustrate, Cozijn et al. (2011) found a speed-up effect directly after the connective and a slow-down effect at the end of the sentence, which may be cancelled out when investigating processing times for the total sentence. Also, the oral reading mode in the current study may have led children to make less regressive eye movements (Vorstius, Radach, & Lonigan, 2014), which mainly accounted for the effects of coherence marking found by Van Silfhout et al. (2014) with a silent reading mode.

In accordance with our expectations, the findings showed that coherence marking did affect online sentence processing times for L2 readers with lower syntactic knowledge, as indicated by the interaction between (a) coherence marking and (c) syntactic knowledge. For the L2 readers, the effect of coherence marking was influenced by individual differences in syntactic knowledge; the lower their syntactic knowledge, the larger the differences in processing times for sentences with and without a connective. The longer processing times for sentences without a connective suggest that L2 readers with lower syntactic knowledge experienced more difficulty with processing causal relations when no connective was present. An explanation that can be put forward is that Dutch syntactic knowledge, which is related to children's linguistic proficiency in general, helps L2 readers to establish the causal coherence relation between two clauses in a more efficient way. Apparently, beginning L2 readers with lower syntactic knowledge benefit from explicit textual support, because they experience more difficulty when integrating information during online sentence processing, which is also found in other studies with young L2 learners (e.g., Felser, Roberts, Marinis, & Gross, 2003) and poor comprehenders (e.g., Barnes, Ahmed, Barth, & Francis, 2015). In line with the recent study of Colenbrander, Kohnen, Smith-Lock and Nickels (2016), our findings demonstrate the importance of considering individual differences, not only in vocabulary knowledge but also in syntactic knowledge, when investigating and remedying poor reading comprehension.

With respect to the second text-related factor (b) linear order of clauses, the results showed that reversing the linear order of clauses did not affect children's online processing time and there was no facilitating effect for the L2 readers. Based on the more flexible word order in their first language (i.e., Turkish) (Nilsson, 1991), we expected a facilitating effect for L2 readers. However, more research is needed to gain insight into possible transfer effects from L2 readers' first language on online

processing in their second language. Furthermore, we did not find support for our prediction that non-iconic (i.e., effect-cause) sentences would be easier to process than iconic (i.e., cause-effect) sentences, based on previous studies (Graesser et al., 1994; Magliano et al, 1993). This might be due to the restricted power in terms of the number of items included in the present study. Another explanation concerns the predictability of the causal relations as a consequence of either the items or the participants in our study. Possibly the high predictability of the causal relations that were used in the current study did not allow for differences in processing times between iconic and non-iconic sentences. Furthermore, the beginning readers in our sample had less linguistic experience, in particular the L2 readers, compared to the university students in previous studies. As a consequence, they may not predict upcoming words during sentence comprehension to the same extent, as found in the study of Martin and colleagues (2013).

Limitations and Suggestions for Future Research

It is important to note that design differences between the current study and previous studies may have consequences for the interpretation of the present findings. First, the children were instructed to read aloud in our study, which may have consequences for the comparability with previous studies with older, more experienced (adult) readers in which silent reading was performed. Because of the young age and the relative little reading experience of the third-grade children in our sample, we aimed to minimize the cognitive load imposed by the sentence reading task as much as possible by instructing them to read aloud, allowing the experimenter to press a button to move to the next item as soon as children finished reading. Although some caution may be warranted due to differences at the operational level, comparisons with previous studies at a conceptual level are still possible, because our choice for the oral reading mode is not likely to affect our results substantially for several reasons: a) the reading mode was similar across items and conditions; b) these young readers are used to reading aloud, which is common practice for reading instruction and assessment in primary education, and the L1 and L2 readers in our study did not differ in word decoding skills; c) most importantly, Van den Boer, Van Bergen and De Jong (2014) investigated oral and silent reading in fourth-grade children and concluded that the two reading modes are fairly similar with respect to the underlying cognitive skills.

Second, we did not relate children's online processing times to their off-line comprehension performance, in contrast to other studies (e.g., Barnes et al., 2015; De Leeuw et al., 2016; Van Silfhout et al., 2014; Wannacott et al., 2015). Previous

studies have already shown that there is a relation between fast reading and better comprehension (Rayner, Chace, Slattery, & Ashby, 2006), and in our study there was also a moderate negative correlation between children's online processing times and their comprehension score on a standardized test for reading comprehension ($r = .32$, $p = .031$). Instead, our focus was on children's online processing and possible differences in that respect between L1 readers and L2 readers. As expected, all children scored at ceiling on the comprehension questions, which allowed us to investigate the efficiency of their online processing rather than the accuracy.

Third, in line with Cain and Nash (2011) who found a speed-up effect of the presence of a connective for the sentence as a whole, we looked at children's processing times at the sentence-level in order to capture all integration processes across the sentence, whereas previous studies have found effects of coherence marking on particular regions of the sentence (Canestrelli et al., 2013; Cozijn et al., 2011; Van Silfhout et al., 2014). It may be the case that coherence marking affected children's processing in a way that was not fully captured by our sentence-level measure. Therefore, we cannot ascertain that L1 readers' processing has not been affected at all. In the current study, not only coherence marking, but also linear order and children's syntactic knowledge were factors of interest. Due to our manipulation of linear order of clauses, the clauses were not comparable in terms of length and position and therefore, a clause-by-clause analysis would not provide reliable effects. Furthermore, syntactic knowledge as a child characteristic of interest was expected to be related to the sentence as a whole rather than a specific region.

For future research it would be informative to further explore the factors that were central to the present study. With respect to coherence marking, it would be interesting to investigate whether similar effects are found when including a larger variety of connectives and a larger number of items, which also allows for more fine-grained analyses at specific regions of the sentence. In the current study we explicitly chose to investigate one specific Dutch causal connective (i.e., *omdat* 'because') which was assumed to be known by all children in our study. As a next step, however, future research should also include other connectives or types of connectives, with some types being more difficult than others (Knott & Sanders, 1998; Sanders, Spooren, & Noordman, 1992), following the studies of Cain and Nash (2011) and Crosson et al. (2008). In that respect, children's syntactic knowledge may also play an important role in comprehending and using more complex connectives during reading. To further explore the effect of word

order, it would be relevant to use causal relations that differ in predictability and to refine the assessment of children's syntactic knowledge, for instance by focusing on children's syntactic knowledge specific to word order violations or by including a measure of L2 readers' syntactic knowledge in their first language. The study by Vulchanova, Foyn, Nilsen and Sigmundsson (2014) showed that lexical knowledge in children's first language was related to sentence comprehension in their second language. In our study, the L2 readers varied in the extent to which they used their first language at home, which could mean that possible transfer effects of children's first language were more profound in children with higher linguistic proficiency in their first language.

Conclusions and Implications

To conclude, the present study demonstrates that both textual and child-related factors influence children's online processing of causal relations. With regard to word-to-text integration processes, this study shows how online measures, such as eye-tracking, can reveal individual differences in online sentence processing of beginning L1 readers and L2 readers that are related to their syntactic knowledge. In contrast to previous studies stating that L2 readers may not be proficient enough to take advantage of coherence markers in the text (Crosson & Lesaux, 2013; Degand & Sanders, 2002), our findings indicated that the presence of a connective can actually help L2 learners with lower syntactic knowledge to establish coherence relations more efficiently as long as they understand the semantic relation expressed by those connectives. By showing that coherence marking can also aid text processing in beginning L2 readers with lower syntactic knowledge, the present study adds on the study of Cain and Nash (2011), who concluded that connectives aid text processing in typically developing readers. Although future research is needed to support and elaborate on these findings, this study provides an initial step toward the use of online measures to gain insight into differences in online sentence processing of beginning L1 and L2 readers, which may partly account for L2 readers' comprehension problems. With respect to implications for educational practice, this study indicates that apart from focusing on decoding skills and vocabulary knowledge, it is also important to pay attention to knowledge and use of connectives as well as to stimulate children's syntactic knowledge in the early stages of reading development.

References

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

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi:10.1016/j.jml.2007.12.005

Barnes, M. A., Ahmed, Y., Barth, A., & Francis, D. J. (2015). The relation of knowledge-text integration processes and reading comprehension in 7th- to 12th-grade students. *Scientific Studies of Reading, 19*, 253-172. doi:10.1080/10888438.2015.1022650

Barr, D. Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48. doi:10.18637/jss.v067.i01

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, *9*, 43-61. doi:10.1207/s1532799xssr0901_4

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish-Dutch children. *Journal of Experimental Child Psychology*, *128*, 105-119. doi:10.1016/j.jecp.2014.06.007

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye-movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye-movements* (pp. 643-662). Oxford, UK: Oxford University Press.

Brus, T. B., & Voeten, M. J. M. (1979). *Een-Minuut-Test*, *Vorm A en B*. *Verantwoording en Handleiding (2e druk)* [One-Minute-Test, Form A and B. Justification and Manual (second edition)]. Lisse, the Netherlands: Harcourt Test Publishers.

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*, 429-441. doi:10.1037/a0022824

Canestrelli, A. R., Mak, W. M., & Sanders, T. J. M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, *28*, 1394-1413. doi:10.1080/01690965.2012.685885

CBS Statistics the Netherlands (2016). Bevolking naar migratieachtergrond [Population by immigration background]. Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Colenbrander, D., Kohnen, S., Smith-Lock, K., & Nickels, L. (2016). Individual differences in the vocabulary skills of children with poor reading comprehension. *Learning and Individual Differences*, *50*, 210-220. doi:10.1016/j.lindif.2016.07.021

Cozijn, R., Noordman, L. G. M., & Vonk, W. (2011). Propositional integration and world

knowledge inference: Processes in understanding because sentences. *Discourse Processes*, 48, 475-500. doi:10.1080/0163853X.2011.594421

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading*, *36*, 241-260. doi:10.1111/j.1467-9817.2011.01501.x

Crosson, A. C., Lesaux, N. K., & Martiniello, M. (2008). Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds. *Applied Psycholinguistics*, *29*, 603-625. doi:10.1017/S0142716408080260

De Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text related characteristics on student's reading behaviour and text comprehension: An eye movement study. *Scientific Studies of Reading*, *20*, 248-263. doi:10.1080/10888438.2016.1146285

De Zeeuw, M., Schreuder, R., & Verhoeven, L. (2013). Processing of regular and irregular past-tense verb forms in first and second language reading acquisition. *Language Learning*, *63*, 740-765. doi:10.1111/lang.12023

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing: An Interdisciplinary Journal*, *15*, 739-757. doi:10.1023/A:1020932715838

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, *38*, 78-103. doi:10.1598/RRQ.38.1.4

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*, 453-489. doi:10.1017.S0142716403000237

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Geva, E., & Zadeh, Z. Y. (2006). Reading efficiency in native English-speaking and English as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading*, *10*, 31-57. doi:10.1207/s1532799xssr1001_3

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395.

Halliday, M. E. K., & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman.

Janssen, C., Segers, E., McQueen, J. M., & Verhoeven, L. (2015). Lexical specificity training effects in second language learners. *Language Learning*, *65*, 358–389. doi:10.1111/lang.12102

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*, 160-212. doi:10.1111/lang.12034

Juhasz, B. J., & Pollatsek, A. (2011). Lexical influences on eye movements in reading. In S. Liversedge, I. Gilchrist, & S. Eberling (Eds.), *Handbook on Eye movements* (pp. 873-894). Oxford, UK: Oxford University Press.

Kaan, E., Kirkham, J., & Wijnen, F. (2016). Prediction and integration in native and second language processing of elliptical structures. *Bilingualism: Language and Cognition*, *19*, 1-18. doi:10.1017/S1366728914000844

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, *94*, 113-147. doi:10.1016/j.cognition.2004.01.002

Knott, A., & Sanders, T. J. M. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, *30*, 135-175. doi:10.1016/S0378-2166(98)00023-X

Kort, W., Schittekatte, M., Dekker, P.H., Verhaeghe, P., Compaan, E.L., Bosmans, M., & Vermeir, G. (2005). *WISC-III NL. Handleiding en Verantwoording* [Manual and Justification]. London, UK: The Psychological Corporation.

Land, J. F. H. (2009). *Zwakke lezers, sterke teksten? Effecten van tekst- en lezerskenmerken op het tekstbegrip en de tekstwaardering van vmbo-leerlingen* [Less-skilled readers, well-built texts? Effects of text and reader characteristics on text comprehension and text appreciation of students in prevocational education]. Delft: Eburon.

Lervåg, A. & Aukrust, V.G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*, 612-620. doi:10.1111/j.1469-7610.2009.02185.x

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, *19*, 99-131. doi:10.1007/s11145-005-4713-6

Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, *39*, 1005-1019. doi:10.1037/0012-1649.39.6.1005

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, *65*, 185-207. doi:10.1111/lang.12117

Lipka, O., & Siegel, L. S. (2007). The development of reading skills in children with English as a second language. *Scientific Studies of Reading*, *11*, 105-131, doi:10.1080/10888430709336555

Mak, W. M., & Sanders, T. J. M. (2013). The role of causality in discourse processing: Effects of expectation and coherence relations. *Language and Cognitive Processes*, *28*, 1414-1437. doi:10.1080/01690965.2012.708423

Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedents and causal consequence inferences. *Discourse Processes*, *16*, 35-53. doi:10.1080/01638539309544828

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority children. *Journal of Educational Psychology*, *102*, 701-711. doi:10.1037/a0019135

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*, 574-588. doi:10.1016/j.jml.2013.08.001

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409-433. doi:10.1037/a0033890

Nagy, W. E., Mcclure, E. F., & Mir, M. (1997). Linguistic transfer and the use of context by Spanish-English bilinguals. *Applied Psycholinguistics*, *18*, 431-452. doi:10.1017/S0142716400010924

Noordman, L. G. M., & De Blijzer, F. (2000). On the processing of causal relations. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause Condition Concession Contrast: Cognitive and Discourse Perspectives* (pp. 35-56). Berlin, Germany: Mouton de Gruyter.

Noordman, L. G. M., & Vonk, W. (1998). Memory-based processing in understanding causal information. *Discourse Processes*, *26*, 191-212. doi:10.1080/01638539809545044

Nilsson, B. (1991). Turkish semantics revisited. In H. Boeschoten and L. Verhoeven (Eds.), *Turkish Linguistics Today* (pp. 93-112). Leiden, the Netherlands: E. J. Brill.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, *18*, 443-468. doi:10.1080/01690960344000008

Pearson, P. D. (1974). The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. *Reading Research Quarterly*, *10*, 155-192. doi:10.2307/747180

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22-37. doi:10.1080/10888438.2013.827687

Perfetti, C., Yang, C., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*, 303-318. doi:10.1002/acp.1419

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rayner, K., Chace, K. H., Slattery, T. J., Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*, 241-255. doi:10.1207/s1532799xssr1003_3

Sanders, T., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*, 37-60. doi:10.1207/S15326950dp2901_3

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1-35. doi:10.1080/01638539209544800

Schaerleakens, A., Kohnstamm, D., Lejaegere, M. (1999). *Streeflijst woordenschat voor zesjarigen* [Vocabulary list for six-year-olds]. Lisse, the Netherlands: Swets en Zeitlinger.

Scheele, A.F., Leseman, P. P. M., & Mayo, A. F. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*, 117-140. doi:10.1017/S0142716409990191

Van den Boer, M., Van Bergen, E., & De Jong, P. F. (2014). Underlying skills of oral and silent reading. *Journal of Experimental Child Psychology*, *128*, 138-151. doi:10.1016/j.jecp.2014.07.008

Van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. D. M. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading*, *31*, 203–223. doi:10.1111/j.1467-9817.2007.00354.x

Van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. M. (2014). Connective and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology*, *106*, 1036-1048. doi:10.1037/a0036293

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*, 313-330. doi:10.1207/S1532799XSSR0404_4

Verhoeven, L.T.W., Keuning, J., Horsels, L., Van Boxtel, H. (2013). *Testinstrumentarium Taalontwikkelingsstoornissen* [Set of Instruments Language Developmental Disorder]. Arnhem, the Netherlands: Cito.

Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral reading. *Visual Cognition*, *22*, 458-485. doi:10.1080/13506285.2014.881445

Vulchanova, M., Foyn, C. H., Nilsen, R. A., Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year-old children. *Learning and Individual Differences*, *35*, 87-95. doi:10.1016/j.lindif.2014.07.016

Wannacott, E., Joseph, H. S. S. L., Adelman, J. S., & Nation, K. (2015). Is children's reading "good enough"? Links between online processing and comprehension as children read syntactically ambiguous sentences. *The Quarterly Journal of Experimental Psychology*, *69*, 855-879. doi:10.1080/17470218.2015.1011176

Appendix A

1A. Omdat Teun honger had, at hij een wafel.
[Because Teun was hungry, he ate a waffle.]
1B. Joep at een appel, omdat hij honger had.
[Joep ate an apple, because he was hungry.]
1C. Daan had honger. Hij at een tosti.
[Daan was hungry. He ate a tosti.]
1D. Guus at een lolly. Hij had honger.
[Guus ate a lolly. He was hungry.]
Comprehension question: Waarom at (…) een (…)? [Why did (…) eat a/an (…)?]

2A. Omdat Roos dorst had, dronk ze een glas cola.
[Because Roos was thirsty, she drank a glass of coke.]
2B. Suus dronk een glas water, omdat ze dorst had.
[Suus drank a glass of water, because she was thirsty.]
2C. Tess had dorst. Ze dronk een glas melk.
[Tess was thirsty. She drank a glass of milk.]
2D. Maud dronk een glas thee. Ze had dorst.
[Maud drank a glass of tea. She was thirsty.]
Comprehension question: Waarom dronk (…) een glas (…)? [Why did (…) drink a glass of (…)?]

3A. Omdat Mark jarig was, kreeg hij een fiets van zijn ouders.
[Because it was Mark's birthday, he got a bike from his parents]
3B. Bart kreeg een trein van zijn ouders, omdat hij jarig was.
[Bart got a train from his parents, because it was his birthday.]
3C. Jens was jarig. Hij kreeg een piano van zijn ouders.
[It was Jens's birthday. He got a piano from his parents.]
3D. Nick kreeg een gitaar van zijn ouders. Hij was jarig.
[Nick got a guitar from his parents. It was his birthday.]
Comprehension question: Waarom kreeg (…) een (…) van zijn ouders? [Why did (...) got a (...) from his parents?]

4A. Omdat Sara gevallen was, had ze veel pijn aan haar neus.
[Because Sara had fallen, she hurt her nose.]
4B. Lara had veel pijn aan haar been, omdat ze gevallen was.
[Lara hurt her leg, because she had fallen.]
4C. Anja was gevallen. Ze had veel pijn aan haar pols.
[Anja had fallen. She hurt her wrist.]
4D. Lisa had veel pijn aan haar knie. Ze was gevallen.
[Lisa hurt her knee. She had fallen.]
Comprehension question: Waarom had (…) veel pijn aan haar (…)? [Why did (...) hurt her (...)?]

Note. Each item consisted of two clauses that were separated by either a comma or full stop. In A-items: connective = present, order = cause-effect. In B-items: connective = present, order = effect-cause. In C-items: connective = absent, order = cause-effect. In D-items: connective = absent, order = effect-cause.
Frequent Dutch names were used and within each block they were matched on gender and number of syllables.

Chapter 5

Online processing of coherence relations in beginning readers from diverse linguistic backgrounds

This chapter is based on:
Van den Bosch, L. J., Segers, E., & Verhoeven, L. (submitted). *Online processing of coherence relations in beginning readers from diverse linguistic backgrounds.* Manuscript submitted for publication.

Abstract

In the present study, two experiments were conducted to investigate online processing of causal coherence relations (Experiment 1) and temporal coherence relations (Experiment 2) in beginning readers from diverse linguistic backgrounds. The participants were 47 Dutch monolingual children (L1 readers) and 41 Turkish-Dutch bilingual children (L2 readers) in third grade. Children's eye movements were recorded to assess online processing of two-clause sentences including a causal or temporal relation and measures of syntactic knowledge and working memory were administered. Results from the two experiments revealed effects of both textual and child-related factors. Children's online processing of coherence relations was affected by the type of connectives, and not by the order of clauses. More familiar causal/temporal connectives led to shorter processing times. Furthermore, L2 readers generally showed more difficulty with processing coherence relations than L1 readers, indicated by longer processing times for the sentence as a whole and for the two clauses separately. Across the two groups of participants, syntactic knowledge explained variation in children's online processing at the sentence level, whereas working memory did so at the clause level. These findings highlight the importance of paying attention to textual demands as well as children's linguistic and cognitive skills to support their comprehension processes.

Introduction

When reading a text, connectives such as 'because' or 'before' help the reader in making meaningful connections between clauses. This is needed to establish local coherence, and, ultimately a coherent mental representation of the text (Graesser, Singer, & Trabasso, 1994; Van Dijk & Kintsch, 1983). Connectives may be specifically challenging for young readers (Cain & Nash, 2011), and even more so for children who read in their second language (Crosson, Lesaux, & Martiniello, 2008). However, despite the frequent occurrence of connectives in text and their importance for text comprehension, it is by no means clear how beginning readers from diverse linguistic backgrounds deal with these linguistic markers to establish coherence relations during reading. In the present study, two experiments were conducted to investigate online processing of causal coherence relations (Experiment 1) and temporal coherence relations (Experiment 2) in beginning first (L1) and second language (L2) readers. The current study aimed to provide more detailed insight into children's online processing of coherence relations by examining their eye movements during reading. To shed light on the role of linguistic diversity, not only children's language background (L1 vs. L2) but also individual differences in syntactic knowledge and working memory capacity were examined in order to explain individual variation in children's online processing.

Background to the Study

Reading comprehension is one of the most crucial skills children acquire at primary school, but it is a very complex activity that involves a range of different skills, processes and types of knowledge. During the process of reading, the readers needs to connect bottom-up word identification processes and top-down comprehension processes within a cognitive system with limited processing resources (Kintsch, 2005) and thereby draws upon linguistic knowledge, orthographic knowledge and general (world) knowledge (Perfetti & Stafura, 2014). To acquire local comprehension, the reader has to integrate word meanings in the context of the sentence or short text, also referred to as word-to-text integration (Perfetti, Yang, & Schmalhofer, 2008). In order to do so, the reader can make use of linguistic markers in the text, such as connectives. Connectives signal the semantic relation between clauses or sentences and can guide the reader in how to construct meaning by coding coherence relations (Sanders & Noordman, 2000). Connectives can be classified according to the type of semantic relationship they signal (i.e., additive, adversative, causal, and temporal) with some types being more difficult than others (Halliday & Hasan, 1976; Sanders, Spooren, & Noordman, 1992), which

is also reflected in the order of acquisition in children's language (Evers-Vermeul & Sanders, 2009). From an early age on, children start to combine simple clauses into complex sentences while they learn about the syntactic structure of sentences, the (syntactic) meaning of connectives and the semantic relation they encode (Diessel, 2004).

With respect to reading, connectives are assumed to be important for text comprehension and frequently occur in written texts, even for beginning readers. A corpus study examining the order of appearance and frequency of connectives in Dutch showed that there is a sharp increase in the number of connectives present in texts for first grade compared to second grade (Tellings & Penning-de Vries, 2019). Nevertheless, connectives can be challenging for young readers who have limited knowledge of how to use connectives (Cain & Nash, 2011; Cain, Patson, & Andrews, 2005). In that respect, previous studies have investigated children's comprehension of temporal coherence relations expressed by the connectives *before* and *after* (e.g., 'The girl ate an apple before she played outside.'). To comprehend this sentence, the reader needs to understand the meaning of the temporal connective to establish the correct order of events and to integrate this information in a mental model of the text (Zwaan, 1996). Acquiring the meaning of *after* is more difficult than *before* due to its linguistic complexity, and sentences that describe events in chronological order are easier to comprehend than non-chronological order assuming that young children follow an order-of-mention strategy (Clark, 1971). This is evidenced in a study with 3-to-7-year-old English children listening to sentences with temporal connectives (Blything & Cain, 2016); the older children (5-7 years) showed more robust understanding of the meaning of *before* and *after* than the younger children (3-5 years), but the older children still showed difficulties with understanding reverse order sentences, which are assumed to have higher processing costs. Moreover, two studies with 8-to-12-year old Finnish and Dutch children reading sentences with temporal connectives demonstrated that even older children were not fully proficient in comprehension depending on sentence chronology and/or the position of the connective (Karlsson, Jolles, Koornneef, Van den Broek, & Van Leijenhorst, 2019; Pyykkönen & Järvikivi, 2012). Of interest, these studies also attempted to examine individual differences by investigating the role of working memory. However, the results were inconclusive and these studies only included monolingual children.

In terms of linguistic diversity, including children from diverse linguistic backgrounds (L1 vs. L2) provides an interesting perspective to study variation

in children's comprehension of coherence relations and their use of connectives. First, connectives may be particularly challenging for bilingual children who learn to read in their second language. These L2 readers often show delays in terms of oral language skills and reading comprehension performance compared to their monolingual peers (Melby-Lervåg & Lervåg, 2014). As a result, L2 readers may not have sufficient linguistic knowledge in their second language to take advantage from the linguistic marking of coherence relations provided by connectives (Degand & Sanders, 2002). Indeed, previous studies showed that no relation or a weaker relation between knowledge of connectives and reading comprehension existed for fifth-grade L2 readers of English and ninth-grade L2 readers of Dutch, while a positive, stronger relation was found for their monolingual peers (Crosson & Lesaux, 2013; Kohnen & Retelsdorf, 2019). Second, in addition to revealing group differences between L1 readers and L2 readers, investigating the role of linguistic diversity can provide insights into the underlying linguistic and cognitive factors, such as syntactic knowledge and working memory. These factors have been associated with children's language background; L2 readers often have lower (L2) syntactic knowledge compared to L1 readers (Farnia & Geva, 2013; Lipka & Siegel, 2007), whereas their working memory capacity is suggested to be larger because of their experience with dual language management (Blom et al., 2014; Grundy & Timmer, 2017). Moreover, these factors have also been associated with children's reading comprehension (Cain, Oakhill, & Bryant, 2004; Muter, Hulme, Snowling, & Stevenson, 2004) and are assumed to be important for processing of complex sentences in terms of syntactic complexity and cognitive capacity (Just & Carpenter, 1992).

Not only our knowledge of the underlying factors, but also our understanding of young children's actual online processing of coherence relations is limited. To gain a full picture of how children process coherence relations online measures are required that expose moment-to-moment comprehension processes. Previous studies have demonstrated that connectives facilitate online processing in adults (Canestrelli, Mak, & Sanders, 2013; Cozijn, Noordman & Vonk, 2011), but studies using online measures to investigate the effect of connectives in younger readers are scarce and restricted to monolingual readers. For adolescents, the study of Van Silfhout, Evers-Vermeul and Sanders (2015) showed that connectives led to faster processing of subsequent words and shorter rereading times of previous words suggesting that they function as immediate processing instructions for 12-to-15-year-old readers. Vorstius, Radach, Mayer and Lonigan (2013) investigated 11-year-old children's processing of coherence relations with a causal connective (because)

or adversative connective (although) and showed that especially the adversative connective was associated with more processing difficulty as demonstrated by longer reading times for both the initial and final clause. One of the few studies with younger children (Cain & Nash, 2011) showed that although knowledge of connectives is still developing, they can aid text processing in 8-to-10-year-old children. For causal (because, so), temporal (before, after) and adversative (but, although) coherence relations, children's reading times were shorter when an appropriate connective was used compared to an inappropriate or neutral connective (and).

Taken together, previous research has shown that connectives are important for reading comprehension, but can be challenging for young readers, especially L2 readers, since children's knowledge of syntactic structures, on the one hand, and of the meaning of connectives and the semantic relation they signal, on the other hand, is still developing. Although studies with monolingual readers showed that connectives affect the reader's online processing, little is known about how beginning readers from diverse linguistic backgrounds actually process coherence relations and how individual variation can be explained. In light of the increasing number of L2 readers in primary education and the growing use of online measures, such as eye-tracking, in children's reading (Blythe & Joseph, 2011), it is both timely and relevant to investigate online processing of coherence relations in a linguistically diverse group of beginning readers. As a first attempt, we conducted a small-scale study (see Chapter 4) to examine the effect of connective presence and order of clauses in beginning L1 and L2 readers (Van den Bosch, Segers, & Verhoeven, 2018). The results showed that L2 readers with lower syntactic knowledge had shorter processing times for sentences with a causal connective (*omdat* 'because') compared to sentences without a causal connective, suggesting that the connective helped them to establish the causal coherence relation more efficiently. The other textual factor of interest, order of clauses (cause-effect vs. effect-cause), did not affect children's online processing. This study was, however, limited to one type of (causal) connective, and did not examine online processing at the clause level, while previous studies revealed differential effects at particular sentence regions rather than the sentence as a whole in adults and adolescents.

The Present Study

The aim of the present study was to provide more detailed insight into online processing of coherence relations in beginning readers from diverse linguistic backgrounds and the factors that explain individual variation among children.

In two separate, but related experiments, two types of semantic relations were investigated that frequently occur in written text. Experiment 1 focused on causal coherence relations and the Dutch causal connectives *omdat* and *doordat* ('because'). Experiment 2 focused on temporal coherence relations and the Dutch temporal connectives *voordat* ('before') and *nadat* ('after'). The two experiments were combined in an intertwined experimental reading task subjected to a group of 47 Dutch monolingual children (L1 readers) and 41 Turkish-Dutch bilingual children (L2 readers) in third grade of primary school. In the Netherlands, Turkish is the largest linguistic minority (CBS, 2016) and the most widely spoken additional language among Dutch primary school children (Extra, 2002). With respect to processing, eye-tracking technology was used as an online measure to gain detailed insight into children's online processing at the sentence and clause level. With respect to individual variation, children's language background (L1 vs. L2) as well as the role of syntactic knowledge and working memory were investigated to gain a better understanding of the factors involved in processing coherence relations. In the remainder of this chapter, the two experiments are presented in separate sections, followed by a general discussion.

Experiment 1: Causal Coherence Relations

The first experiment was conducted to elaborate on a previous small-scale study on online processing of causal coherence relations (e.g., 'Tom ate an apple, because he was hungry.') in beginning L1 and L2 readers (Van den Bosch et al., 2018; see Chapter 4). By using a similar but larger sample and an eye-tracker with higher temporal resolution, the present study allows for conducting more reliable and fine-grained analyses of the effect of coherence marking (i.e., presence vs. absence of a connective) and order of clauses (i.e., cause-effect vs. effect-cause). Apart from children's sentence processing times, we now also examined children's processing times for the two clauses and the connectives specifically to gain more detailed insight into the effect of connectives on particular sentence regions. Furthermore, in addition to the two-clause sentences with the Dutch causal connective *omdat* retrieved from our previous study, we constructed an extra set of sentences with the Dutch causal connective *doordat*, which has a comparable meaning ('because'), but is less frequent (Tellings, Hulsbosch, Vermeer, & Van den Bosch, 2014).

The following research questions were addressed:

1. To what extent are L1 and L2 readers' processing of causal coherence relations affected by connective presence and order of clauses?
2. To what extent can L1 and L2 readers' variation in processing of causal coherence relations be explained from syntactic knowledge and working memory?

Based on previous studies, we expected that the presence of a causal connective would aid children's online processing, as evidenced by shorter processing times, particular for the clause directly after the connective. Furthermore, it was expected that any processing differences between L1 readers and L2 readers could be largely explained by individual differences in syntactic knowledge.

Method

Participants. A total of 88 third-grade children from six primary schools in the Netherlands participated. These were monolingual schools with Dutch as the language of instruction, located in urban areas with a mixed student population. The current sample was a representative subsample from our longitudinal study on reading comprehension in beginning first and second language readers (see Chapter 3). The six (out of thirteen) schools with the largest number of participants were selected to take part in the present study. The children (51 girls and 37 boys) were

aged between 8 and 10 years (*M* = 8;11, *SD* = 0;5). The sample included monolingual children who had Dutch as their home language (L1 readers; *n* = 47, 57% boys) and bilingual children who had Turkish as their home language (L2 readers; *n* = 41, 58% boys). The majority of all L2 readers in this sample where born in the Netherlands (95.1%) and spoke both Turkish and Dutch at home, with more than half having Turkish as their main home language (58.5%). The L2 readers indicated to read almost exclusively in Dutch; the majority was not literate in Turkish (82.9%) since they received formal reading instruction only in Dutch at school, starting from first grade onward.

The two language groups did not differ with respect to age ($t(86) = -0.05$, $p = .958$, $d = 0.01$) and non-verbal reasoning ($t(86) = 0.17$, $p = .866$, $d = 0.03$), and differences in terms of their socioeconomic status (i.e., averaging parents' level of education measured on a four-point scale ranging from 1 = primary school to 4 = higher education/university) were negligible ($U = 734.50$, $z = -1.97$, $p = .049$, $r = -.21$). Since the broader longitudinal study included a large number of measures, we were also able to check for between-group differences in terms of language and literacy skills (see Table 5.1). As expected, the L2 readers scored considerably lower on standardized tests for receptive vocabulary knowledge (Verhoeven, Keuning, Horsels, & Van Boxtel, 2013), connective understanding (Verhoeven & Vermeer, 1993) and reading comprehension (Cito B.V., 2014) compared to the L1 readers, whereas no significant group differences were found for (pseudo) word decoding efficiency (Verhoeven & Keuning, 2017).

Table 5.1 Descriptive Statistics and Mean Differences for L1 readers and L2 readers

	L1 readers (*n* = 47)		L2 readers (*n* = 41)		Mean differences		
Background variables	*M (SD)*	*range*	*M (SD)*	*range*	*t*	*p*	*d*
Decoding skills: words	218 (44)	123-326	205 (46)	92-284	1.28	.205	0.27
Decoding skills: pseudo	148 (42)	40-257	150 (37)	59-213	-0.28	.783	0.06
Vocabulary knowledge	50 (4)	40-59	41 (6)	26-57	8.15	< .001	1.70
Connective understanding	28 (6)	18-39	23 (5)	11-36	3.97	< .001	0.85
Reading comprehension	154 (28)	111-280	137 (20)	86-188	3.34	.001	0.72

Measures

Stimuli causal coherence relations. Children were presented with 32 two-clause Dutch sentences describing a causal relation. Sentences were written in past tense and ranged from seven to thirteen words. In order to minimize the effect of

vocabulary knowledge, the sentences only included words that were assumed to be familiar to all third-grade children based on a Dutch wordlist for six-year olds (Schaerleakens, Kohnstamm, & Lejaegere, 1999) and the described events were commonly occurring and familiar to children this age. The two events depicted in the causal relations reflected a real-world causality between two events/states in the world. Sixteen sentences with the connective *omdat* 'because' were retrieved from our previous study (Van den Bosch et al., 2018; see Chapter 4) and we constructed a similar, additional set of 16 sentences with another, less frequent causal connective *doordat* (see Appendix B for an overview of all items). Two factors were manipulated: Connective presence (present vs. absent) and order of clauses (effect-cause vs. cause-effect) resulting in a 2x2 design with four within-subject conditions (see Table 5.2). Each participant was presented with eight items in each condition (32 items in total), in pseudo-randomized order. After each sentence, children answered a simple recall question to check whether they performed the task seriously. Out of two alternatives they had to select the correct object ('*What did he eat? An apple/a pizza*') or subject name ('*Who was too late for school? Sam/Eva*'). This turned out to be the case; only 57 out of 2816 responses were incorrect (2.0%) and these were unrelated to any of the three factors of interest. Furthermore, our previous study (see Chapter 4) with a very similar sample already showed that children were highly able to establish the causal coherence relation between clauses (Van den Bosch et al., 2018), therefore only the online processing data were subjected to statistical analyses.

Table 5.2 Overview of the Four Within-Subject Conditions for Causal Coherence Relations

Condition	Item	Connective	Linear order
1	Tom at een appel, omdat hij honger had. [Tom ate an apple, because he was hungry.]	present	effect-cause
2	Omdat Tom honger had, at hij een appel. [Because Tom was hungry, he ate an apple.]	present	cause-effect
3	Tom at een appel. Hij had honger. [Tom ate an apple. He was hungry.]	absent	effect-cause
4	Tom had honger. Hij at een appel. [Tom was hungry. He ate an apple.]	absent	cause effect

Syntactic knowledge. A standardized grammaticality judgment task was used to assess children's syntactic knowledge (T-TOS; Verhoeven et al., 2013). Children listened to 55 sentences and had to indicate whether the sentences were grammatically correct (17 sentences) or incorrect (38 sentences). Incorrect

sentences included incorrect inflection of nouns (10 sentences), incorrect verb conjugation (10 sentences), incorrect order of words (10 sentences), or incorrect subject-verb agreement (8 sentences). The total score is determined by the number of items answered correctly with a maximum score of 55 (Cronbach's α = .85 in the present study).

Working memory. The Digit Span Backwards (WISC-NL-III; Kort et al., 2005) was administered to assess children's working memory capacity. In this standardized task, children are presented with a spoken sequence of digits and have to repeat the sequence in reverse order. Each item consists of two sequences of the same length, starting with two-digit sequences. When the child answered at least one sequence of an item correctly, the next item was presented with a longer sequence (one extra digit). The task stopped when a child gave an incorrect answer to both sequences of one item. The score for children's working memory capacity was determined by the number of sequences repeated correctly with a maximum score of 16 (Cronbach's α = .64, Kort et al., 2005).

Apparatus. While children were reading the two-clause sentences, their eyemovements were recorded with an SMI RED500 eye-tracking system (SensoMotoric Instruments, Germany). This is a portable, screen-based eye-tracker that captures gaze data with a 500Hz sampling rate by using a binocular tracking technique. The distance between children's eyes and the monitor (22 inch, resolution 1680x1050) was 70cm. For the purpose of maximizing ecological validity and because of the system's head movement compensation, we did not use a chinrest. Single-line sentences were successively presented in the center of the screen in the monospaced font Courier New in black against a light grey background. A nine-point calibration procedure was used to locate the position of the eyes. In case of poor quality calibration (accuracy > 1.0) recalibration was done. Afterwards, the data were inspected visually to detect any problems. Data of two children (one L1 reader and one L2 reader) were excluded because of overall low tracking ratio (< 50%). By means of SMI BeGaze analysis software, we determined children's total fixation times at three specific regions of the sentence: 1) the first clause, 2) the final clause, and 3) the connective (if present), to illustrate: [Because]^{3} [Tom was hungry]^{1} [he ate an apple]^{2}.

Procedure. Prior to the start of the study, ethical approval was obtained from our institution's ethics committee and we asked for parental consent and children's assent. The data were collected halfway the third-grade school year

by the first author and two trained research assistants. The children took part in two individual test sessions (30 minutes each) in a quiet room at their school. In the first session, children completed several standardized tests to assess their language and literacy skills, including the syntactic knowledge test. In the second session, the experimental reading task was administered. The items of Experiment 1 (causal relations) and Experiment 2 (temporal relations) were mixed in two blocks, each consisting of 16 causal items and 16 temporal items, presented in pseudo-randomized order. Alternating causal and temporal items within one block increased task variation and by doing so the different types of items served as filler items for each other. In between blocks, children completed the working memory test. The first block was preceded by four example items and at the start of each block there was a calibration procedure to locate the position of the eyes. Children were instructed to read silently while comprehending what they were reading. Once they completed reading the sentence, they had to press an orange button and then then a comprehension question appeared. After answering this question by pressing either the corresponding left or right-sided grey button, a neutral screen with a fixation point was presented for 1000 milliseconds and then the next sentence appeared. Children's eye-movements, sentence reading times (ms), responses and response times (ms) were recorded.

Data Analysis. A total of 88 children reading 32 sentences resulted in 2816 observations. First, sentence reading times were analyzed and then fixation times per region (first clause, final clause, connective) were analyzed. Prior to the analyses, we checked all outcome variables for outliers and deviations from normality. All values two standard deviations above or below item *and* subject mean were excluded from the analysis (cf. Van den Bosch et al., 2018). As a result, we excluded 64 observations (2.3%) for sentence reading times, 44 observations (1.6%) for first clause fixation times, 53 observations (1.9%) for final clause fixation times, and 31 observations (2.2%) for connective fixation times. For all outcome variables, the remaining data were within the range of normal distribution, therefore no log-transformations were required.

Linear mixed-effects regression analyses (Baayen, Davidson, & Bates, 2008) were performed in the R statistics environment (version 3.4.4; R Core Team, 2018) using the lmer() function from the lme4 package (version 1.1-17; Bates, Maechler, Bolker & Walker, 2015) in a similar, stepwise way for each outcome variable. To address the first research question, we tested a model with Connective, Order, Language background, and their interactions as fixed effects. Random effects were fit using

a maximal random effects structure (Barry, Levy, Scheepers, & Tily, 2013) with random intercepts for participants and items, and by-subject random slopes for Connective, Order, and their interaction, and a by-item random slope for Language background. Then, regarding the second research question, we added Syntactic knowledge and Working memory (mean-centered) as fixed effects to the model. Models were fit using restricted maximum likelihood technique. A fixed effect was considered significant if the absolute value of the *t*-statistic was larger than or equal to 2.0 (Gelman & Hill, 2007). Since lmer does not automatically provide *p*-values, these were derived using Satterthwaite approximations from the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) which produces acceptable Type 1 error rates even for smaller samples (Luke, 2017).

Results

First, we examined group differences between L1 readers and L2 readers with respect to the child-related factors of interest. Independent samples t-tests showed that the L2 readers (M = 34.5, SD = 6.1) scored significantly lower than L1 readers (M = 42.5, SD = 5.0) on syntactic knowledge with a large effect size, $t(86) = 6.69$, $p < .001$, $d = 1.43$. No significant difference between L1 readers (M = 4.1, SD = 1.6) and L2 readers (M = 4.6, SD = 1.6) existed for working memory capacity although a small effect size was observed, $t(86) = -1.37$, $p = .173$, $d = 0.31$. Next, we conducted linear mixed effects analyses for all outcome variables; regression coefficients, standard errors and t-values for all fixed effects are shown in Table 5.3.

Table 5.3 Results Causal Relations: Sentence Reading Times and Clause Fixation Times

	Model 1			Model 2		
	β	*SE*	*t*	*β*	*SE*	*t*
Sentence reading times						
Connective	-721.77	297.49	-2.43*	-721.57	297.42	-2.42*
Order	-372.74	297.52	-1.25	-371.88	297.29	-1.25
LB	909.37	317.09	2.87**	456.79	368.58	1.24
Connective x Order	335.90	420.86	0.80	335.21	420.54	0.80
Connective x LB	-139.17	207.91	-0.67	-138.67	207.78	-0.67
Order x LB	272.59	208.11	1.31	273.38	207.50	1.32
Connective x Order x LB	99.65	294.44	0.34	98.54	293.58	0.34
WM				-15.01	87.96	-0.17
SK				-57.88	24.73	-2.34*
First clause: effect-cause						
Connective	-47.97	156.13	-0.31	-47.53	155.81	-0.31
LB	739.40	166.29	4.45***	488.20	202.96	2.41*
Connective x LB	4.09	110.02	0.04	3.68	109.43	0.03
WM				-79.63	50.11	-1.59
SK				-34.31	14.48	-2.37*
First clause: cause-effect						
Connective	-119.50	215.49	-0.56	-119.57	215.52	-0.56
LB	574.50	154.34	3.72***	388.73	178.56	2.18*
Connective x LB	9.63	20.67	0.08	10.56	120.98	0.09
WM				-95.00	40.13	-2.37*
SK				-27.22	11.61	-2.34*
Final clause: effect-cause						
Connective	101.33	129.39	0.78	101.06	129.47	0.78
LB	315.27	101.93	3.09**	314.66	120.90	2.60*
Connective x LB	98.08	98.73	0.99	98.63	98.82	1.00
WM				-69.14	28.09	-2.46*
SK				-3.51	8.14	-0.43
Final clause: cause-effect						
Connective	21.86	142.96	0.15	21.72	143.02	0.15
LB	432.01	132.87	3.25**	366.80	162.65	2.26*
Connective x LB	-31.75	88.91	-0.36	-31.73	88.97	-0.36
WM				-89.04	40.53	-2.20*
SK				-12.32	11.70	-1.05
Connective fixation time						
Order	-31.51	36.45	-0.87	-32.13	36.47	-0.88
LB	84.93	44.15	1.93	41.24	50.29	0.82
Order x LB	33.96	39.80	0.85	34.40	39.84	0.86
WM				-14.12	50.29	-1.30
SK				-6.02	10.90	-1.91

Note. Connective = Connective presence (0 = present, 1 = absent), Order = Order of clauses (0 = cause-effect, 1 = effect cause), LB = Language background (0 = L1, 1 = L2), WM = Working memory, SK = Syntactic knowledge.
Formula model 1: outcome ~ Connective*Order*LB + (1+Connective*Order | subject) + (1| item); model 2: outcome ~ Connective*Order*LB + WM + SK + (1+Connective*Order | subject) + (1| item).
*p < .05, **p < .01, ***p < .001.

Sentence reading times. The initial model showed that there was a significant main effect of Connective; children's sentence reading times were longer when a causal connective was present. There was also a significant main effect of Language Background; the L2 readers had longer sentence reading times compared to the L1 readers. In the second model, where syntactic knowledge and working memory were added as fixed effects, there was again a significant main effect of Connective, but the main effect of Language background was no longer significant. Instead, there was now a significant main effect of syntactic knowledge; lower syntactic knowledge was associated with longer reading times.

It should be noted, however, that differences in sentence length existed across conditions due to our manipulation of connective presence. When controlling for this by calculating the average reading time per word, the main effect of connective presence was no longer significant in both models (M1: β = -22.84, *SE* = 27.21, *t* = -0.84, *p* = .407; M2: β = -22.86, *SE* = 27.22, *t* = -0.84, *p* = .407), indicating that the longer reading times for sentences with a connective could be attributed to the inclusion of an additional word.

Clause fixation times. Next, we investigated children's fixation times for the two clauses separately to gain more detailed insight into their online processing. Due to our manipulation of order of clauses, the first and final clause are not comparable in terms of content and length across the four conditions. Therefore, the analyses were split-up by order so that separate models were tested for effect-cause sentences and cause-effect sentences.

Fixation times first clause. For effect-cause sentences, only a significant main effect of Language background was found in the initial model, indicating that L2 readers had longer fixation times at the first clause. In the second model, the main effect of Language background was still significant and in addition, there was a significant main effect of Syntactic knowledge, indicating that children with lower syntactic knowledge had longer fixation times on the first clause. For cause-effect sentences, similar effects were found, but in the second model there was an additional significant effect of working memory, showing that children with lower working memory capacity had longer fixation times on the first clause.

Fixation times final clause. For effect-cause sentences, only a significant main effect of Language Background was found, indicating that L2 readers had longer fixation times at the final clause. In the second model, the main effect of Language

background was still significant and in addition there was a significant main effect of working memory, indicating that children with lower working memory had longer fixation times at the final clause. For cause-effect sentences, similar effects were found.

Connective fixation times. Finally, we also investigated children's fixations times at the connective itself for half of the items in which a causal connective was present. Comparable patterns were observed, but no significant effects were found. In the initial model, there was only a marginally significant effect of Language background, whereas there was a marginally significant effect of Syntactic knowledge in the second model.

Although it was not the focus of the current experiment, we performed some exploratory analyses to check for differences between the two types of causal connectives (*omdat* vs. *doordat*) by including connective type as an additional fixed effect in the linear mixed effect models. The results for sentence reading times, clause fixation times, and connective fixation times remained similar and an additional significant effect of connective type was found. These additional analyses showed that children had longer processing times for the less frequent connective (*doordat*) compared to more frequent connective (*omdat*), which was evidenced for the sentence as whole ($p = .041$) as well as for the two separate clauses (all p's $< .05$) and the connective itself ($p = .003$).

Conclusions and Discussion

Regarding the first research question, the results of the first experiment showed that children's online processing of causal coherence relations was not affected by order of clauses and only minimally by connective presence. For both L1 readers and L2 readers, the presence of a causal connective (*omdat/doordat* 'because') led to longer sentence reading times because of the increased sentence length. No facilitating effects of connective presence were found for the two clauses, which is in contrast to previous studies with adults (Canestrelli et al., 2013, Cozijn et al., 2011) and adolescents (Van Silfhout et al., 2015) that showed that the presence of a connective led to faster processing of subsequent words. This, however, corresponds to our previous study (see Chapter 4) with young L1 and L2 readers using similar stimuli but with a smaller sample size and a lower temporal resolution eye-tracking system, focusing on sentence processing in an oral reading mode rather than a silent reading mode (Van den Bosch et al., 2018).

Consequently, we can rule out some of our previous explanations for not finding a main effect of coherence marking in beginning readers. On the one hand, these results may indicate that no explicit textual support is needed to establish local coherence in these relatively short and simple sentences expressing high predictable causal relations. On the other hand, it may be that children, contrary to adolescents and adults, showed no speed-up effect since they are less experienced readers and their decoding skills are still developing (Schaars, Segers, & Verhoeven, 2019) and show more variation than in older, more experienced readers.

Regarding the second research question, the results revealed that in general the L2 readers had longer processing times compared to the L1 readers, which was the case for sentence reading times as well as clause fixation times, but to a lesser extent for connective fixation times. This suggests that L2 readers had more difficulties in terms of comprehension as these longer processing times could not be attributed to lower decoding skills (i.e., no group differences existed). These difficulties seemed to be related to their lower syntactic knowledge, given that L1-L2 processing differences at the sentence level could completely be explained by individual differences in syntactic knowledge. At the clause level, however, different patterns were observed across clauses; syntactic knowledge was related to fixation times at the first clause, whereas working memory was mainly related to fixation times at the second clause. These findings suggest that syntactic knowledge is required for incremental processing and parsing these complex sentences as a whole and that working memory is involved in integrating information across clauses, which is most likely to take place at the end of the sentence (Millis & Just, 1994) as evidenced by increased reading times at the final clause (Rayner, Kambe & Duffy, 2000).

Experiment 2: Temporal Coherence Relations

The second experiment was conducted to extend the scope of the present study by investigating another type of coherence relations that frequently appears in texts for young readers: temporal coherence relations (e.g., 'Tom ate an apple before he played outside'). In line with Experiment 1, we again investigated the effect of coherence marking and order of clauses on children's online processing in Experiment 2, but some differences existed. First, we could not manipulate the presence of temporal connectives in a similar way as done for causal connectives in Experiment 1. Therefore, in Experiment 2, we investigated the type of connective (i.e., *voordat* 'before' vs. *nadat* 'after) and the order of clauses (i.e., chronological vs. reverse) as done in previous studies on temporal coherence relations (Blything & Cain, 2016; Pyykkönen & Järvikivi, 2012). These two temporal connectives differ in frequency: *nadat* ('after') is less frequent than *voordat* ('before') (Tellings et al., 2014). Second, in addition to examining children's online processing times for sentences, clauses and connectives as done in Experiment 1, we also examined children's scores and response times on comprehension questions in Experiment 2, which was the main focus of previous studies on temporal coherence relations.

The following research questions were addressed:

1. To what extent are L1 and L2 readers' processing and comprehension of temporal relations affected by connective type and order of clauses?
2. To what extent can L1 and L2 readers' variation in processing and comprehension of causal relations be explained from syntactic knowledge and working memory?

We expected that the L2 readers would experience more difficulties with temporal coherence relations than the L1 readers (especially for *nadat* ('after') sentences and reverse order sentences), as evidenced by longer processing times, lower accuracy, and slower responses. Furthermore, it was expected that in particular working memory was related to L1 and L2 readers' online processing and comprehension of temporal coherence relations, as found in previous studies with monolingual children (Blything & Cain, 2016; Karlsson et al., 2019).

Method

Participants. As described in Experiment 1, the participants were 88 third-grade children (8-to-10 years old) of which 47 Dutch monolinguals (L1 readers) and 41 Turkish-Dutch bilinguals (L2 readers).

Measures.

Stimuli temporal coherence relations. Children were presented with 32 two-clause Dutch sentences including a temporal coherence relation We constructed these sentences in a similar way as done by Blything and Cain (2016) for English and Pyykkönen and Järvikiv (2012) for Finnish. We used the Dutch equivalents of before (*voordat*) and after (*nadat*). The sentences were written in past tense and included nine to eleven words (see Appendix C for an overview of all items). In order to minimize the effect of vocabulary knowledge, the sentences only included words that were assumed to be familiar to all third-grade children based on a Dutch wordlist for six-year olds (Schaerleakens et al., 1999) and the described events were commonly occurring and familiar to children this age. The two events depicted in the temporal relations were not causally related (or frequently co-occurring) to prevent inferential reasoning on the basis of world knowledge. Two factors were manipulated: Connective type (i.e., *voordat* vs. *nadat*) and order of clauses (i.e., chronological vs. reverse) resulting in a 2x2 design with four within-subject conditions (see Table 5.4).

Table 5.4 Overview of the Four Within-Subject Conditions for Temporal Coherence Relations

Condition	Sentence with temporal relation	Connective	Order
1	Sara bakte koekjes, voordat ze televisie ging kijken. [Sara baked biscuits, before she watched television.]	voordat [before]	chronological
2	Voordat Sara televisie ging kijken, bakte ze koekjes. [Before Sara watched television, she baked biscuits.]	voordat [before]	reverse
3	Nadat Sara koekjes bakte, ging ze televisie kijken. [After Sara baked biscuits, she watched television.]	nadat [after]	chronological
4	Sara ging televisie kijken, nadat ze koekjes bakte. [Sara watched television, after she baked biscuits.]	nadat [after]	reverse

Each participant was presented with eight items in each condition (32 items in total), presented in pseudo-randomized order. After each sentence, children answered a comprehension question. In line with previous studies, children had to select the event that happened first out of the two alternatives that were described (e.g., 'What did she do first? Baking biscuits or watching television'). In terms of reliability, the internal consistency was good (Cronbach's α = .82). In terms of validity, we checked whether children's comprehension scores for temporal relations correlated with their scores on a standardized test for connective understanding (Verhoeven & Vermeer, 1993). This task consists of two cloze texts with twenty items each, where children have to select the correct connective out of three alternatives

for which they have to understand the meaning of the connectives as well as the relation between clauses/sentences. There was a moderate positive correlation between children's comprehension score for temporal relations and their score for connective understanding ($r = .45$, $p < .001$), which confirms the validity of our experimental task.

Syntactic knowledge. As described in Experiment 1, we used a standardized grammaticality judgment task to assess children's syntactic knowledge.

Working memory. As described in Experiment 1, the Digit Span Backwards was administered to assess children's working memory capacity.

Apparatus. As described in Experiment 1, children's eye movements were recorded with an SMI RED500 eye-tracking system. Again, we used BeGaze analysis software to determine children's total fixation times at three specific regions of the sentence: 1) the first clause, 2) the final clause, and 3) the connective, to illustrate:
[Sara baked biscuits][1] [before][3] [she watched television.][2]

Procedure. As described in Experiment 1, children took part in two individual test sessions and the items of both experiments were mixed into one experimental reading task to increase task variation and to serve as filler items for each other.

Data Analysis. A total of 88 children reading 32 sentences resulted in 2816 observations per outcome variable. Similar to Experiment 1, sentence reading times were analyzed first and then fixation times per region (first clause, final clause, connective) were analyzed. In addition, comprehension scores and response times were examined. Prior to the analyses, we checked all continuous outcome variables for outliers and deviations from normality. All values two standard deviations above or below item *and* subject mean were excluded from the analysis (cf. Van den Bosch et al., 2018). As a result, we excluded 62 observations (2.2%) for sentence reading times, 64 observations (2.3%) for first clause fixation times, 54 observations (1.9%) for final clause fixation times, 49 observations (1.7%) for connective fixation times, and 85 observations (3.0%) for response times. For all variables, the remaining data were within the range of normal distribution; no log-transformations were required.

Linear mixed-effects regression analyses (Baayen et al., 2008) were performed in the R statistics environment (version 3.4.4; R Core Team, 2018) using the lmer() function (for continuous outcome variables) and the glmer() function (for binomial

outcome variable) from the lme4 package (version 1.1-17; Bates et al., 2015). In line with Experiment 1, the analyses were performed in a similar, stepwise way for each outcome variable. To address the first research question, we tested a model with Connective type, Order, Language background, and their interactions as fixed effects. Random effects were fit using a maximal random effects structure (Barry et al., 2013) with random intercepts for participants and items, and by-subject random slopes for Connective type, Order, and their interaction, and a by-item random slope for Language background. Then, regarding the second research question, we added Syntactic knowledge and Working memory (mean-centered) as fixed effects to the model. Models were fit using restricted maximum likelihood technique. A fixed effect was considered significant if the absolute value of the *t*-statistic was larger than or equal to 2.0 (Gelman & Hill, 2007). Since the lmer function does not automatically provide *p*-values, these were derived using Satterthwaite approximations from the lmerTest package (Kuznetsova et al., 2017) which produces acceptable Type 1 error rates even for smaller samples (Luke, 2017).

Results

The results of the linear mixed effects analyses for the processing outcome variables (sentence reading times, clause fixation times and connective fixation times) are shown in Table 5.5 and the results for comprehension outcome variables (accuracy and response times) are reported in Table 5.6.

Sentence reading times. The initial model showed that there was a significant main effect of Connective type, indicating that children took longer to process *after* sentences than *before* sentences. There was also a significant main effect of Language Background; L2 readers had longer sentence processing times than L1 readers. In the second model, where syntactic knowledge and working memory were added as fixed effects, there was again a main effect of Connective type, but the main effect of Language Background was no longer significant. Instead, as in Experiment 1, there was now a significant main effect of syntactic knowledge; lower syntactic knowledge was associated with longer reading times.

Table 5.5 Results Temporal Relations: Sentence Reading Times and Clause Fixation Times

	Model 1			Model 2		
	β	*SE*	*t*	*β*	*SE*	*t*
Sentence reading times						
Connective	541.02	254.51	2.13*	540.93	254.05	2.13*
Order	-3.29	243.62	-0.01	-3.21	243.30	-0.01
LB	1068.01	359.97	2.97**	541.44	435.51	1.24
Connective x Order	-103.74	346.03	-0.30	-103.50	345.56	-0.30
Connective x LB	-66.32	263.12	-0.25	-66.75	262.05	-0.26
Order x LB	254.59	239.57	1.06	254.45	238.73	1.07
Connective x Order x LB	-636.17	341.87	-1.86	-635.94	340.67	-1.87
WM				-88.17	111.34	-0.79
SK				-71.48	31.30	-2.28*
First clause: chronological						
Connective	190.94	102.56	1.86	191.06	102.49	1.86
LB	597.78	166.99	3.58***	454.30	199.03	2.28*
Connective x LB	8.47	130.25	0.07	7.78	130.25	0.06
WM				-108.49	47.77	-2.27*
SK				-22.83	13.82	-1.65
First clause: reversed						
Connective	35.28	120.72	0.29	34.70	120.54	0.29
LB	530.52	150.96	3.51***	359.15	177.33	2.03*
Connective x LB	-213.65	98.25	-2.18*	-213.46	98.20	-2.17*
WM				-100.90	42.20	-2.39*
SK				-25.78	12.19	-2.12*
Final clause: chronological						
Connective	130.89	105.31	1.24	131.43	105.03	1.25
LB	439.49	100.14	4.39***	351.30	121.48	2.89**
Connective x LB	1.97	97.50	0.02	1.81	97.66	0.02
WM				-54.76	29.68	-1.85
SK				-13.43	8.59	-1.56
Final clause: reversed						
Connective	40.18	105.46	0.38	40.48	105.31	0.38
LB	514.71	147.73	3.48***	416.12	171.16	2.43*
Connective x LB	-113.44	109.66	-1.03	-113.72	109.46	-1.04
WM				-55.93	37.25	-1.50
SK				-14.71	10.75	-1.37
Connective fixation times						
Connective	-19.14	33.20	-0.58	-19.49	33.25	-0.59
Order	-61.34	33.52	-1.83	-60.85	33.58	-1.81
LB	171.36	45.67	3.75***	138.06	53.18	2.60*
Connective x Order	92.74	51.03	1.82	92.86	51.14	1.82
Connective x LB	-22.80	43.48	-0.52	-22.29	43.48	-0.51
Order x LB	-19.51	43.68	-0.45	-19.69	43.70	-0.45
Connective x Order x LB	-53.96	67.82	-0.80	-54.15	67.88	-0.80
WM				-23.50		-2.02*
SK				-5.22		-1.55

Note. Connective = Connective type (0 = voordat 'before', 1 = nadat 'after'), Order = Order of clauses (0 = chronological, 2 = reverse), LB = Language background (0 = L1, 1 = L2), WM = Working memory, SK = Syntactic knowledge.
Formula model 1: outcome ~ Connective*Order*LB + (1+Connective*Order | subject) + (1| item); model 2: outcome ~ Connective*Order*LB + WM + SK + (1+Connective*Order | subject) + (1| item).
*p < .05, **p < .01, ***p < .001.

Clause fixation times. Next, we investigated children's fixation times for the two clauses separately to gain more detailed insight into their online processing. Due to our manipulation of order of clauses, the first and final clause are not comparable in terms of content and length across the four conditions. Therefore, the analyses were split-up by order so that separate models were tested for chronological sentences and reverse sentences.

Fixation times first clause. For chronological sentences, only a significant main effect of Language background was found in the initial model, indicating that L2 readers had longer fixation times for the first clause than L1 readers. In the second model, the main effect of Language background remained significant and in addition there was a significant main effect of working memory, indicating that children with lower working memory had longer fixation times on the first clause. For reverse sentences, a significant main effect of Language background and a significant Connective type x Language background interaction effect were found in the initial model, indicating that the differences in first clause fixation times between *before* sentences and *after* sentences were larger for L2 readers compared to L1 readers. In the second model these effects remained significant and additional significant main effects were found for working memory and syntactic knowledge, indicating that children with lower working memory and children with lower syntactic knowledge had longer fixation times at the first clause.

Fixation times final clause. For chronological sentences, only a significant main effect of Language Background was found in the initial model as well as in the second model, indicating that L2 readers had longer fixation times at the final clause than L1 readers. The results were similar for reverse sentences.

Connective fixation times. In the initial model, there were no significant effects for Connective type and Order, but there was a significant main effect of Language background, indicating that the L2 readers fixated longer on the connective than L1 readers. In the second model, where syntactic knowledge and working memory were added as fixed effects, the main effect of Language background remained significant and in addition there was a significant main effect of working memory, indicating that children with lower working memory had longer fixation times at the connective.

Table 5.6 Results Temporal Relations: Comprehension Scores and Response Times

	Model 1			Model 2		
	β	*SE*	*z/t*	*β*	*SE*	*z/t*
Comprehension scores						
Connective	-0.10	0.42	-0.24	-0.19	0.44	-0.45
Order	0.08	0.42	0.20	0.03	0.44	0.06
LB	-0.59	0.31	-1.88	-0.31	0.38	-0.82
Connective x Order	-0.87	0.53	-1.66	-0.87	0.54	-1.61
Connective x LB	-0.86	0.43	-1.98*	-0.83	0.45	-1.87
Order x LB	-0.25	0.43	-0.58	-0.23	0.44	-0.53
Connective x Order x LB	0.82	0.46	1.78	0.83	0.48	1.72
WM				0.15	0.08	1.90
SK				0.05	0.02	2.01*
Response times						
Connective type	269.73	112.16	2.41*	269.47	118.46	2.28*
Order	3.07	109.76	0.03	3.56	116.61	0.03
LB	120.71	169.32	0.71	206.88	193.77	1.07
Connective x Order	-98.96	167.49	-0.59	-99.15	176.16	-0.56
Connective x LB	88.54	172.67	0.51	90.09	155.87	0.58
Order x LB	285.48	169.06	1.69	285.95	152.55	1.87
Connective x Order x LB	-393.34	256.31	-1.54	-394.08	234.14	-1.68
WM				10.42	46.93	0.22
SK				11.53	13.19	0.87

Note. Connective = Connective type (0 = voordat 'before', 1 = nadat 'after'), Order = Order of clauses (0 = chronological, 2 = reverse), LB = Language background (0 = L1, 1 = L2), WM = Working memory, SK = Syntactic knowledge.
Formula model 1: outcome ~ Connective*Order*LB + (1+Connective*Order | subject) + (1| item); model 2: outcome ~ Connective*Order*LB + WM + SK + (1+Connective*Order | subject) + (1| item).
*p < .05.

Comprehension scores. The initial model showed that there were no significant main effects of Connective type, Order, Language background, but a significant Connective type x Language background interaction effect existed. As can be concluded from Figure 5.1, L1 readers scored on average relatively high on comprehension for both connectives (*before*: 90% correct, *after*: 80% correct), whereas L2 readers scored on average particularly lower on *after* sentences (65% correct). Additional analyses indicated that the L2 readers scored significantly lower on *after* sentences than *before* sentences (p = .018), whereas this was not the case for the L1 readers (p = .919). Also, L2 readers scored significantly lower than L1 readers for *after* sentences (p < .001), but not for *before* sentences (p = .068). In the second model, where syntactic knowledge and working memory were added as fixed effects, the main effect of Language Background was not significant anymore and the Connective type x Language background interaction effect was only

marginally significant now. Instead, there was a significant main effect of syntactic knowledge, indicating children with higher syntactic knowledge scored higher on the comprehension questions.

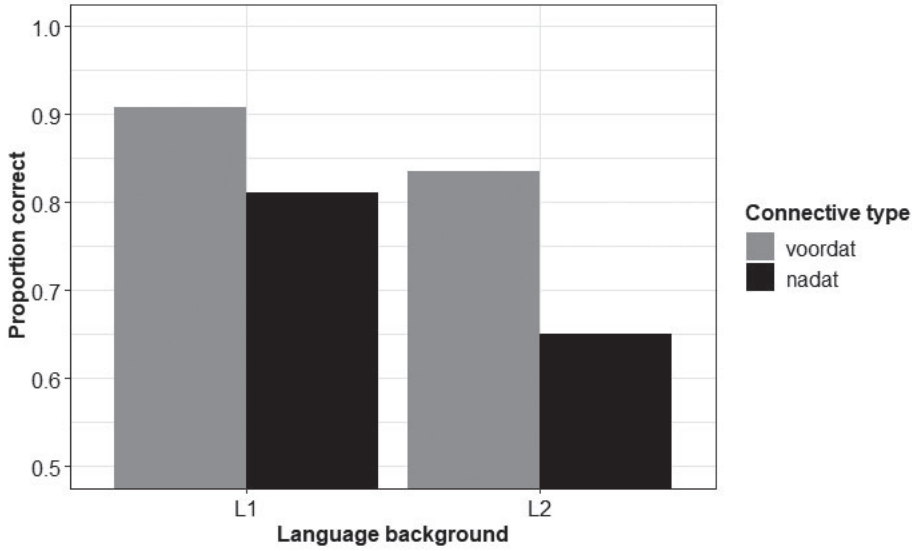

Figure 5.1 Comprehension score (i.e., mean proportion correct across items) for *voordat* ('before') and *nadat* ('after') sentences split by language background (L1 readers vs. L2 readers).

Response times. The initial model showed that there was only a significant main effect of Connective type, indicating that response times were longer for *after* sentences compared to *before* sentences. In the second model, where syntactic knowledge and working memory were added as fixed effects, similar results were found.

Given that children's comprehension of temporal relations was not at ceiling, we conducted additional analyses to investigate whether processing differences existed between correct and incorrect items. When we included Correctness (item correct vs. item incorrect) as an additional fixed effect in the linear mixed effect models (see also Vorstius et al., 2013), the results for sentence reading times, clause fixation times, connective fixation times and response times remained similar. Only for connective fixation times and response times a significant effect of Correctness was found, indicating that children's fixation times on the connective were longer for correct items (M1: $\beta = 0.08$, $SE = 0.04$, $t = 2.36$, $p = .018$; M2: $\beta = 0.08$, $SE = 0.04$,

t = 2.47, p = .013) and their response times for correct items were shorter (M1: β = -445.32, SE = 75.39, t = -5.91, p < .001; M2: β = -449.11, SE = 75.37, t = -5.96, p < .001) compared to incorrect items. The fact that longer fixation times on the connective were found for correct items suggests that children paid attention to connectives during reading to establish the correct order of events.

Conclusions and Discussion

Regarding the first research question, the results of the second experiment showed that children's online processing and comprehension of temporal coherence relations was affected by the type of connective. For both L1 readers and L2 readers, sentence reading times were longer for *after* sentences compared to *before* sentences, but no effect was found at particular regions of the sentences. Moreover, children's response times were longer for *after* sentences compared to *before* sentences. In addition, for L2 readers (but not L1 readers) a difference between connectives was also observed for comprehension; L2 readers were less accurate on *after* sentences than *before* sentences. Furthermore, L2 readers' proportion correct answers for *after* sentences (65%) was not far from chance level (50%) and considerably lower than L1 readers (81%). These findings are in line with our expectations and the notion that *after* is linguistically more complex and less frequent than *before* and thus more difficult to process and comprehend (Clark, 1971).

Against expectations, the other text-related factor of interest, order of clauses, had no impact on children's online processing or comprehension contrary to previous studies that showed that reverse sentences are more difficult than chronological sentences (Blything & Cain, 2016; Clark, 1971). This discrepancy may either relate to fact that we included children from diverse linguistic backgrounds rather than monolingual children only, or to the specific language of the study. Karlsson et al. (2019), who also investigated temporal relations in Dutch, neither found an effect of order, whereas they did find an effect of the position of the connective (sentence initially vs. sentence medially) on children's processing and comprehension. Although we did not explicitly manipulate the position of the connective, this varied as a result of manipulating the other two textual factors. Additional analyses, however, did not yield any significant effects of connective position in the present study.

Regarding the second research question, the results showed that, as expected, the L2 readers had longer processing times than the L1 readers, which was found

for sentence reading times as well as clause fixation times and connective fixation times. Similar to Experiment 1, group differences in sentence reading times (but also comprehension scores) could largely be explained by individual differences in syntactic knowledge. At the clause level, however, working memory was related to children's fixation times at the first clause and the connective, but not the second clause. These findings suggest that syntactic knowledge is required for parsing these complex sentences as a whole and that lower working memory capacity probably leads to longer (or more fixations/re-reading) on the initial clause and the connective itself in order to establish the coherence relation between clauses.

General Discussion

In the present study, two experiments were conducted to investigate online processing of coherence relations in beginning first language (L1) and second language (L2) readers. We focused on the processing of two types of semantic relations: causal coherence relations (Experiment 1) and temporal coherence relations (Experiment 2). The two experiments examined children's online processing of coherence relations at the sentence level as well as the clause level and investigated the role of both text-related factors (connective presence/type and order of clauses) and child-related factors (language background, syntactic knowledge, working memory). Together, these experiments aimed to shed light on how beginning readers establish local coherence during reading and on the factors that explain individual variation in children's online comprehension processes.

With respect to textual factors, the two experiments together showed that not the order of clauses or the presence (or position) of a connective, but the type of connective primarily affected beginning readers' online processing of coherence relations. For both causal and temporal coherence relations, children had more processing difficulties (indicated by longer processing times) with a less frequent connective (*doordat*, *nadat*) compared to a more frequent connective (*omdat*, *voordat*), as can be explained by the difference in familiarity and/or linguistic complexity of these connectives. Also, this finding relates to word frequency effects evidenced in eye-movement studies with children, showing that less frequent words have longer reading times (Joseph, Nation, & Liversedge, 2013). Importantly, this finding indicates that beginning readers actually paid attention to these coherence markers during reading, which could not be concluded yet from previous research using offline measures of reading comprehension to investigate the effect of connectives (Crosson et al., 2008; Crosson & Lesaux, 2013; Kohnen

& Retelsdorf, 2019). Additional evidence for children's attention to and usage of connectives during reading comes from the findings that the children had longer sentence reading times when a connective was present (Experiment 1) and that the children fixated longer on the connectives for correct items (Experiment 2). It can be concluded that although beginning readers' knowledge of connectives is still developing, eye-tracking technology revealed that connectives affected their online processing to some extent. Yet, no immediate speed-up effects could be evidenced in line with previous studies with adolescents and adults (Cozijn et al., 2011; Van Silfhout et al., 2015).

With respect to child-related factors, the two experiments demonstrated that processing differences existed between beginning L1 readers and L2 readers, even though the two groups were comparable in terms of decoding skills. Based on their longer processing times in general, it can be concluded that L2 readers experienced more difficulties with processing causal and temporal coherence relations, which may relate a delay in oral language and reading comprehension. Apart from children's language background, also individual differences in syntactic knowledge and working memory affected their online processing. Interestingly, syntactic knowledge largely explained these group differences at the sentence level, whereas working memory was mainly related to children's online processing at the clause level. This latter finding adds on previous studies on temporal relations showing that working memory capacity was related to monolingual children's sentence reading times (Karlsson et al., 2019) and their comprehension scores and response times (Blything & Cain, 2016) and are in accordance with a memory-capacity constrained framework (Just & Carpenter, 1992).

The present study has some limitations that should be mentioned. First, we investigated children's processing of coherence relations by presenting isolated sentences in an experimental setting. To arrive at a better understanding of the processing of coherence relations in text it is important to conduct follow-up studies using reading tasks that include paragraphs or short texts to investigate the effect of connectives in a more natural way. Moreover, other ways of manipulating coherence marking could be applied in future studies focusing on beginning readers from diverse linguistic backgrounds, such as using appropriate vs. inappropriate connectives (Cain & Nash, 2011) or highlighting connectives (Geva & Ryan, 1985). Second, we focused on children's total fixation times as an eye-tracking measure as a first attempt to investigate children's online processing in more detail, but it is recommended to use a larger variety of processing measures in future studies (see

e.g., De Leeuw, Segers, & Verhoeven, 2016), such as first pass/second pass duration and the number of regression to elaborate on the findings of the present study.

To conclude, this is one of the first studies that used eye-tracking to provide more detailed insight into how beginning readers from diverse linguistic backgrounds process coherence relations and deal with coherence markers, such as connectives, during reading. Building on previous studies that mainly focused on offline comprehension and/or more experienced readers, the present study demonstrated that both textual and child-related factors affected beginning readers' online processing of causal and temporal coherence relations. Furthermore, by focusing on beginning readers from diverse linguistic backgrounds, the present study revealed online processing differences between L1 readers and L2 readers and highlighted the role of individual differences in syntactic knowledge and working memory. Together, these findings emphasize the importance of paying attention to textual demands as well as children's linguistic and cognitive skills to support their comprehension processes during reading.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412. doi:10.1016/j.jml.2007.12.005

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48. doi:10.18637/jss.v067.i01

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish-Dutch children. *Journal of Experimental Child Psychology*, *128*, 105-119. doi:10.1016/j.jecp.2014.06.007

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye-movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.). *The Oxford Handbook of Eye-movements* (pp. 643–662). Oxford, UK: Oxford University Press.

Blything, L. P., & Cain, K. (2016). Children's processing and comprehension of complex sentences containing temporal connectives: The influence of memory on time course of accurate responses. *Developmental Psychology*, *52*, 1517-1529. doi:10.1037/dev0000201

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*, 429-441. doi:10.1037/a0022824

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31-42. doi:10.1037/0022-0663.96.1.31

Cain., K., Patson, N., & Andrews., L. (2005). Age- and ability-related differences in young readers' use of conjunctions. *Journal of Child Language*, *32*, 877-892. doi:10.1017/S0305000905007014

Canestrelli, A. R., Mak, W. M., & Sanders, T. J. M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, *28*, 1394-1413. doi:10.1080/01690965.2012.685885

CBS Statistics the Netherlands (2016). Bevolking naar migratieachtergrond [Population By immigration background]. Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Cito B.V. (2014). *Begrijpend lezen 3.0 voor groep 5: M5-toets [Reading Comprehension 3.0 for third grade: M5 test]*. Arnhem, the Netherlands: Cito B.V.

Clark, E. V. (1971). On the acquisition of the meaning of before and after. *Journal of Verbal Learning and Verbal Behavior*, *10*, 266-275. doi:10.1016/S0022-5371(71)80054-3

Cozijn, R., Noordman, L. G. M., & Vonk, W. (2011). Propositional integration and world knowledge inference: Processes in understanding because sentences. *Discourse Processes*, 48, 475-500. doi:10.1080/0163853X.2011.594421

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading*, *36*, 241-260. doi:10.1111/j.1467-9817.2011.01501.x

Crosson, A. C., Lesaux, N. K., & Martiniello, M. (2008). Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds. *Applied Psycholinguistics*, *29*, 603-625. doi:10.1017/S0142716408080260

De Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text related characteristics on student's reading behaviour and text comprehension: An eye movement study. *Scientific Studies of Reading*, *20*, 248-263. doi:10.1080/10888438.2016.1146285

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing: An Interdisciplinary Journal*, 15, 739-757. doi:10.1023/A:1020932715838

Diessel, H. (2004). *The acquisition of complex sentences in English*. Cambridge, UK: Cambridge University Press.

Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives; How cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language*, *36*, 829-854. doi:10.1017/S0305000908009227

Extra, G. (2002). *De andere talen van Nederland: Thuis en op school. [The other languages of the Netherlands: At home and at school]*. Bussum, the Netherlands: Coutinho.

Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*, 389-421. doi:10.1111/jrir.12003

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Geva, E., & Ryan, E. B. (1985). Use of conjunctions in expository texts by skilled and less skilled readers. *Journal of Literacy Research*, *17*, 331-346. doi:10.1080/10862968509547549

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*, 325-340. doi:10.1177/0267658316678286

Halliday, M. E. K., & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman.

Joseph, H. S. S. L., Nation, K., & Liversedge, S. (2013). Using eyemovements to investigate word frequency effects in children's sentence reading. *School Psychology Review*, *42*, 207-222.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122-149. doi:10.1037/0033-295X.99.1.122

Karlsson, J., Jolles, D., Koornneef, A., Van den Broek, P., & Van Leijenhorst, L. (2019). Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory updating. *Journal of Experimental Child Psychology*, *185*, 1-18. doi:10.1016/j.jecp.2019.04.007

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes*, *39*, 125-128. doi:10.1080/0163853X.2005.9651676

Kohnen, N., & Retelsdorf, J. (2019). The role of knowledge of connectives in comprehension of a German narrative text. *Journal of Research in Reading*, *42*, 371-388. doi:10.1111/1467-9817.12273

Kort, W., Schittekatte, M., Dekker, P. H., Verhaeghe, P., Compaan, E. L., Bosmans, M., & Vermeir, G. (2005). *WISC-III NL. Handleiding en Verantwoording* [Manual and Justification]. London: The Psychological Corporation.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1-26. doi:10.18637/jss.v082.i13

Lipka, O., & Siegel, L. S. (2007). The development of reading skills in children with English as a second language. *Scientific Studies of Reading*, *11*, 105-13. doi:10.1080/10888430709336555

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*, 1494-1502. doi:10.3758/s13428-016-0809-y

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409-433. doi:10.1037/a0033890

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language*, *33*, 128-147. doi:10.1006/jmla.1994.1007

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundation of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*, 665– 681. doi:10.1037/0012-1649.40.5.665

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22-37. doi:10.1080/10888438.2013.827687

Perfetti, C., Yang, C., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*, 303-318. doi:10.1002/acp.1419

Pyykkönen, P., & Järvikivi, J. (2012). Children and situation models of multiple events. *Developmental Psychology*, *48*, *2*, 521-529. doi:10.1037/a0025526

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rayner, K., Kambe, G., & Duffy, S. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology*, *53A*, 1061-1079. doi:10.1080/713755934

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1-35. doi:10.1080/01638539209544800

Sanders, T., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*, 37-60. doi:10.1207/S15326950dp2901_3

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2019). Cognitive and linguistic precursors of early first and second language reading development. *Learning and Individual Differences*, *72*, 1-14. doi:10.1016/j.lindif.2019.03.008

Schaerleakens, A., Kohnstamm, D., Lejaegere, M. (1999). *Streeflijst woordenschat voor zesjarigen* [Vocabulary list for six-year-olds]. Lisse: Swets en Zeitlinger.

Tellings, A., Hulsbosch, M., Vermeer, A., & Van den Bosch, A. (2014). BasiLex: an 11.5 million words corpus of Dutch texts written for children. *Computational Linguistics in the Netherlands Journal*, *4*, 191-208.

Tellings A., & Penning-de Vries, B. (2019). Connective frequency in child-directed texts: Dutch corpus analysis. Manuscript submitted for publication.

Van den Bosch, L. J., Segers, E., & Verhoeven, L. (2018). Online processing of causal relations in beginning first and second language readers. *Learning and Individual Differences*, *61*, 59-67. doi:10.1016/j.lindif.2017.11.007

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, 52, 47-76. doi:10.1080/0163853X.2014.905237

Verhoeven, L., & Keuning, J. (2017). The nature of developmental dyslexia in a transparent orthography. *Scientific Studies of Reading*, 1-17. doi:10.1080/10888438.2017.1317780

Verhoeven, L.T.W., Keuning, J., Horsels, L., Van Boxtel, H. (2013). *Testinstrumentarium Taalontwikkelingsstoornissen* [Set of Instruments Language Development Disorder]. Arnhem, the Netherlands: Cito.

Verhoeven, L., & Vermeer, A. (1993). *Taaltoets Allochtone Kinderen Bovenbouw* [Language Test Immigrant Children Upper Primary Grades]. Tilburg, the Netherlands: Zwijsen.

Vorstius, C., Radach, R., Mayer, M. B., & Lonigan, C. J. (2013). Monitoring local comprehension monitoring in sentence reading. *School Psychology Review*, *32*, 191-206.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, 1196-1207.

Appendix B

1A. Omdat Daan honger had, at hij een wafel.
[Because Daan was hungry, he ate a waffle.]
1B. Luuk at een appel, omdat hij honger had.
[Luuk ate an apple, because he was hungry.]
1C. Joep had honger. Hij at een tosti.
[Joep was hungry. He ate a sandwich.]
1D. Koen at een pizza. Hij had honger.
[Koen ate a pizza. He was hungry.]
Comprehension question: Wat at hij? [What did he eat?]

2A. Omdat Anna dorst had, dronk ze een glas cola.
[Because Anna was thirsty, she drank a glass of coke.]
2B. Anne dronk een glas water, omdat ze dorst had.
[Anne drank a glass of water, because she was thirsty.]
2C. Tess had dorst. Ze dronk een glas melk.
[Tess was thirsty. She drank a glass of milk.]
2D. Roos dronk een glas thee. Ze had dorst.
[Roos drank a glass of tea. She was thirsty.]
Comprehension question: Wat dronk ze? [What did she drink?]

3A. Omdat Jens jarig was, kreeg hij een fiets van zijn ouders.
[Because it was Jens birthday, he got a bike from his parents.]
3B. Lars kreeg een trein van zijn ouders, omdat hij jarig was.
[Lars got a train from his parents, because it was his birthday.]
3C. Bram was jarig. Hij kreeg een piano van zijn ouders.
[It was Bram's birthday. He got a piano from his parents.]
3D. Bart kreeg een gitaar van zijn ouders. Hij was jarig.
[Nick got a guitar from his parents. It was his birthday.]
Comprehension question: Wat kreeg hij? [What did he get?]

4A. Omdat Lisa gevallen was, had ze veel pijn aan haar neus.
[Because Sara had fallen, she hurt her nose.]
4B. Nina had veel pijn aan haar been, omdat ze gevallen was.
[Lara hurt her leg, because she had fallen.]
4C. Sara was gevallen. Ze had veel pijn aan haar pols.
[Anja had fallen. She hurt her wrist.]
4D. Vera had veel pijn aan haar knie. Ze was gevallen.
[Lisa hurt her knee. She had fallen.]
Comprehension question: Wat deed pijn? [What did she hurt?]

5A. Doordat Sam ziek was, ging hij niet naar school.
[Because Sam was ill, he did not go to school.]
5B. Sem ging niet naar school, doordat hij ziek was.
[Sem did not go to school, because he was ill.]
5C. Tom was ziek. Hij ging niet naar school.
[Tom was ill. He did not go to school.]
5D. Tim ging niet naar school. Hij was ziek.
[Tim did not go to school. He was ill.]
Comprehension question: Wie ging niet naar school? [Who did not go to school?]

6A. Doordat Lara een lekke fietsband had, moest ze met de bus.
[Because Lara got a flat, she had to take the bus.]
6B. Mila moest met de bus, doordat ze een lekke fietsband had.
[Mila had to take the bus, because she got a flat.]
6C. Nora had een lekke fietsband. Ze moest met de bus.
[Nora got a flat. She had to take the bus.]
6D. Lina moest met de bus. Ze had een lekke fietsband.
[Lina had to take the bus. She got a flat.]
Comprehension question: Wie moest met de bus? [Who had to take the bus?]

7A. Doordat Rick niet geleerd had, haalde hij een slecht cijfer.
[Because Rick did not learn, he got a low grade.]
7B. Nick haalde een slecht cijfer, doordat hij niet geleerd had.
[Nick got a low grade, because he did not learn for the test.]
7C. Siem had niet geleerd. Hij haalde een slecht cijfer.
[Siem did not learn. He got a low grade.]
7D. Teun haalde een slecht cijfer. Hij had niet geleerd. [Teun got a low grade. He did not learn.]
Comprehension question: Wie haalde een slecht cijfer? [Who got a low grade?]

8A. Doordat Eva zich verslapen had, kwam ze te laat op school.
[Because Eva overslept, she was late at school].
8B. Evi kwam te laat op school, doordat ze zich verslapen had.
[Evi was late at school, because she overslept].
8C. Isa had zich verslapen. Ze kwam te laat op school.
[Isa overslept. She was late at school.]
8D. Kim kwam te laat op school. Ze had zich verslapen.
[Kim was late at school. She overslept.]
Comprehension question: Wie kwam te laat? [Who was too late?]

Note. Each item consisted of two clauses that were separated by either a comma or full stop. In A-items: connective = present, order = cause-effect. In B-items: connective = present, order = effect-cause. In C-items: connective = absent, order = cause-effect. In D-items: connective = absent, order = effect-cause.
Frequent Dutch names were used and within each block they were matched on gender and length.

Appendix C

9A. Luuk at een boterham, voordat hij met de hond ging wandelen.
[Luuk ate a sandwich before he took the dog for a walk.]
9B. Voordat Daan met de hond ging wandelen, at hij een boterham.
[Before Daan took the dog for a walk, he ate a sandwich.]
9C. Nadat Koen een boterham at, ging hij met de hond wandelen.
[After Koen at a sandwich, he took the dog for a walk.]
9D. Joep ging met de hond wandelen, nadat hij een boterham at.
[Joep took the dog for a walk after he ate a sandwich.]
Comprehension question: Wat deed hij eerst? [What did he do first?]

10A. Bart dronk een kopje thee, voordat hij een boek ging lezen.
[Bart drank a cup of tea before he read a book.]
10B. Voordat Bram een boek ging lezen, dronk hij een kopje thee.
[Before Bram read a book, he drank a cup of tea.]
10C. Nadat Lars een kopje thee dronk, ging hij een boek lezen.
[After Lars drank a cup of tea, he read a book.]
10D. Jens ging een boek lezen, nadat hij een kopje thee dronk.
[Jens read a book, after he drank a cup of tea.]
Comprehension question: Wat deed hij eerst? [What did he do first?]

11A. Roos kamde haar haren, voordat ze haar tanden ging poetsen.
[Roos combed her hair, before she brushed her teeth.]
11B. Voordat Tess haar tanden ging poetsen, kamde ze haar haren.
[Before Tess brushed her teeth, she combed her hair.]
11C. Nadat Anne haar haren kamde, ging ze haar tanden poetsen.
[After Anne combed her hair, she brushed her teeth.]
11D. Anna ging haar tanden poetsen, nadat ze haar haren kamde.
[Anna brushed her teeth, after she combed her hair.]
Comprehension question: Wat deed ze eerst? [What did she do first?]

12A. Nina maakte een tekening, voordat ze naar school ging.
[Nina made a drawing before she went to school.]
12B. Voordat Lisa naar school ging, maakte ze een tekening.
[Before Lisa went to school, she made a drawing.]
12C. Nadat Vera een tekening maakte, ging ze naar school.
[After Vera made a drawing, she went to school.]
12D. Sara ging naar school, nadat ze een tekening maakte.
[Sara went to school after she made a drawing.]
Comprehension question: Wat deed ze eerst? [What did she do first?]

13A. Lina maakte haar huiswerk, voordat ze buiten ging spelen.
[Lina did her homework, before she played outside.]
13B. Voordat Nora buiten ging spelen, maakte ze haar huiswerk.
[Before Nora played outside, she did her homework.]
13C. Nadat Mila haar huiswerk maakte, ging ze buiten spelen.
[After Mila did her homework, she played outside.]
13D. Lara ging buiten spelen, nadat ze haar huiswerk maakte.
[Lara played outside, after she did her homework.]
Comprehension question: Wat deed ze eerst? [What did she do first?]

14A. Evi bakte lekkere koekjes, voordat ze televisie ging kijken.
[Evi baked nice biscuits, before she watched television.]
14B. Voordat Eva televisie ging kijken, bakte ze lekkere koekjes.
[Before Eva watched television, she baked nice biscuits.]
14C. Nadat Kim lekkere koekjes bakte, ging ze televisie kijken.
[After Kim baked nice biscuits, she watched television.]
14D. Isa ging televisie kijken, nadat ze lekkere koekjes bakte.
[Isa watched television after she baked nice biscuits.]
Comprehension question: Wat deed ze eerst? [What did she do first?]

15A. Sem belde met zijn oma, voordat hij een hut ging bouwen.
[Sem called his grandma before he built a hut.]
15B. Voordat Sam een hut ging bouwen, belde hij met zijn oma.
[Before Sam built a hut, he called his grandma.]
15C. Nadat Tim met zijn oma belde, ging hij een hut bouwen.
[After Tim called his grandma, he built a hut.]
15D. Tom ging een hut bouwen, nadat hij met zijn oma belde.
[Tom built a hut after he called his grandma.]
Comprehension question: Wat deed hij eerst? [What did he do first?]

16A. Teun dronk een glas cola, voordat hij ging zwemmen.
[Teun drank a glass of coke, before he went to the pool.]
16B. Voordat Siem ging zwemmen, dronk hij een glas cola.
[Before Siem went to the pool, he drank a glass of coke.]
16C. Nadat Nick een glas cola dronk, ging hij zwemmen.
[After Nick drank a glass of coke, he went to the pool.]
16D. Rick ging zwemmen, nadat hij een glas cola dronk.
[Rick went to the pool after he drank a glass of coke.]
Comprehension question: Wat deed hij eerst? [What did he do first?]

Note. Each item consisted of two clauses that were separated by either a comma or full stop. In A-items: connective = voordat, order = chronological. In B-items: connective = voordat, order = reversed. In C-items: connective = nadat, order = chronological. In D-items: connective = nadat, order = reversed.
Frequent Dutch names were used and within each block they were matched on gender and number of syllables.

Chapter 6

General discussion

Reading is a crucial skill that children acquire at primary school, but comprehending written text is a complex process. Reading comprehension can be particularly difficult for children who learn to read in their second language (L2 readers). Yet, little is known about the early stages of reading comprehension development and the process of reading comprehension in beginning readers from diverse linguistic backgrounds. The present dissertation investigated the role of linguistic diversity in early reading comprehension. The general aim of this dissertation was to examine reading comprehension performance, development and online processes in beginning L1 readers and L2 readers and to explain between-group variation as well as individual variation. The first part of this dissertation (Chapters 2 and 3) contained a cross-sectional study and a longitudinal follow-up study on early reading comprehension performance and its development over time. The second part of this dissertation (Chapters 4 and 5) presented two experimental studies on online processing of coherence relations. Based on the three research questions that were central to the present dissertation (see Chapter 1), this final chapter provides a summary of the main findings. Furthermore, theoretical implications are described followed by limitations and recommendations for future research, and implications for educational practice.

Early Reading Comprehension: Variation between L1 readers and L2 readers

With regard to the first research question, a comparison was made between L1 readers and L2 readers in terms of early reading comprehension performance (Chapter 2), development (Chapter 3), and online processes (Chapters 4 and 5). Children's reading comprehension performance was first assessed mid second grade and then follow-up measures were administered at the end of second grade and mid third grade. The cross-sectional study presented in Chapter 2 showed that halfway second grade, when children have moved from reading isolated words to comprehending sentences and short texts, L2 readers performed on average considerably lower on reading comprehension than L1 readers. The longitudinal study described in Chapter 3 showed that this L1-L2 performance gap in reading comprehension was also observed at the end of second grade and mid third grade, whereas the two groups were comparable in terms of their decoding skills. More specifically, L2 readers' performance in third grade was comparable to L1 readers' performance in second grade, indicating an average gap (or delay) of one year.

These findings are in line with the meta-analysis on reading comprehension in L1 and L2 readers conducted by Melby-Lervåg and Lervåg (2014) and demonstrate that the L1-L2 performance gap that was found in the intermediate (Droop &

Verhoeven, 2003; Verhoeven, 2000) and higher primary grades (Farnia & Geva, 2013) already exists in lower primary grades. Yet, the longitudinal study in Chapter 3 also revealed that the L1-L2 performance gap in reading comprehension decreased over time, in contrast to previous studies that showed no change or an increased gap in the intermediate/upper primary grades (e.g., Farnia & Geva, 2013). In the present dissertation, the L2 readers showed larger growth in reading comprehension than the L1 readers, particularly in third grade where reading instruction is focused more explicitly on comprehension skills. It may be that L2 readers are able to catch up with their monolingual peers initially as most texts in the lower primary grades are narrative texts and require only a very basic level of vocabulary (Gottardo & Mueller, 2009). In the higher primary grades, children are confronted with more complex, expository texts including academic vocabulary (Snow, 2010) which may be particularly challenging to L2 readers and may result in larger differences (i.e., a larger performance gap) with their monolingual peers again.

In addition to differences in reading comprehension outcomes, the experimental studies described in Chapters 4 and 5 demonstrated that also some differences in online reading comprehension processes existed between L1 readers and L2 readers. The first experimental study (Chapter 4) showed that the absence of a causal connective (*omdat* 'because') led to longer sentence processing times for L2 readers with lower syntactic knowledge, whereas this was not found for L1 readers. Apparently, L2 readers with limited syntactic knowledge benefitted from coherence marking provided by a connective, whereas L1 readers did not need explicit textual support to establish coherence during reading. Furthermore, the second experimental study (Chapter 5) showed that L2 readers had longer processing times compared to the L1 readers, both at the sentence level and the clause level, despite accurate decoding skills. This indicates that L2 readers experienced more difficulties with processing causal and temporal coherence relations in general, which may be related to their delay in oral language skills and reading comprehension performance (Melby-Lervåg & Lervåg, 2014).

Individual Variation in Early Reading Comprehension Outcomes

With regard to the second research question, the present dissertation investigated how variation in early reading comprehension performance and development among L1 readers and L2 readers could be explained. To explain between-group variation, the study described in Chapter 2 examined several linguistic and cognitive factors that were determined on the basis of prominent reading

comprehension theories. The results showed that the L2 readers scored lower on vocabulary knowledge and morphosyntactic knowledge than the L1 readers on average, whereas the two groups were comparable in terms of decoding skills and working memory capacity. Importantly, the classical regression analysis demonstrated that the key predictors were similar across groups (Lesaux, Lipka, & Siegel, 2006; Verhoeven & Van Leeuwe, 2012); decoding skills, vocabulary knowledge, and morphosyntactic knowledge explained unique variance in early reading comprehension performance for both L1 readers and L2 readers. By contrast, children's language background (L1 vs. L2) did not explain unique variance, indicating that individual differences in decoding and oral language skills were sufficient to explain the L1–L2 performance gap, which is in line with the study of Babayiğit (2014). Moreover, considerable variation within groups existed in terms of early reading comprehension performance. As can be concluded from Figure 2.1, L1 readers and L2 readers were both represented in the lower and higher ends of the distribution. Consequently, it can be argued that taking into account individual differences among children is more relevant than categorizing them based on their language background (L1 readers vs. L2 readers). As an attempt to move away from this categorization, the present dissertation explored other ways to explain individual variation in reading comprehension that move beyond the binary view on linguistic diversity.

First, a distinction was made based on children's reading comprehension ability level instead of their language background (Chapter 2), assuming that such a distinction would do more justice to individual variation among children. By means of quantile regression analysis, the relation between linguistic/cognitive factors and reading comprehension performance was investigated at multiple points (quantiles) across the distribution that are indicative of different ability levels (poor vs. average vs. good). Interestingly, the results showed that the key predictors of early reading comprehension performance differed across ability levels; decoding skills and morphosyntactic knowledge were unique predictors at all ability levels, but vocabulary knowledge was uniquely related for poor comprehenders only and working memory for good comprehenders only. These findings are in line with recent studies using quantile regression to show that differential relations exist across the distribution of reading comprehension scores in older, monolingual readers (Hua & Keenan, 2017; LRRC & Logan, 2017; Tighe & Schatschneider, 2016). By extending the use of quantile regression to a linguistically diverse group of beginning readers, this is the first study demonstrating that the explanatory factors of early reading comprehension performance are more likely to vary as a

function of children's ability level than their language background. In addition, these findings again provide support for the idea that not children's language background but their actual skills are important to take into account.

Second, subgroups of L2 readers where distinguished based on their vocabulary knowledge in both languages (Chapter 3), presuming that bilinguals should not be considered as one and the same group (Luk & Bialystok, 2013) and that both languages need to be taken into account to provide a full picture of their proficiency (Bialystok, Luk, Peets, & Yang, 2010). By means of cluster analysis, four subgroups of L2 readers with varying levels of L1 (Turkish) and L2 (Dutch) vocabulary knowledge were identified. These four subgroups corresponded to the four types of bilingualism distinguished by Cummins (1979). Of interest, the performance gap in reading comprehension between L1 readers and L2 readers differed across these L2 subgroups. The overall performance gap was smaller for the two L2 subgroups with higher L2 vocabulary than for the two L2 subgroups with lower L2 vocabulary. Furthermore, the performance gap decreased over time for the two L2 subgroups with higher L1 vocabulary; these subgroups showed an increase in growth as compared to the L1 readers. Remarkably, the L2 subgroup with higher vocabulary in both their languages even closed the gap with L1 readers halfway third grade. Together, these findings stress the importance of considering individual variation among bilingual children instead of categorizing them as one and the same group and reveal that both L1 and L2 vocabulary have an impact on their early reading comprehension development.

Individual Variation in Early Reading Comprehension Processes

Apart from reading comprehension outcomes, the present dissertation also examined online reading comprehension processes to gain more detailed insights into how beginning readers from diverse linguistic backgrounds construct meaning during reading. With regard to the third research question, the experimental studies reported in Chapters 4 and 5 investigated to what extent textual and child-related factors explained variation among L1 readers and L2 readers in online processing of coherence relations. First, a small-scale study (Chapter 4) was conducted on causal coherence relations containing the Dutch connective *omdat* ('because'), followed by a more extensive study (Chapter 5) on causal and temporal coherence relations including the Dutch connectives *omdat* or *doordat'* ('because') and *voordat* ('before') or *nadat* ('after').

With respect to textual factors, both studies investigated the effect of connective presence and/or connective type and the order of clauses. A similar facilitating effect of connective presence as has been found for monolingual adults and adolescents in previous studies (Canestrelli, Mak, & Sanders, 2013; Cozijn, Noordman & Vonk, 2011; Van Silfhout, Evers-Vermeul, Mak, & Sanders, 2014; Van Silfhout, Evers-Vermeul, & Sanders, 2015) was not found for beginning readers in general in the present dissertation. This may be explained by the fact that there is more variation in decoding skills among beginning readers (Schaars, Segers, & Verhoeven, 2017) as compared to older/more experiences readers, which makes it more difficult to detect an actual effect. Nevertheless, Chapter 4 showed that the presence of a familiar causal connective supported online processing in beginning L2 readers with lower syntactic knowledge, which extends previous studies with young monolingual readers (Cain & Nash, 2011; Vorstius, Radach, Mayer, & Lonigan, 2013).

Furthermore, Chapter 5 demonstrated that the type of connective affected children's online processing; causal and temporal coherence relations were easier to process when they contained a more frequent/familiar connective. This was particular true for L2 readers' processing of temporal relations, which extends the findings of previous studies that investigated comprehension of temporal relations in monolingual children (Blything & Cain, 2016; Karlsson et al., 2019; Pyykkönen & Järvikivi, 2012). Although children's knowledge of connectives is still developing, the present dissertation demonstrated that beginning readers actually paid attention to connectives during reading and that these coherence markers affected their online processing to some extent, which could not yet be concluded from previous studies using offline measures to investigate the impact of connectives on children's reading comprehension (Crosson & Lesaux, 2013; Kohnen & Retelsdorf, 2019). By contrast, children's online processing of coherence relations was not affected by the order of clauses in both studies; online processing times were similar for causal relations presented in iconic and non-iconic order (Chapters 4 and 5) and for temporal relations presented in chronological and reversed order (Chapter 5).

With respect to child-related factors, the study presented in Chapter 4 examined the role of syntactic knowledge, and the study presented in Chapter 5 additionally investigated the role of working memory capacity. Both studies showed that syntactic knowledge was involved in children's processing of coherence relations. In Chapter 4, individual differences in syntactic knowledge explained variation

in L2 readers' processing of causal coherence relations. In Chapter 5, individual differences in syntactic knowledge explained variation in children's processing of causal and temporal coherence relations and fully accounted for processing differences between L1 readers and L2 readers at the sentence level. The latter finding confirms the results of Chapter 2, by showing that differences in oral language skills explain the reading comprehension performance gap as well as online processing differences between L1 readers and L2 readers. In addition, the study in Chapter 5 indicated that working memory capacity affected children's online processing at the clause level, which is in accordance with a memory-capacity constrained framework (Just & Carpenter, 1992).

The Role of Linguistic Diversity Revisited

With respect to early reading comprehension, the results of the four studies in the present dissertation together indicate that the role of linguistic diversity, when considered as a dichotomy (L1 readers vs. L2 readers), is less prominent than assumed. First, the prediction of L1 and L2 reading comprehension is found to be universal given that similar explanatory factors apply to L1 readers and L2 readers. Second, between-group variation in offline reading comprehension outcomes and online reading comprehension processes could be completely explained by these factors, suggesting that children's language background is no longer relevant once the underlying linguistic and cognitive skills are taken into account. Third, apart from between-group variation in reading comprehension, also considerable within-group variation existed, which asks for other, more refined ways of explaining variation across children from diverse linguistic backgrounds. In that respect, the present dissertation showed that distinguishing children based on their reading comprehension ability (Chapter 2) or levels of L1 and L2 vocabulary knowledge (Chapter 3), or taking into account individual differences in syntactic knowledge and working memory capacity (Chapters 4 and 5) is more relevant than focusing on their language background only.

In contrast to the minor role of linguistic diversity, the studies in the present dissertation confirmed the major role of oral language skills in early reading comprehension. First, the cross-sectional study (Chapter 2) showed that vocabulary knowledge and (morpho)syntactic knowledge were the most important explanatory factors for early reading comprehension performance in both L1 and L2 readers. Second, with respect to vocabulary knowledge, the longitudinal study (Chapter 3) demonstrated that both L1 and L2 vocabulary affected early reading comprehension development in L2 readers. Third, with respect to syntactic

knowledge, the experimental studies (Chapters 4 and 5) showed that individual differences in syntactic knowledge explained variation in online processing of coherence relations among beginning L1 readers and L2 readers.

In light of the Reading System Framework (Perfetti & Stafura, 2014) as a central, overarching theory on reading comprehension, it can be concluded that the various components and processes involved in reading comprehension (as included in the framework) apply to L1 readers as well as L2 readers. This implies that there is no need to adapt this framework for L2 readers or to add linguistic diversity to the framework as an additional factor to better explain variation in reading comprehension. Yet, with respect to the lexicon as a central component in this framework, a refinement could be applied. In case of L2 reading comprehension, lexical quality (Perfetti, 2007) should be determined on the basis of children's first and second language, given that their mental lexicon is integrated across languages (Dijkstra & Van Heuven, 2012) and that their semantic knowledge is depended on both their languages (Monsrud, Rydland, Geva, Thurmann-Moe, & Lyster, 2019).

In terms of theoretical implications, it can thus be argued that the research field should move away from the binary view on linguistic diversity (i.e., comparing L1 and L2 readers to each other) and the homogeneous view on L2 learners (i.e., considering L2 readers as one category). Instead, research on reading comprehension should focus more explicitly on individual differences in the actual linguistic and cognitive skills that are important, ultimately assessing bilingual children in both their languages so that a full picture of children's proficiency is provided.

Limitations and Directions for Future Research

The present dissertation has some limitations that ask for future research to expand on. The first limitation considers the generalizability. All four studies were conducted with Turkish-Dutch bilingual children as L2 readers. It was a well-considered choice to compose an L2 sample with a similar native language across children to exclude variance caused by differences across native languages and to facilitate the assessment of vocabulary knowledge in the native language. However, this limits the generalizability of the findings of the current dissertation. Therefore, future studies with other, more diverse samples of L2 readers are needed to reinforce the claim that the binary view on linguistic diversity does not do justice to individual variation. Moreover, it should be noted that the findings of the experimental studies reported in Chapters 4 and 5 may be specific to our sample

with Turkish-Dutch bilingual children, assuming that the effect of connectives and order of clauses might also depend on the language distance between languages (Nicoladis, 2012) and the amount of word order flexibility languages permit (Kaiser & Trueswell, 2004). Turkish and Dutch are both alphabetic languages with a relatively transparent orthography, but the syntactic structure differs and few cognates exist across the two languages, which may have affected the amount of cross-linguistic influence (Şahin, 2015).

A second limitation is that only one measure of L2 readers' native language skills was administered in the present dissertation (see Chapter 3). Although this receptive vocabulary measure provided important insights into the individual variation among L2 readers, future studies should incorporate a broader range of measures in both languages. For instance, using receptive *and* productive vocabulary measures in both languages to assess breadth as well as depth of L2 readers' vocabulary knowledge (Vermeer, 2001) would be a better indication of their semantic knowledge (Monsrud et al., 2019) and could reveal unique and shared domains of L1 and L2 vocabulary (Wolter, 2006). Also, it would be informative to assess L2 readers syntactic knowledge in both their languages, given that this was an important explanatory factor in children's reading comprehension performance (Chapter 2) as well as their online comprehension processes (Chapters 4 and 5). Furthermore, although some insight was provided into L2 readers' language exposure at home (see Chapter 3), children's home language environment should be investigated in more depth in future studies. Assessing the quantity and quality of language exposure (Paradis, 2017; Unsworth, 2016) and reading and oral language activities (Scheele, Mayo, & Leseman, 2010) in both languages can reveal more detailed insights into the dynamics of bilingualism as a continuum rather than a dichotomy.

A third limitation relates to the assessment of reading comprehension, which was a central construct in the present dissertation. On the one hand, standardized paper-and-pencil tests with multiple-choice questions were used as offline measures of reading comprehension outcomes in Chapters 2 and 3. On the other hand, an experimental sentence-reading task with eye-tracking technology was used as an online measure of reading comprehension processes in Chapters 4 and 5. Given the complexity of reading comprehension as a construct and the assumption that outcomes may differ depending on the type of test that is used (Cutting & Scarborough, 2006; Keenan, Betjemann, & Olson, 2008), it is recommended to use multiple measures and different types of measures to reliably assess children's

reading comprehension in future research. To elaborate on the experimental studies in the present dissertation, an avenue for future research would be to examine online comprehension processes at a more global level by including text passages rather than isolated sentences and/or to focus on other aspects than processing of coherence relations. For instance, eye-tracking technology could also provide more detailed insight into anaphor/pronoun resolution (Eilers, Tiffin-Richards, & Schroeder, 2018), inference making (Sansosti, Was, Rawson, & Remaklus, 2013), and comprehension monitoring (Kim, Vorstius, & Radach, 2018) in beginning readers from diverse linguistic backgrounds.

Implications for Educational Practice

Based on the findings of the present dissertation, several implications for educational practice can be formulated. First, caution needs to be exercised when classifying or selecting children only on the basis of their language background (L1 readers vs. L2 readers), as this does not do justice to individual variation across children and may be off target when it comes to differentiation or allocation of extra support in education. Instead, insights into children's linguistic and cognitive skills and their reading comprehension performance from a developmental perspective should guide educational professionals to better meet individual needs in the classroom and to enable children's full potential.

Second, apart from focusing on decoding skills in the early stages of reading acquisition, it is also crucial to pay explicit attention to children's oral language skills from early onwards. Vocabulary knowledge and syntactic knowledge were found to be the most important explanatory factors for L1 readers as well as L2 readers in the lower primary grades, indicating that knowledge about word meanings and sentence structures is a prerequisite for early reading comprehension. Furthermore, more attention could be devoted to integration processes in formal reading instruction to better support beginning readers in the transition they make from reading isolated words to comprehending sentences and short texts. In that respect, explicit instruction about sentence structures and the meaning and use of connectives may support beginning readers to construct meaning during reading more efficiently.

Finally, bilingualism (and multilingualism) should receive more recognition in education, regardless of the language status. Children from a linguistic minority usually speak more than one language, but minimal attention is paid to their home language environment and their native language skills at school. When only the

language of instruction is taken into account, an incomplete picture of their linguistic proficiency is provided and the impact of their native language is ignored. In terms of assessment, it is important to have an impression of children's native language skills to gain a full picture of their proficiency. This could be fairly simple obtained by means of digital vocabulary tests or educational applications that are available in many languages nowadays. Moreover, to promote reading comprehension in children from a linguistic minority background, their native language skills could be stimulated at home while their second language skills, in particular their vocabulary knowledge, should be stimulated at school to jointly support their semantic knowledge. Taken together, to assess and to stimulate bilingual children's linguistic proficiency and to support their academic performance at school both their languages should be considered and individual differences among children should be taken into account.

References

Babayiğit, S. (2014). The role of oral language skills in reading and listening comprehension of a text: A comparison of monolingual (L1) and bilingual (L2) speakers of English Language. *Journal of Research in Reading*, *37*, 22-47. doi:10.1111/j.1467-9817.2012.01538.x

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*, 525-531. doi:10.1017/S1366728909990423

Blything, L. P., & Cain, K. (2016). Children's processing and comprehension of complex sentences containing temporal connectives: The influence of memory on time course of accurate responses. *Developmental Psychology*, *52*, 1517-1529. doi:10.1037/dev0000201

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*, 429-441. doi:10.1037/a0022824

Canestrelli, A. R., Mak, W. M., & Sanders, T. J. M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, *28*, 1394-1413. doi:10.1080/01690965.2012.685885

Cozijn, R., Noordman, L. G. M., & Vonk, W. (2011). Propositional integration and world knowledge inference: Processes in understanding because sentences. *Discourse Processes*, 48, 475-500. doi:10.1080/0163853X.2011.594421

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading*, *36*, 241-260. doi:10.1111/j.1467-9817.2011.01501.x

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*, 222-251. doi:10.3102/00346543049002222

Cutting, L. E., & Scarborough, H. S. (2009). Prediction of reading comprehension: Contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*, 277-299. doi:10.1207/s1532799xssr1003_5

Dijkstra, T., & Van Heuven, W. J. B. (2012). Word recognition in the bilingual brain. In M. Faust (Ed.), *Handbook of Neuropsychology of Language*. Wiley-Blackwell.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, *38*, 78-103. doi:10.1598/RRQ.38.1.4

Eilers, S., Tiffin-Richards, S. P., & Schroeder, S. (2018). Individual differences in children's pronoun processing during reading: Detection of incongruence is associated with higher reading fluency and more regressions. *Journal of Experimental Child Psychology*, *173*, 250-267. doi:10.1016/j.jecp.2018.04.005

Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*, 389-4212. doi:10.1111/jrir.12003

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, *10*, 330-344. doi:10.1037/a0014320

Hua, A. N., & Keenan, J. M. (2017). Interpreting reading comprehension test results: Quantile regression shows that explanatory factors can vary with performance level. *Scientific Studies of Reading*, *21*, 225-238. doi:10.1080/10888438.2017.1280675

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122-149. doi:10.1037/0033-295X.99.1.122

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, *94*, 113-147. doi:10.1016/j.cognition.2004.01.002

Karlsson, J., Jolles, D., Koornneef, A., Van den Broek, P., & Van Leijenhorst, L. (2019). Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory updating. *Journal of Experimental Child Psychology*, *185*, 1-18. doi:10.1016/j.jecp.2019.04.007

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral language comprehension. *Scientific Studies of Reading*, *12*, 281-300. doi:10.1080/10888430802132279

Kim, Y. S. G., Vorstius, C., & Radach, R. (2018). Does online comprehension monitoring make a unique contribution to reading comprehension in beginning readers? Evidence from eye movements. *Scientific Studies of Reading*, *22*, 367-383. doi:10.1080/10888438.2018.1457680

Kohnen, N., & Retelsdorf, J. (2019). The role of knowledge of connectives in comprehension of a German narrative text. *Journal of Research in Reading*, *42*, 371-388. doi:10.1111/1467-9817.12273

Language and Reading Research Consortium (LRRC), & Logan, J. (2017). Pressure points in reading comprehension: A quantile multiple regression analysis. *Journal of Educational Psychology*, *109*, 451-464. doi:10.1037/edu0000150

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, *19*, 99-131. doi:10.1007/s11145-005-4713-6

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction

between language proficiency and usage. *Journal of Cognitive Psychology*, *25*, 605-621. doi: 10.1080/20445911.2013.795574

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409-433. doi:10.1037/a0033890

Monsrud, M. B., Rydland, V., Geva, E., Thurmann-Moe, A. C., Lyster, S. A. H. (2019). The advantages of jointly considering first and second language vocabulary skills among emergent bilingual children. *International Journal of Bilingual Education and Bilingualism*. doi:10.1080/13670050.2019.1624685

Nicoladis, E. (2012). Cross-linguistic influence in French-English bilingual children's Possessive constructions. *Bilingualism: Language and Cognition*, *15*, 320-328. doi:10.1017/S1366728911000101

Paradis, J. (2017). Parent report data on input and experience reliably predict bilingual development and this is not trivial. *Bilingualism: Language and Cognition*, *20*, 27-28. doi:10.1017/S136672891600033X

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

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22-37. doi:10.1080/10888438.2013.827687

Pyykkönen, P., & Järvikivi, J. (2012). Children and situation models of multiple events. *Developmental Psychology*, *48*, *2*, 521-529. doi:10.1037/a0025526

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, *328*, 450-452. doi:10.1126/science.1182597

Sansosti, F. J., Was, C., Rawson, K. A., & Remaklus, B. L. (2013). Eye movements during processing of text requiring bridging inferences in adolescents with higher functioning autism spectrum disorders: A preliminary investigation. Research in Autism Spectrum Disorders, *7*, 1535-1542. doi:10.1016/j.rasd.2013.09.001

Şahin, H. (2015). *Cross-linguistic influences: Dutch in contact with Papiamento and Turkish* (Doctoral dissertation). Retrieved from https://www.lotpublications.nl/Documents/405_fulltext.pdf

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2017). Word decoding development in incremental phonics instruction in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal*, *30*, 1529-1550. doi:10.1007/s11145-017-9735-3

Scheele, A. F., Leseman, P. P. M., & Mayo, A. F. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*, 117-140. doi:10.1017/S0142716409990191

Tighe, E. L., & Schatschneider, C. (2016). A quantile regression approach to understanding the relations among morphological awareness, vocabulary and reading comprehension in adult basic education students. *Journal of Learning Disabilities*, *49*, 424-436. doi:10.1177/0022219414556771

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Lifespan perspectives on bilingualism* (pp. 136-196). Mouton de Gruyter/APA.

Van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. M. (2014). Connective and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology*, *106*, 1036-1048. doi:10.1037/a0036293

Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, 52, 47-76. doi:10.1080/0163853X.2014.905237

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*, 313-330. doi:10.1207/S1532799XSSR0404_4

Verhoeven, L., & Van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing: An Interdisciplinary Journal*, *25*, 1805-1818. doi:10.1007/s11145-011-9346-3

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*, 217-234.

Vorstius, C., Radach, R., Mayer, M. B., & Lonigan, C. J. (2013). Monitoring local comprehension monitoring in sentence reading. *School Psychology Review*, *32*, 191-206.

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics*, *27*, 741-747. doi:10.1093/applin/aml036

Nederlandse samenvatting
Dankwoord
Curriculum Vitae

Nederlandse Samenvatting

Lezen is een van de belangrijkste vaardigheden die kinderen leren op de basisschool. Het stelt hen in staat om geschreven tekst te kunnen decoderen en begrijpen, hetgeen cruciaal is op school én daarbuiten. Begrijpend lezen is echter een complexe vaardigheid; relatief veel kinderen op de basisschool hebben moeite met het begrijpen van teksten. Dit blijkt met name te gelden voor kinderen die thuis een andere taal spreken en op school leren lezen in hun tweede taal. Veel van deze tweede-taallezers ondervinden problemen bij begrijpend lezen terwijl zij wel over een goede technische leesvaardigheid beschikken. Al in de onderbouw van de basisschool, waar leren lezen centraal staat en teksten nog relatief eenvoudig zijn, lijken deze problemen zich voor te doen. Eerder onderzoek naar begrijpend lezen bij eerste- en tweede-taallezer richtte zich echter vooral op de bovenbouw, waar teksten complexer en kinderen meer ervaren lezers zijn. Er is veel minder bekend over de vroege fasen van begrijpend lezen in de onderbouw, waar kinderen de overstap maken van het leren lezen van letters en woorden naar het begrijpen van zinnen en korte teksten. Gezien het belang en de complexiteit van begrijpend lezen, is het belangrijk om inzicht te verkrijgen in verschillen tussen beginnende lezers in de vroege fasen van begrijpend lezen, hoe deze verschillen te verklaren zijn en welke rol hun taalachtergrond daarbij speelt.

Het huidige proefschrift is gericht op begrijpend lezen in de onderbouw van de basisschool. Er is onderzoek gedaan naar begrijpend leesprestaties, begrijpend leesontwikkeling en begrijpend leesprocessen van beginnende lezers met verschillende taalachtergronden, namelijk eentalige Nederlandse kinderen (eerste-taallezers) en tweetalige Turks-Nederlandse kinderen (tweede-taallezers). Daarbij is onderzocht hoe groepsverschillen én individuele verschillen verklaard kunnen worden. Er is voor Turks-Nederlandse kinderen gekozen, omdat Turks een van de meest gesproken andere thuistalen is onder basisschoolleerlingen in Nederland. Er zijn in totaal vier studies uitgevoerd, waarvan de opzet en de belangrijkste bevindingen achtereenvolgens besproken zullen worden.

Begrijpend Leesprestaties

In hoofdstuk 2 wordt een cross-sectionele studie beschreven over de begrijpend leesprestaties van leerlingen in groep 4. Allereerst is er gekeken naar verschillen tussen de eerste- en tweede-taallezers. Daarbij bleek dat er in de beginfase van begrijpend lezen (halverwege groep 4) al een zogenaamde prestatiekloof bestaat; gemiddeld genomen laten de tweede-taallezers een achterstand zien in begrijpend

lezen ten opzichte van hun eentalige klasgenoten. Vervolgens is nagegaan welke factoren gerelateerd zijn aan begrijpend leesprestaties en in hoeverre deze verklarende factoren verschilden voor eerste- en tweede-taallezers. De resultaten lieten zien dat technische leesvaardigheid (het kunnen decoderen van woorden), woordenschat (kennis over de betekenis van woorden) en morfosyntactische kennis (kennis over de structuur van woorden en zinnen) gerelateerd waren aan begrijpend lezen in groep 4, terwijl het werkgeheugen en de taalachtergrond van leerlingen hier niet aan bijdroegen. Deze factoren waren van belang voor begrijpend lezen bij zowel eerste-taallezers als tweede-taallezers. Bovendien bleek er behalve variatie in leesprestaties tussen de groepen ook grote variatie in leesprestaties te bestaan binnen de groepen. Tezamen suggereren deze resultaten dat het maken van een tweedeling op basis van de taalachtergrond van leerlingen dan ook minder relevant is dan gedacht werd, als het gaat om het verklaren van verschillen in begrijpend lezen.

Daarom is vervolgens nagegaan of het relevanter is om een onderscheid te maken tussen leerlingen op basis van hun niveau van begrijpend lezen in plaats van hun taalachtergrond. Wederom is gekeken welke factoren gerelateerd waren aan begrijpend lezen en in hoevewe deze verschilden voor leerlingen met een laag, gemiddeld of hoog niveau van begrijpend lezen. Er bleken inderdaad enkele verschillen te bestaan tussen de niveaus wat betreft de factoren die een rol speelden bij begrijpend lezen. Technische leesvaardigheid en morfosyntactische kennis waren van belang ongeacht het niveau van begrijpend lezen, woordenschat was vooral belangrijk voor leerlingen met een laag niveau en werkgeheugen voor leerlingen met een hoog niveau van begrijpend lezen.

Kortom, de bevindingen van deze studie laten zien dat de factoren die van belang zijn voor begrijpend lezen niet verschillen op basis van de taalachtergrond van kinderen, maar wel op basis van hun niveau van begrijpend lezen. Het is dan ook belangrijk om individuele verschillen in de begrijpend leesprestaties van beginnende lezers in kaart te brengen.

Begrijpend Leesontwikkeling

In hoofdstuk 3 wordt een longitudinale studie beschreven over de ontwikkeling van begrijpend lezen bij leerlingen in groep 4 en 5. Om na te gaan of de prestatiekloof tussen eerste- en tweede-taallezers verandert naarmate zij meer ervaren worden in het lezen, is er gekeken naar de ontwikkeling van hun begrijpend leesprestaties gedurende twee schooljaren. Hoewel deze kloof in groep 5 nog steeds bestond,

werd het verschil tussen de begrijpend leesprestaties van eerste- en tweede-taallezers wel kleiner. Met name in groep 5, waar begrijpend lezen een centralere rol krijgt in het leesonderwijs, lieten de tweede-taallezers een grotere groei in begrijpend lezen zien dan hun eentalige klasgenoten. Het gaat hier echter om groepsgemiddelden van eerste- en tweede-taallezers, terwijl eerder is gesuggereerd dat een tweedeling op basis van taalachtergrond niet altijd relevant is en juist het kijken naar individuele verschillen van groot belang is.

Daarom is er in deze studie onderzocht hoe individuele verschillen binnen de groep van tweede-taallezers verklaard kunnen worden op basis van hun woordenschat. Woordenschat speelt immers een belangrijke rol bij begrijpend lezen. In tegenstelling tot veel bestaande studies, is er in deze studie gekeken naar de woordenschat van kinderen in hun beide talen, namelijk Nederlands en Turks. Op basis daarvan konden er vier subgroepen worden geïdentificeerd; (1) kinderen met een lage woordenschat in beide talen, (2) kinderen met een hoge Turkse woordenschat en een lage Nederlandse woordenschat, (3) kinderen met een lage Turkse woordenschat en een hoge Nederlandse woordenschat, en (4) kinderen met een hoge woordenschat in beide talen. Vervolgens is gekeken naar de begrijpend leesontwikkeling van deze subgroepen in vergelijking met hun eentalige klasgenoten. Er bleken enkele verschillen te bestaan tussen de subgroepen. Met name de subgroepen met een hoge Turkse woordenschat lieten een grotere groei in begrijpend lezen zien dan hun eentalige klasgenoten; de prestatiekloof werd kleiner gedurende groep 4 en 5. De subgroep met hoge woordenschat in beide talen liep de achterstand op hun eentalige klasgenoten zelfs in; halverwege groep 5 was er geen sprake meer van een prestatiekloof.

Samengevat laten de resultaten van deze studie zien dat het belangrijk is om tweetalige kinderen, net als eentalige kinderen, niet als één homogene groep te zien maar individuele verschillen in woordenschat in acht te nemen. Het is daarbij essentieel om beide talen te betrekken om een zo compleet mogelijk beeld te krijgen van hun mentale lexicon en om variatie tussen tweetalige kinderen beter in kaart te kunnen brengen.

Begrijpend Leesprocessen

In de hoofdstukken 4 en 5 worden experimentele studies beschreven over de begrijpend leesprocessen van leerlingen in groep 5. Met behulp van eye-tracking technologie zijn de oogbewegingen van leerlingen gemeten tijdens het lezen om inzicht te krijgen in hun cognitieve processen. Omdat het beginnende lezers

betrof, werd er ingezoomd op begrijpend leesprocessen op zinsniveau waarbij het leggen van een semantische relatie tussen twee zinsdelen centraal stond. De eerste, kleinschaligere studie in hoofdstuk 4 ging specifiek over causale relaties (bijv. 'Tom at een appel omdat hij honger had.') en de tweede, uitgebreidere studie in hoofdstuk 5 ging daarnaast ook over temporele relaties (bijv. 'Voordat Laura buiten ging spelen, maakte ze haar huiswerk.'). In beide studies werd onderzocht welke factoren (m.b.t. de tekst en de lezer) van invloed waren op de begrijpend leesprocessen van beginnende lezers met verschillende taalachtergronden.

Hoewel de focus lag op het verklaren van individuele verschillen tussen beginnende lezers in het algemeen, is er in eerste instantie nagegaan in hoeverre er groepsverschillen tussen eerste- en tweede-taallezers bestonden wat betreft hun begrijpend leesprocessen. De resultaten duiden erop dat de tweede-taallezers in vergelijking met de eerste-taallezers in het algemeen meer moeite hadden met het verwerken van zinnen met een causale of temporele relatie (hoofdstuk 5). De tweede-taallezers hadden namelijk langere gemiddelde leestijden voor de zinnen in het geheel en voor de zinsdelen apart dan de eerste-taallezers, terwijl de groepen niet verschilden in niveau van technisch lezen. Dit wijst erop dat niet zozeer hun decodeervaardigheid, maar wél hun begripsprocessen tijdens het lezen minder efficiënt zijn.

Net als bij begrijpend leesprestaties (hoofdstuk 2) en begrijpend leesontwikkeling (hoofdstuk 3), is er ook bij begrijpend leesprocessen (hoofdstuk 4 en 5) onderzocht hoe individuele verschillen tussen beginnende lezers verklaard konden worden. Daarbij is de rol van verschillende tekst- en lezerskenmerken bestudeerd. Wat betreft tekstkenmerken bleek er in beide experimentele studies geen invloed te zijn van de volgorde van zinsdelen op de begrijpend leesprocessen. Hoewel er eveneens geen effect van een causaal verbindingswoord ('omdat'/'doordat') op de begrijpend leesprocessen van beginnende lezers in het algemeen werd gevonden, bleek er wel een effect te bestaan voor tweede-taallezers met minder grammaticale kennis (hoofdstuk 4). Dit suggereert dat tweede-taallezers baat hebben bij het expliciet maken van de relatie tussen zinsdelen met een verbindingswoord wanneer zij over weinig grammaticale kennis beschikken. Tevens bleek er een effect te zijn van het type verbindingswoord. De verwerking van temporele relaties verliep efficiënter bij een meer frequent voorkomend verbindingswoord ('voordat') dan bij een minder frequent verbindingswoord ('nadat'), vooral bij tweede-taallezers (hoofdstuk 5).

Wat betreft lezerskenmerken bleek met name de grammaticale kennis van kinderen, maar ook hun werkgeheugen invloed te hebben op hun begrijpend leesprocessen. Zoals eerder beschreven is, liet de studie in hoofdstuk 4 zien dat het effect van een causaal verbindingswoord ('omdat') afhankelijk was van de grammaticale kennis bij de tweede-taallezers. Daarnaast bleek dat de grammaticale kennis van beginnende lezers ook van invloed was op het verwerken en begrijpen van zinnen met causale/temporele relaties; hoe meer grammaticale kennis hoe efficiënter de begripsprocessen op zinsniveau verliepen (hoofdstuk 5). Bovendien verklaarden individuele verschillen in grammaticale kennis in veel gevallen de groepsverschillen in begrijpend leesprocessen tussen eerste- en tweede-taallezers. Ook individuele verschillen in werkgeheugen speelden een rol bij het verwerken/begrijpen van causale en temporele relaties; hoe groter de capaciteit van het werkgeheugen, hoe efficiënter de begripsprocessen op zinsdeelniveau verliepen.

De resultaten van deze twee experimentele studies laten zien dat zowel kenmerken van de tekst als kenmerken van de lezer de leesprocessen van beginnende lezers beïnvloeden. Met name individuele verschillen in grammaticale kennis blijken van invloed te zijn op hoe efficiënt hun begripsprocessen op zinsniveau verlopen tijdens het lezen.

Conclusies

Allereerst kan op basis van de bevindingen geconcludeerd worden dat de rol van taalachtergrond minder prominent is dan verondersteld werd wanneer uitgegaan wordt van een tweedeling (d.w.z. eentalig vs. tweetalig). Hoewel er gemiddeld genomen sprake is van groepsverschillen in begrijpend leesprestaties, begrijpend leesontwikkeling en begrijpend leesprocessen tussen eentalige en tweetalige kinderen, blijkt er een drietal factoren van belang te zijn voor begrijpend lezen ongeacht taalachtergrond, te weten: technisch lezen, woordenschat en morfosyntactische kennis. Bovendien bleek dat de groepsverschillen geheel te verklaren waren door individuele verschillen in deze factoren, waardoor het label 'eentalig/tweetalig' geen toegevoegde waarde meer heeft. Dit laatste wordt nog eens versterkt door het feit dat er behalve verschillen tussen groepen ook aanzienlijke verschillen binnen de twee groepen bestonden wat betreft begrijpend lezen; een dergelijke tweedeling doet dan ook onvoldoende recht aan de grote variatie tussen kinderen.

Ten tweede bevestigen de bevindingen van de vier studies de essentiële rol van (mondelinge) taalvaardigheid bij begrijpend lezen. Zo blijkt dat ook bij beginnende lezers woordenschat en grammaticale kennis zeer belangrijke factoren zijn wat betreft begrijpend leesprestaties, begrijpend leesontwikkeling en begrijpend leesprocessen. Kennis van woorden en begrip van zinsstructuren zijn voorwaardes om een tekst te kunnen verwerken en begrijpen, ook in de onderbouw waar teksten nog relatief kort en eenvoudig lijken. Daarbij geldt dat voor tweetalige kinderen niet alleen de woordenschat in de schooltaal, maar ook de woordenschat in de thuistaal een rol speelt bij het lezen in de tweede taal. Tezamen geeft de woordenschat in de eerste en tweede taal immers een completer beeld van hun mentale lexicon, in het bijzonder hun conceptuele kennis, hetgeen van groot belang is bij de ontwikkeling van begrijpend lezen in de tweede taal op school.

Met betrekking tot de theorie kan gesteld worden dat het categoriseren van kinderen op basis van hun taalachtergrond, zoals gedaan wordt in veel bestaande studies naar begrijpend lezen, niet wenselijk is. In plaats van een tweedeling te maken op basis van taalachtergrond en tweetalige kinderen als één homogene groep te zien, dient er binnen het onderzoeksveld meer aandacht te komen voor individuele verschillen om recht te kunnen doen aan variatie tussen kinderen met diverse taalachtergronden. Om deze individuele verschillen goed in kaart te kunnen brengen dient er niet alleen gekeken worden naar taalvaardigheid in de schooltaal, maar ook naar de taalvaardigheid in de thuistaal van tweetalige kinderen. Dit kan eraan bijdragen dat talige diversiteit niet langer als dichotomie maar als continuüm beschouwd wordt in het onderzoeksveld.

Implicaties voor de Onderwijspraktijk

Tot slot kunnen er op basis van de bevindingen van dit proefschrift enkele implicaties voor de onderwijspraktijk beschreven worden. Zo is het essentieel dat er naast technisch lezen ook voldoende aandacht wordt besteed aan het ontwikkelen van de (mondelinge) taalvaardigheid van leerlingen in de onderbouw. In de vroege fase van begrijpend lezen blijken woordenschat en grammaticale kennis al een belangrijke rol te spelen bij zowel eerste- als tweede-taallezers. Bij woordenschat gaat het niet alleen om kennis van inhoudswoorden, maar ook van functiewoorden, zoals verbindingswoorden. Kennis over de betekenis en functie van verbindingswoorden kan bijdragen aan efficiëntere begripsprocessen en betere begrijpend leesprestaties. Bovendien zouden beginnende lezers er baat bij kunnen hebben wanneer er expliciet aandacht besteed wordt aan het leggen

van verbanden binnen en tussen zinnen om hen te ondersteunen in de overstap van het lezen van woorden naar het lezen van zinnen en korte teksten.

Wat betreft differentiatie in de klas en het bieden van extra ondersteuning bij begrijpend lezen is het niet raadzaam om de taalachtergrond van leerlingen als leidraad te nemen. Er is immers sprake van grote variatie tussen kinderen; tweede-taallezers hebben niet per definitie moeite met begrijpend lezen en eerste-taallezers zijn niet per definitie goede begrijpend lezers. Daarentegen zou de focus moeten liggen op individuele verschillen in taal- en leesvaardigheid om te kunnen voldoen aan individuele behoeften van leerlingen in de klas.

Ook dient op scholen erkenning en waardering te zijn voor andere thuistalen van leerlingen, ongeacht de taalstatus. In de huidige samenleving is er immers steeds vaker sprake van culturele én talige diversiteit binnen klassen. Niet alleen de schooltaal, maar ook de thuistaal maakt deel uit van de identiteit en vaardigheden van tweetalige leerlingen. Wat betreft assessment is het dus van belang om bij tweetalige leerlingen een indicatie te hebben van hun taalvaardigheid in *beide* talen om zo individuele leerpotenties beter in kaart te kunnen brengen. Hierbij kan ICT een belangrijke rol spelen door het inzetten van digitale toetsen en het gebruik van educatieve applicaties. Tenslotte is het raadzaam om naast het stimuleren van de taalvaardigheid in de tweede taal op school, ook de taalvaardigheid in de eerste taal te stimuleren in de thuisomgeving om zo de conceptuele kennis van kinderen te vergroten, hetgeen van essentieel belang is voor hun begrijpend leesvaardigheid.

Dankwoord

Tijdens mijn studie Onderwijswetenschappen ontdekte ik mijn passie voor onderzoek doen. Na het schrijven van twee scripties en het doen van twee onderzoeksstages, kreeg ik de kans om deze passie voort te zetten als promovenda. Mijn promotietijd is enorm waardevol geweest; ik heb me gedurende deze jaren kunnen ontwikkelen op academisch én persoonlijk vlak. Een woord van dank aan degenen die hieraan een bijdrage hebben geleverd.

Allereerst wil ik mijn promotoren bedanken voor de kansen die ze mij hebben geboden en voor het vertrouwen in mij. Jullie gaven mij vanaf het begin af aan de vrijheid om dit project vorm te geven en gaven mij waardevolle input op de momenten waarop ik dat nodig had. Ludo, je vermogen om de grote lijnen in kaart te brengen, je kennis van de theorie en de praktijk en je enthousiasme inspireren mij. Eliane, je scherpe opmerkingen, snelle manier van denken en oog voor detail stimuleerden mij om steeds een (denk)stapje extra te zetten. Samen vormden jullie een ideaal begeleidingsteam en ik kijk dan ook met veel plezier terug op onze tientallen overlegmomenten. Jullie expertise, betrokkenheid en positieve woorden waren van grote waarde bij de totstandkoming van mijn proefschrift en bij mijn persoonlijke ontwikkeling. Bedankt voor het delen van jullie netwerk en voor alle wijze raad die jullie mij als 'wetenschapper in de dop' gegeven hebben.

I would like to thank Prof. Caroline Rowland, Prof. Elma Blom and Prof. Ad Backus for being part of the manuscript committee. Ook wil ik de overige leden van de commissie, Dr. Mienke Droop, Dr. Sharon Unsworth en Dr. Anne Helder, bedanken voor hun bereidheid om tijdens de verdediging met mij van gedachten te wisselen over mijn proefschrift.

Uiteraard wil ik de basisscholen bedanken die hebben deelgenomen aan dit onderzoek voor hun medewerking, gastvrijheid en enthousiasme. Het contact met de onderwijspraktijk, de leerkrachten en de leerlingen zorgden voor afwisseling en uitdaging en gaf mij extra motivatie om dit project tot een goed einde te brengen. Daarnaast wil ik de bachelor- en masterstudenten bedanken voor hun hulp bij het testen van alle leerlingen. Ook bedankt voor de ondersteuning die ik kreeg vanuit het PWO secretariaat (Lanneke en Christel) en de BSI directie (Toon, Meta en Ralph) en voor de hulp van de Technical Support Group (Hubert, Gerard en Ronny) en Nina Davids bij het gebruik van de eye-tracker.

Mijn collega's hebben mijn promotietijd in Nijmegen extra plezierig gemaakt. Ik kwam als nieuwkomer op de afdeling, maar voelde mij hier al snel op mijn plek. Een hechte groep promovendi bij wie ik veel gezelligheid en steun vond. Een aantal hiervan wil ik in het bijzonder bedanken. Petrie, wij startten tegelijkertijd met ons promotieonderzoek en al snel raakten wij goed bevriend. De afgelopen jaren hebben we veel kunnen delen, zowel op het werk als daarbuiten, en daar ben ik je heel dankbaar voor. Joyce, met jou heb ik aan het begin van mijn promotie veel tijd doorgebracht als kantoorgenoten. Bedankt voor de gezelligheid en je nuttige adviezen door de jaren heen. Carolien, jouw enthousiasme en gedrevenheid zijn aanstekelijk. Bedankt voor alle gezellige theemomentjes en onze fijne samenwerking op het gebied van eye-tracking. Henriëtte, onze promotietrajecten liepen parallel en hadden bovendien veel gelijkenissen. Ik vond het heel fijn dat wij niet alleen onze kennis en ervaringen, maar ook vele congresbezoeken met elkaar konden delen. Ook wil ik de dames van de Consumptieclub (Nicole, Moniek, Evelien, Sanne en Helen) bedanken voor de gezelligheid en de support. Onder het genot van een hapje en een drankje hebben wij veel gedeeld (en gelachen!) tijdens deze bijzonder periode. Sanne, onze 'Westerse roots' schepten meteen een band en wat fijn dat ik altijd bij je terecht kon met een moeilijke vraag en voor een gezellig praatje. Helen, wat was het fijn om met jou een kantoor te delen en tegelijkertijd onze promotieprojecten af te ronden. Dankjewel voor de aanmoediging en de fijne gesprekken die over zoveel meer dan alleen ons onderzoek gingen.

Daarnaast zijn er nog een aantal andere collega's die ik graag wil noemen. Femke, Wanda, Heleen, Nicole, Christel, Mienke en Eliane, bedankt voor de prettige samenwerking binnen het 'Tel mee met taal' project. Het was voor mij erg leerzaam om een praktijkgericht onderzoek uit te voeren in teamverband. Henny, Marja, Marijke en Marianne, onder jullie begeleiding heb ik mijn eerste stappen gezet als docent op de universiteit. Bedankt voor jullie enthousiasme, coaching en vertrouwen in mij. Dit heeft eraan bijgedragen dat ik na mijn promotietijd de stap nam om mij verder te ontwikkelen als docent op de universiteit. Ook wil ik mijn nieuwe collega's van de programmagroep Onderwijswetenschappen aan de Universiteit Leiden bedanken voor het warme welkom en de belangstelling voor mijn promotieonderzoek. Het geeft me veel voldoening om studenten te begeleiden en mijn passie voor onderzoek doen aan hen over te dragen.

Ik prijs me gelukkig met de lieve vriendinnen om me heen en wil hen bedanken voor alle gezellige momenten samen, dit zorgde voor de nodige ontspanning en afleiding. Marloes en Marisse, wij kennen elkaar al voordat we konden lezen en

hebben sindsdien enorm veel samen meegemaakt (van het schoolplein tot aan het moederschap), jullie zijn heel belangrijk voor me. Diede, Daniëlle, Eveline, Carina, Jessica en Juliëtte, hoewel we na onze middelbare schooltijd allemaal een andere richting op gingen (letterlijk en figuurlijk), ben ik heel blij dat we nog steeds zo'n hecht clubje zijn. Maaike en Aliza, samen beleefden wij onze studententijd in Utrecht, daar begon onze vriendschap en die is me heel dierbaar. Maaike, wat bijzonder dat ik met jou als vriendin alles kan delen op persoonlijk én academisch gebied, met een lach én een traan.

Een speciaal woord van dank gaat uit naar mijn familie. Ik vind het heel bijzonder dat mijn opa's en oma's al zo lang in mijn leven zijn en het maakt me trots dat jullie deze mijlpaal met mij mee vieren. Jan en El, ik waardeer jullie behulpzaamheid en betrokkenheid en vind het ontzettend leuk om Aniek en Sven zo groot en wijs te zien worden. Familie Van den Berg, ik kom inmiddels al heel wat jaren over de vloer bij jullie 'mannengezin' en voel me hier heel welkom. Geert en Astrid, bedankt dat jullie altijd voor ons klaar staan en een geweldige opa en oma zijn voor Jens. Dat geldt ook voor mijn ouders. Lieve pap en mam, jullie hebben mij geleerd wat hard werken is en hebben mij altijd gestimuleerd om mijn eigen keuzes te maken. Ik ben heel dankbaar voor jullie liefde, steun en betrokkenheid bij alles wat ik doe. Ook mijn twee zussen wil ik bedanken. Lieve Roxette, ik vind het heel bijzonder dat ik het promoveren met jou kan delen als geen ander. Dankjewel voor je steun, je luisterend oor en je nuttige feedback en voor de gezellige momenten met jou en Rieuwert. Lieve Nicole, jouw motto 'Het komt goed' blijkt wederom te kloppen. Dankjewel voor je vertrouwen, de afleiding en de lol die wij samen hebben.

Tot slot richt ik me tot mijn eigen gezinnetje. Lieve Jens, jouw komst maakte het laatste jaar van mijn promotietijd extra bijzonder en uitdagend. Je bent een verrijking van ons leven en wat is het mooi om te zien hoe geliefd je bent bij iedereen om ons heen. Je bent nog maar klein, maar nu al zo'n nieuwsgierig en ijverig ventje. Ik kijk ernaar uit om je te zien opgroeien en samen nog veel meer mooie avonturen te beleven! Lieve Erwin, *time flies when you're having fun*. Het ene moment stuur je een berichtje om me succes te wensen met mijn eindexamen, het andere moment zit je op de eerste rij bij mijn verdediging. In de tussentijd, een periode van ruim tien jaar, was jij mijn trouwste supporter. Je hebt altijd het volste vertrouwen in mij en stimuleert me om het beste uit mezelf te halen, zelfs toen ik verhuisde naar de andere kant van het land. Dankjewel voor je onvoorwaardelijke liefde en steun. Wat is het fijn om nu écht thuis te komen bij jou en samen het leven te vieren. Ik kijk uit naar onze toekomst!

Curriculum Vitae

Liza van den Bosch is geboren op 27 augustus 1991 te Nootdorp. Na het behalen van haar vwo-diploma aan het Erasmus College in Zoetermeer, begon ze in 2009 met de bachelor Onderwijswetenschappen aan de Universiteit Utrecht. Tijdens haar bachelor nam ze deel aan het honoursprogramma van de faculteit Sociale Wetenschappen en behaalde zij de prijs voor 'Beste bachelor thesis'. Vervolgens startte Liza in 2012 met de research master 'Educational Sciences: Learning in Interaction', eveneens aan de Universiteit Utrecht. Tijdens haar master deed zij een buitenlandse stage bij de onderzoeksgroep van Prof. Kate Cain (Lancaster, UK) en een onderzoeksstage bij Cito. Haar interesse in begrijpend lezen en meertaligheid vertaalde zich in het opzetten van een multidisciplinair onderzoeksproject voor haar master thesis in samenwerking met Prof. Paul Leseman en Prof. Ted Sanders.

Na het behalen van de research master (cum laude) begon Liza in september 2014 met haar promotieonderzoek bij het Behavioural Science Institute (BSI) aan de Radboud Universiteit in Nijmegen. Zij schreef een eigen projectvoorstel over begrijpend lezen bij beginnende lezers met verschillende taalachtergronden, dat onder begeleiding van Prof. Ludo Verhoeven en Prof. Eliane Segers leidde tot het huidige proefschrift. Haar bevindingen presenteerde ze aan onderzoekers en onderwijsprofessionals op nationale en internationale conferenties. Naast haar promotieonderzoek was Liza ook betrokken bij een NRO-project in het kader van 'Tel mee met Taal' over leesmotivatie bij vmbo-leerlingen in samenwerking met het Expertisecentrum Nederlands en het ITTA. Tevens verzorgde Liza onderwijs aan bachelor- en masterstudenten bij Pedagogische Wetenschappen en Onderwijswetenschappen en was zij lid van de Onderdeelcommissie (OC) van de Faculteit Sociale Wetenschappen. In 2018 behaalde zij haar Basiskwalificatie Onderwijs (BKO).

Na het afronden van haar promotieonderzoek vervolgde Liza haar carrière als docent Onderwijswetenschappen bij de Universiteit Leiden. Sinds augustus 2019 verzorgt zij hier onderwijs aan bachelor- en masterstudenten binnen de opleidingen Pedagogische Wetenschappen en de Academische Pabo.

Behavioural
Science
Institute

Radboud University

Reading is one of the most crucial skills that children acquire at primary school. Yet, comprehending written text is a complex process and considerable variation across children exists. Some children easily understand texts whereas others struggle to construct meaning. Reading comprehension can be particularly challenging for children who learn to read in their second language.

The present dissertation examined variation in early reading comprehension in first and second language readers. By means of cross-sectional, longitudinal and experimental studies, new insights are provided into reading comprehension performance, development and online processes in beginning readers from diverse linguistic backgrounds.