The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/204424

Please be advised that this information was generated on 2021-07-04 and may be subject to change.
Cognitive and linguistic precursors of early first and second language reading development

Moniek M.H. Schaars, E. Segers, L. Verhoeven

Behavioural Science Institute, Radboud University, the Netherlands

Abstract

The present large-scaled longitudinal prediction study examined cognitive and linguistic precursors of early word decoding and reading comprehension from kindergarten to Grade 3 in 613 first language (L1) and 109 second language (L2) learners of Dutch. L1 learners outperformed L2 learners on reading comprehension, and on kindergarten vocabulary, rapid naming (RAN), and phoneme segmentation. No differences were found on word decoding across the grades, kindergarten grapheme knowledge, phoneme isolation, or short term memory (STM). Despite L2 learners’ delay in reading comprehension and language-related precursors, the developmental paths and structural relations of L2 learners were highly similar to those of L1 learners. For both groups, RAN, grapheme knowledge and STM predicted word decoding development. Word decoding, phonemic awareness, vocabulary, and STM predicted reading comprehension. There were strong autoregressive effects of both word decoding and reading comprehension. In kindergarten, L2 learners showed delays in RAN, phonological awareness, and vocabulary. These measures were all indicative of future reading.

1. Introduction

Decoding and comprehending written text are two core abilities in literate societies which have been shown to be highly related (e.g., Garcia & Cain, 2014; Gough & Tunmer, 1986). Word decoding is the conversion of orthographic into phonetic code, and reading comprehension is the understanding of the message that has been read. Children nowadays often learn to read in a second language (Durgunoglu & Verhoeven, 2013). Most of these children are sequential bilinguals, who start to acquire their second language (L2) in preschool or kindergarten. Research has indicated that the development of word decoding skills is highly comparable for second language (henceforth L2) and first language (henceforth L1) learners across orthographies (Geva, Yaghoub-Zadeh, & Schuster, 2000; Lesaux & Siegel, 2003; Lipka & Siegel, 2012; Mancilla-Martinez & Lesaux, 2010; Melby-Lervåg & Lervåg, 2014; Raudszus, Segers, & Verhoeven, 2018). However, L2 learners often lag behind in reading comprehension (English as L1; Babayigit, 2014; Farnia & Geva, 2013; Melby-Lervåg & Lervåg, 2014).

Prior to reading development, children start to develop cognitive and linguistic precursor skills. Cognitive skills are the mental actions or processes of acquiring knowledge and understanding through thought and experience, such as lexical retrieval and short-term memory. Linguistic skills concern language related skills such as phonological awareness, letter knowledge, and vocabulary. These cognitive and linguistic skills have been found to strongly impact the reading development in both L1 and L2 learners during the first years in primary school (Caravolas, Lervåg, Defior, Målková, & Hulme, 2013; Farnia & Geva, 2013). Nonetheless, comparative longitudinal studies on L1 and L2 reading have not taken into account initial differences in cognitive and linguistic abilities before formal reading instruction, whereas these initial differences may very well explain the later differences in reading comprehension (Adlof, Catts, & Lee, 2010).

The interrelations between early word decoding and early reading comprehension have not comprehensively been studied from a developmental perspective (but see Verhoeven & van Leeuwe, 2008). As a consequence, the early parallel development and the reciprocal relationships during early development have not yet been studied. It is still unclear to what extent the development of L1 and L2 word decoding and reading comprehension during the early primary grades can be explained from children’s cognitive and linguistic skills measured before formal reading instruction. In the present study, therefore, we examined reading development in children learning to read Dutch as L1 and French as L2 and their relations with linguistic and cognitive precursor skills.
and L2 in Grades 1 to 3, and related their development to their cognitive and linguistic skills in kindergarten.

1.1. Development of word decoding and reading comprehension

Learning to read involves the development of word decoding and reading comprehension skills (Ehri, 2005). Children start with learning how graphemes systematically correspond to phonemes and how the latter can be combined to construct words. This conversion of print into spoken language by systematically mapping and blending the phonological elements within words facilitates word decoding. One model to describe word decoding development is the restrictive-interactive model by Perfetti (1992, 2007), which is based on an incremental acquisition process. The correct identification of written text is supposedly influenced by the precision and redundancy of underlying representations in the mental lexicon (Perfetti, 1992). After the basic principles of word decoding have been acquired, further word decoding skills typically develop in a self-teaching manner, and gradually the mental representations become more precise, redundant, and efficient (Share, 2004; Tucker, Castles, Laroche, & Deacon, 2016). In other words, typically developing children become more accurate and more fluent in word reading over time, (i.e., they increase in word decoding efficiency; Share, 1995, 1999, 2004).

From the start of word decoding development, children's word decoding efficiency increases over time (e.g., Verhoeven & van Leeuwe, 2009). This is especially the case in learning to read in transparent orthographies (e.g., Spanish; Baker, Park, & Baker, 2010). In transparent orthographies first graders generally master the alphabetic principle within six months of phonics-based reading instruction (Authors, 2017). With increasing word decoding skills, the connections between orthography (graphemes), phonology (phonemes), and semantics (word meanings) become augmented and more coherent (Plaut, McClelland, Seidenberg, & Patterson, 1996; Van Orden, Pennington, & Stone, 1990), and children learn how to efficiently store and retrieve written words from memory. According to Perfetti (2007), phonology, orthography, semantics, and the coherence between these components influences the specificity and quality of the mental representations. According to this Lexical Quality theory (Perfetti, 1992), a higher specificity of mental representations facilitates the ultimate next developmental step for the children: To comprehend written text.

Automation of word decoding skills has indeed been found to be essential for the development of reading comprehension skills (Perfetti, 1992; Stanovich, 2000). Therefore, once efficient word decoding skills have been obtained, this heralds the gradual development of reading comprehension skills. In the literature, the relation between word decoding skills and reading comprehension has been established in both adult and child readers (see the metanalysis of Garcia and Cain (2014); and the Simple View of Reading, Gough & Tunmer, 1986). Foorman (1997) found a strong correlation between English word decoding and reading comprehension during all years of elementary school, and word decoding in lower grades predicted reading comprehension in later grades. Foorman, Petscher, and Herrera (2018) showed that the contribution of decoding to English reading comprehension decreased across the primary grades, whereas instead contribution of language skills increased across the grades. The impact of language skills, in terms of vocabulary, syntax, and listening comprehension, on reading comprehension was found to be stable over time (Cutting & Scarborough, 2006).

1.2. Precursors of word decoding

Research indicated that reading development is predicted by individual variation in cognitive and linguistic pre-reading characteristics in children. Across orthographies, word decoding development has typically been found to be determined by kindergarten precursor measures of phonological awareness, grapheme-to-phoneme knowledge, rapid naming, vocabulary, and visual and verbal memory skills (Al Otaiba & Fuchs, 2002; Kirby, Desrochers, Roth, & Lai, 2008; Landerl et al., 2013; Melby-Lervåg, Lyster, & Hulme, 2012; Moll et al., 2014; Van den Boer, de Jong, & Haentjens-van Meeteren, 2013). The influence of these skills on later word decoding development has been suggested to be more or less universal, and was found in both deep (opaque) alphabetic orthographies such as English, and in shallow (transparent) orthographies such as Dutch (e.g., Caravolas et al., 2013), despite the fundamental differences in linguistic and orthographic complexity that can be found cross-linguistically (Seymour, Ari, & Erskine, 2003).

1.3. Precursors of reading comprehension

A large amount of studies revealed the influence of kindergarten skills on early word decoding development. In comparison, there is little research on the kindergarten influence on early reading comprehension. Furthermore, research primarily focused on the predictive value of lexical and semantic components, such as word reading, language comprehension, and vocabulary (Gough & Tunmer, 1986; Lerkkanen, Rasku-Puttonen, Aunola, & Nurmi, 2004; Oakhill & Cain, 2012), leaving other kindergarten skills out of consideration. Neuroimaging studies have also primarily focused on semantic components, and found additional evidence that these components were highly related to reading comprehension (e.g., see the ERP-study by Landi & Perfetti, 2007).

Consequently, the important role of semantic components in reading comprehension is well-established. To a lesser extent, also orthographical and phonological precursors have been related to reading comprehension (Perfetti, 1992; Perfetti, 2007; Richter, Isberner, Naumann, & Neeb, 2013). This finding of an additional contribution of phonology and orthography to reading comprehension, along with semantics, is in line with the triangle model of reading by Harm and Seidenberg (2004) and with Perfetti and Hart (2002). They stated that the source of individual variation in terms of reading comprehension is that readers vary in the full range of their lexical representations. In other words, semantic representations result from fully specified orthographic representations and redundant phonological information (Perfetti, 2007).

In addition to skills related to lexical quality, a review study by Kirby, Georgiou, Martinussen, and Parrila (2010) showed that, across orthographies, rapid naming was an independent predictor of reading comprehension. Furthermore, readers need good memory skills to be able to coherently grasp the meaning of texts. Memory skills have been shown to be independently predictive of reading comprehension skills (e.g., in both English and French; Cain, Oakhill, & Bryant, 2004; Seigneuric & Ehrlich, 2005; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000). The exact contributions of rapid naming and memory skills to the prediction of reading comprehension are still under debate. For example, Schatschneider, Harrell, and Buck (2007) and McCallum et al. (2006) did not find independent contributions of memory skills to reading comprehension in English orthography. Similarly, Cutting and Scarborough (2006) did not find independent contributions of memory and rapid naming in their prediction analyses of an English speaking sample, despite correlations of the skills with reading comprehension measures. It should be noted, however, that the contributions of rapid naming and memory could have been subsumed by word decoding and language proficiency in their models.

Adlof et al. (2010) studied the prediction of English reading comprehension in Grade 2 and Grade 8 and demonstrated that inclusion of kindergarten language related skills and nonverbal cognitive skills would add to the prediction of early reading comprehension problems. However, they did not assess visual or verbal memory skills in kindergarten, as possible predictors of later reading comprehension, whereas in other studies, a certain relation between memory skills and reading comprehension skills was found (Cain et al., 2004; Haarmann,
Davelaar, & Usher, 2003; Nouwens, Groen, & Verhoeven, 2017).

Leppänen, Aunola, Niemi, and Nurmi (2008) found a contribution of visual attention to reading comprehension in later grades. They also found a predictive contribution of phonological awareness to Grade 4 reading comprehension, although mediated by other reading measures in kindergarten and Grade 1. Although the kindergarten cognitive and linguistic assessment by Leppänen et al. was more comprehensive and wider in scope compared to most other reading comprehension studies, no early measures of rapid naming and memory skills were included in the kindergarten test battery. Therefore, no complete picture of kindergarten contributions to reading comprehension has been provided yet. What remains uncertain is whether and to what extent the wide range of kindergarten precursors co-act and contribute to the development of reading comprehension skills.

1.4. Measures of early reading comprehension

It is difficult to compare study outcomes for reading comprehension development, since different levels of reading comprehension have been studied. The general measurement intervals with focus on later reading comprehension may have overlooked the first fundamental development at the tipping point of learning to read for comprehension (Little, 2013). In previous research, reading comprehension and its predictors have primarily been assessed in later primary school grades. However, reading comprehension problems might already emerge during initial phases of reading development, and precursor contributions might vary over time (De Jong & van der Leij, 2002). Furthermore, most studies did not measure full reading comprehension during early grades. For example, the Danish study by Frost, Madsbjerg, Niedersse, Olofsson, and Möller Srensen (2005) did measure (pre-) kindergarten skills of language comprehension and phonological awareness in relation to Grade 9 full reading comprehension. In Grade 3, 4 and 6, they also measured aspects of reading, though restricted to sentence level comprehension. Sentence level comprehension skills measured different aspects of reading comprehension as compared to text-and discourse-level comprehension skills (higher order comprehension and inference making skills; Hogan, Bridges, Justice, & Cain, 2011; Silva & Cain, 2015).

In addition, other large-scale prediction studies of reading comprehension mainly assessed text- and discourse level reading comprehension in later grades, and hence discarded the early development of full reading comprehension. For example, the longitudinal Finnish study by Leppänen et al. (2008) indicated that letter knowledge assessed in kindergarten was a good predictor of reading comprehension from kindergarten for later L2 reading comprehension. The impact of kindergarten measures is highly comparable for L1 and L2 learners (Geva et al., 2006; Verhoeven & van Leeuwe, 2012), and associative semantic links between words have shown to be weaker in the L2 (Vermeer, 2001).

It has been found that the prediction of word decoding from kindergarten measures is highly comparable for L1 and L2 learners (Geva & Yaghoub-Zadeh, 2006). Phonological awareness and rapid naming have been found to be the best predictors of word decoding in both L1 and L2 learners (Durgunoglu, Nagy, & Hancin-Bhatt, 1993; Geva et al., 2000; Lesaux & Siegel, 2003).

1.5. Learning to read in a second language

Although the literature on second language learning is growing, reading development in L2 learners is still less comprehensively studied as compared to L1 reading development (Durgunoglu & Verhoeven, 2013). Therefore, it is not yet clear to what extent first language and second language word decoding acquisition and reading comprehension are comparable. Understanding individual variation in early word decoding development, reading comprehension, and the prediction from kindergarten skills is critical to determine the prerequisites for establishing L2 literacy. It has been found that L1 and L2 learners develop comparable word decoding skills already from the start of formal education (Lesaux & Siegel, 2003; Mancilla-Martinez & Lesaux, 2010; Melby-Lervåg and Lervåg, 2014; Verhoeven, 2000). Geva and Yaghoub-Zadeh (2006) even found word decoding efficiency to be higher in L2 learners.

In contrast to word decoding skills, L2 learners often face difficulties in developing reading comprehension skills (Babyajit, 2014; Burgoyne, Whiteley, & Hutchinson, 2011; Cain, Oakhill, & Bryant, 2000; Geva & Farnia, 2012; Melby-Lervåg & Lervåg, 2014). This has also been found in samples of L2 learners in Dutch (Cremers & Schoonen, 2013; Verhoeven & van Leeuwe, 2012). Nevertheless, some studies found similar reading comprehension skills as well as similar underlying skills for both L1 and L2 learners (Lipka & Siegel, 2012). Similar to what was established by the Simple View of Reading in L1 learners, word decoding skills are highly predictive of reading comprehension skills in L2 learners too (both accuracy and fluency; e.g., Baker, Park, & Baker, 2012; Geva & Farnia, 2012; Verhoeven & van Leeuwe, 2012; Saiegh-Haddad, 2003; Yaghoub-Zadeh, Farnia, & Geva, 2012).

1.6. Precursors of word decoding in L2

Previous research has indicated that L1 and L2 learners in elementary grades perform similarly on a wide range of basic cognitive and linguistic skills measuring precursors of lexical quality and of memory, such as nonverbal reasoning, rapid naming, phonological awareness, and short term memory (Geva et al., 2000; Lipka & Siegel, 2012; Raudszus et al., 2018). However, results of group comparisons might differ across stages of reading development. For example, Geva et al. found differences between L1 and L2 learners on rapid naming assessments in Grade 1, which disappeared in Grade 2. In addition, similar performances were generally only found in skills that did not involve any semantic components. Bilingual students often lag behind in semantic skills in their second language, as compared to their first language learning peers (Farnia & Geva, 2013; Raudszus et al., 2018; Verhoeven & van Leeuwe, 2012), and associative semantic links between words have shown to be weaker in the L2 (Vermeer, 2001).

It has been found that the prediction of word decoding from kindergarten measures is highly comparable for L1 and L2 learners (Geva & Yaghoub-Zadeh, 2006). Phonological awareness and rapid naming have been found to be the best predictors of word decoding in both L1 and L2 learners (Durgunoglu, Nagy, & Hancin-Bhatt, 1993; Geva et al., 2000; Lesaux & Siegel, 2003).

1.7. Precursors of reading comprehension in L2

Compared to the vast body of literature on the predictors for word decoding skills, little is known about the scope of predicting patterns from kindergarten for later L2 reading comprehension. The impact of language skills on reading comprehension established in L1 learners has also been found in studies of second language learners (Farnia & Geva, 2013; Geva & Farnia, 2012). It has been assumed that lexical and semantic skills such as vocabulary are related to reading comprehension, and group comparison studies mostly focused on the pronounced differences in these skills between first and second language readers (Droop & Verhoeven, 2003; Farnia & Geva, 2013; Kieffer & Lesaux, 2012; Verhoeven & van Leeuwe, 2012). Indeed, less developed semantic skills of L2 learners are associated with their lower reading comprehension skills (e.g., Babyajit, 2014; Burgoyne et al., 2011; Cain et al., 2006; Lervåg & Aukrust, 2010; Melby-Lervåg & Lervåg, 2014).

Although valuable, the prediction studies in L2 samples mainly focused on semantically related predictors, thereby overlooking the role of other kindergarten precursors that might be different for L2 learners as compared to L1 learners. Orthographic quality, phonological quality (Perfetti, 1992) and memory might be additionally relevant in predicting reading comprehension development (Yaghoub-Zadeh et al., 2012; also see the review by Genesee & Jared, 2008). It could be suggested that L2 learners compensate for lower semantically related skills by relying on better developed phonologically, orthographically, or memory related skills, which may, for example, be measured by
phonemic awareness, grapheme-to-phoneme knowledge, and verbal short term memory respectively. However, this has not been studied in a single design thus far.

1.8. Measures of reading in L2

Whereas a variety of first languages were involved in most studies on second language learners, English was the second language in most cases. English has an opaque orthography, which might modulate the developmental interactions and the impact of precursors measures (Share, 2008). Dutch has a transparent orthography (Seymour, Aro, & Erskine, 2003) and is therefore particularly suitable for studying the interrelation between word decoding, reading comprehension, and kindergarten precursors in first and second language learners.

To conclude, literacy studies vary widely with respect to the composition of predictor variables they have considered. In addition, prediction of early reading comprehension problems has been studied less than word decoding. It remains unclear how cognitive and linguistic kindergarten skills predict early word decoding development, early reading comprehension, and their integrated development throughout the initial elementary grades, and it remains debatable how the prediction and development of reading differ for first and second language learners. A more accurate and complete picture of how the development of word decoding and reading comprehension are interrelated and how the integrated development of the two can be predicted from cognitive and linguistic skills assessed before formal reading instruction is required.

1.9. The present study

In the present longitudinal study, we examined how early word decoding and reading comprehension build on cognitive and linguistic precursors from kindergarten in 613 first language (L1) and 109 second language (L2) learners of Dutch. First of all, it was examined how first and second language learners perform on a wide range of precursor measures of lexical quality, memory, and of early word decoding in Grades 1 and 2, and early reading comprehension in Grades 2 and 3. Second, the longitudinal prediction model of word decoding and reading comprehension in terms of L1 and L2 reading was analysed.

The two main research questions were:

1) To what extent do L2 learners of Dutch differ from L1 learners in kindergarten cognitive and linguistic skills, first and second grade word decoding development, and second and third grade reading comprehension development?

2) How can the integrated development of word decoding and reading comprehension in the early elementary grades be predicted in L1 and L2 learners from kindergarten measures of cognitive and linguistic skills?

Concerning the first research question, we expected L2 learners to differ from L1 learners in kindergarten skills concerning semantic components. Therefore, differences were expected in vocabulary (L2 learners scoring below L1 learners) and reading comprehension skills. It could be expected that L2 learners scored slightly below L1 learners on rapid naming of objects too, since this lexical retrieval task contains a semantic component. No differences were expected for skills tapping into orthography or phonology, so grapheme-to-phoneme knowledge and word decoding efficiency were expected to be similar across groups. Likewise, no differences between the groups were expected for memory skills and nonverbal reasoning.

With regard to the second research question, we expected that development of word decoding and reading comprehension in the early elementary grades would increase over time, and that both developmental paths would be highly autoregressive in nature. We expected that reading comprehension in Grades 2 and 3 was predicted by word decoding in Grades 1 and 2, and that both word decoding and reading comprehension could be predicted from kindergarten measures of grapheme-to-phoneme knowledge, rapid naming of objects, and memory skills. Vocabulary was expected to contribute to the prediction of reading comprehension, independent from the prediction by word decoding skills. Although differences in literacy performances were expected between L1 and L2 learners, it was hypothesized that developmental structural relations and predictive values would not be different between the two groups of learners.

2. Method

2.1. Participants

All children in this study (N = 722) were participants in a larger longitudinal study (N = 1006) on Dutch reading development, which started in 2013 in 37 schools throughout the Netherlands (Authors, 2017). All Dutch regions and both rural and urban areas were represented. The sample was treated in accordance with institutional guidelines and APA ethical standards and no outside approval by a governing board was required. The data collection was non-invasive, since it was based on regular educational practices, curricula, and methods in daily educational settings. Schools, parents, and children were informed about the purpose of the research, the expected durations of the tasks, and the procedures. They were informed about whom to contact for questions about the research. Schools gave active consent to participate in the longitudinal study. Prior to testing, informed passive consent was obtained from the parents of all participating children. Both schools and parents were aware of their right to decline participation and to withdraw from the research any time before or during the research project. After each academic year, the schools were asked if they were willing to maintain their participation. Schools were debriefed with the results and conclusions of the research.

For current analyses, five schools (90 children) were excluded from the total cohort, because they missed two or more of the measurement moments for reading comprehension. Of the 90 excluded children, sixteen children were L2 learners (18%), which is highly comparable to the 15% L2 learners in the remaining sample. Children in these five schools had similar scores on kindergarten measures as the children in the remaining sample (all independent sample t-tests p > .20). Furthermore, 35 children were excluded from the analyses, since they missed all reading comprehension measurement moments. Conclusively, analyses were conducted with a representative subsample of 722 Dutch children in 32 schools (379 boys; 343 girls) of which 109 (15.1%) L2 learners and 613 L1 learners.

The percentage of L2 learners is comparable with the percentage of immigrants in the Dutch society (Centraal Bureau voor de Statistiek, 2018; CBS). The second language learners in the current study came from the same schools and classrooms as the first language learners. The children came from a variety of ethnic and linguistic backgrounds. Our sample included > 20 languages. The predominant languages were Moroccan (Arabic and Berber; about 30%), and Turkish (about 30%). The others spoke a wide variety of languages including, among others, Polish, Somali, Spanish, and English. Children were operationally defined as second language learners if they spoke at least one language other than Dutch at home with their parents, siblings, and others who lived at home with them (e.g., grandparents). First language learners were defined as children who spoke exclusively one language at home, which was the same language as during instruction at school (Dutch). Language information was obtained through school records.

The L2 learners did not differ from the L1 learners on a raw score measure of nonverbal logical reasoning (Raven Coloured Matrices assessment; Raven, 1958) at the end of Grade 1, M_{L2 learners} (SD) = 27.62 (4.44); M_{L1 learners} (SD) = 27.66 (4.96). T-test for independent samples (two-sided, equal variances assumed) showed t (720) = −0.08, p = .93. The L2 learners, however, did differ from the L1 learners on
socio-economic status of the home environment (represented by the educational level of the first care giver), on a categorical scale of 1 (elementary school) to 4 (vocational education and university), M_{2.learners} (SD) = 2.22 (1.07); M_{1.learners} (SD) = 3.28 (0.67). T-test for independent samples (two-sided, equal variances not assumed) showed t (123.12) = 9.94, p < .001. Therefore, we controlled for socio-economic status in the comparative statistics.

All children attended regular classroom education and they all spoke Dutch at school. The children were first assessed at the end of the second year in kindergarten (M_{age} = 6;1, SD = 0;4). First grade reading instruction was comparable across the schools, since all schools were using the same highly structured, systematic and phonics based reading curriculum (Veilig Leren Lezen, “Learning to read safely”, Mommers et al., 2003).

2.2. Measures

2.2.1. Precursor measures

We administered cognitive and linguistic measures at the end of kindergarten. The tasks were designed and analysed for the purpose of the larger longitudinal study (Authors, 2017). All cognitive and linguistic measures, except for the Grapheme-Phoneme Knowledge task, were preceded by some practice items during which feedback was allowed. Furthermore, Phoneme Segmentation, Visual and Verbal Short Term Memory, and Vocabulary contained a cut off score to avoid further frustration if the performance level of a child was reached.

2.2.1.1. Phoneme Isolation. The child was asked to sound out the first phoneme of 10 orally presented monosyllabic CVC-structured words (e.g., muis, sop; mouse, soup). The score on this task was the amount of correct responses, with a maximum score of 10. Reliability of the task was good (Cronbach’s α = 0.83).

2.2.1.2. Phoneme Segmentation. The child was asked to serially pronounce each phoneme of an orally presented word. The 10 presented words increased in internal complexity, starting with CVC-structured words, and followed by CCVC- or CVCC-structured words and CCCVC- or CVCCC-structured words. The score on this task was the amount of correct total responses, with a maximum score of 10. The reliability of the task was good (Cronbach’s α = 0.85).

We expected a high interrelationship between initial phoneme isolation and phoneme segmentation, since both are assumed to measure the phonemic awareness of the participating children. Both tasks require one manipulation step of sounds in spoken stimuli (isolating or segmenting phonemes) and they are in the middle range of complexity. These tasks score high on criterion validity with reading acquisition as the criterion (Yopp, 1988). Principal Axis Factoring was conducted on initial phoneme isolation and phoneme segmentation. Indeed, one component with relatively high loadings was revealed (0.72 and 0.72). Therefore, construct validity of the component was indicated to be good. The component explained 51.83% of the variance. Adequacy of this analysis was low but acceptable (KMO = 0.50; Hutcheson & Sofroniou, 1999) and together with the strong theoretical hypothesis, it was decided to be accepted as one factor representing phonemic awareness. Factor scores were calculated by the regression method and used as one combined variable in path analyses of the current study.

2.2.1.3. Grapheme-phoneme knowledge. The child was asked to sound out 34 graphemes used in Dutch. Only the grapheme sound was considered correct in this task. If the child named the grapheme, the child was asked once to also give the sound of the grapheme. The score on this task was the amount of graphemes that were sounded out correctly. Reliability of the task was excellent (Cronbach’s α = 0.93).

2.2.1.4. Active vocabulary. The child was asked to complete a little sentence that was orally presented accompanied by a picture in a booklet. An example of a sentence was: “the man is …” with the correct answer: “fishing”. In total, 29 picture-word combinations were assessed. The task was based on the Vocabulary task in the Taaltoets Allochtone Kinderen (“Language test Ethnic Minority Children”; TAK; Verhoeven & Vermeer, 1986). The amount of correct words was the score on this task. The reliability was good (Cronbach’s α = 0.83).

2.2.1.5. Rapid naming (RAN). To measure lexical retrieval of objects, the child was asked to name visually presented objects as accurate and quickly as possible during 1 min. The task material consisted of a card with repeated rendering of five highly familiar pictures, which were practiced once before actual measurement. The five pictures corresponded with one-syllable, high frequent Dutch words (viz., saw, pot, thumb, trousers, tent). The amount of correct named pictures was the lexical retrieval score. Reliability of the task was excellent (Cronbach’s α = 0.95).

2.2.1.6. Verbal Short-Term Memory (STMver). To assess verbal short term memory, the child was asked to repeat orally presented pseudowords. The task consisted of 20 pseudowords increasing from 1 to 4 syllables. The entire words had to be repeated correctly to be considered correct. The score on this task was the amount of correct repetitions, with a maximum score of 20. The reliability of the task was good (Cronbach’s α = 0.77).

2.2.1.7. Visual Short-Term Memory (STMvis). We asked the child to remember and rebuild the order of a series of visual figures (viz., fish, cow, ship, chicken, sock) that was presented shortly by the test assistant. The amount of visual figures in a series increased from two to five figures to remember. The complete task consisted of 15 series. This task followed the task design of a sub task of the RAKIT-test kit, called “Visual Memory Span: Concrete Figure Sequences” (Pieters, Dek, & Kooij, 2013). This task contained concrete figures, which was specifically constructed for young children. Using abstract figures would be too complex for the participant age group. The entire series had to be remembered to be considered correct. The amount of correct series was the score on this task, with a maximum score of 15. The reliability of the Visual Short-Term Memory task was good (Cronbach’s α = 0.77).

2.2.1.8. Raven coloured progressive matrices. To measure non-verbal reasoning, the child was asked to make sense of items with increasing complexity. This test was chosen for its non-verbal character and suitability for use with 1:2 learners. The task material consisted of three sets of 12 puzzles with one piece left out. Six answer options of missing pieces were provided with each puzzle. The child was asked to pick the right answer option by ticking the right box on an answer sheet. The amount of correct answered pictures was the non-verbal reasoning score. Reliability of the task as conducted in the current study was measured in a representative sample of 1006 children, and was considered good (Cronbach’s α = 0.81). This resembles the standardized task, which has been shown to have good reliability (Cronbach’s α = 0.90; Van Bon, 1986).

2.2.2. Standardized word decoding

Biannually, children’s ability to decode words was assessed with a standardized test (Drie-minutentests; “Three-minute test”; Krom, Jongen, Verhelst, Kamphuis, & Kleinjens, 2010). The total task consisted of three word cards with varying internal complexities: the first card contained one-syllable simple-structured words; the second card contained words with one or two consonant clusters; the third card contained words with at least two syllables. Per card, we asked the child to accurately read as many words as possible during 1 min. The amount of correct read words per timed element was the score on a reading card. The composite score of the three card scores was the composite word decoding efficiency score. In line with the manual of the standardized test, the third card was not included halfway Grade 1 for complexity
reasons. The combined card scores were considered reliable, (Cronbach’s $\alpha = 0.97$; Krom et al., 2010).

2.2.3. Reading comprehension

From halfway Grade 2, reading comprehension was measured biannually by the standardized reading comprehension test of the CITO (Centraal Instituut Toets Ontwikkeling; Central Institute for Test Development;, 2015). The CITO material was used by all participating schools. Each test consists of several short texts followed by multiple choice questions addressing both literal information and inference making. The written texts used for the tests include formal explanatory reading texts, fiction, narratives, and literature sections. For each measurement moment, other sets of reading texts and test items were used to adapt to the required norm level of the children, and to avoid memory and learning effects caused by repetition of texts.

One test moment consisted of two sub-parts. All children first completed the first sub-part. The individual score on this first sub-part determined whether a child continued with an easier or more difficult version of the second sub-part. This two-step method was used in order to arrive at a precise indication of reading comprehension level (Rasch, 1960). Per test moment, one total score was calculated out of the two sub-parts. Final test results were compared between and within (viz., subsequently) individuals. Scores were calculated on one and the same scale, regardless of the items that have been administered for the individuals. Scores were standardized based on national norm scores of representative groups of Dutch children. For each measurement moment, the test was considered reliable (Cronbach’s $\alpha$'s respectively 0.86, 0.83, 0.84; CITO, 2015).

2.3. Procedure

In the Netherlands, kindergarten is a two-year program prior to first grade. No formal literacy instruction is provided in kindergarten. However, children are stimulated to playfully discover grapheme-phoneme correspondences and phonological awareness. We first assessed the cognitive and linguistic measures at the end of kindergarten. The cognitive and linguistic measures were administered by the first author and eight trained test assistants with Bachelor or Master degrees in Educational Science, Psychology or Linguistics. The test assistants were qualified for assessing behavioural tasks. All assistants were fully trained prior to any data collection. All tasks were administered individually in a quiet room at school during regular school hours. To test nonverbal reasoning skill of the children, Raven coloured progressive matrices was assessed at the end of Grade 1. The test was conducted with a paper and pencil task in classroom setting. Instruction was orally provided in classroom setting, and one item was practised in the group after instruction.

In Grade 1, children in the Netherlands start receiving formal reading instruction. Word decoding instruction was provided in general classroom setting following the daily reading curriculum. All participating schools made use of the same systematic incremental phonics reading curriculum, called Veilig Leren Lezen (“Learning to Read Safely”; Mommers et al., 2003). In Grade 2 and Grade 3, reading instruction gradually shifts attention towards explicit instruction on reading comprehension instead of decoding. However, training and practising on decoding efficiency continues. Standardized word decoding (WD) tasks were administered halfway (i.e., in February; word decoding middle of the year; WDM) and at the end (i.e., in June; word decoding end of the year; WDE) of Grade 1 and Grade 2 (i.e., four measurement moments, comprising WDM1, WDE1, WDM2, WDE2). In accordance with the guidelines of the standardized tests used, standardized word decoding measurements were assessed individually by certified and well instructed classroom teachers or internal remedial teachers of the participating schools.

Reading comprehension (RC) was administered in classroom setting halfway and at the end of Grade 2, and halfway Grade 3 (i.e., three measurement moments; comprising RCM2, RCE2, RCM3).

Fig 1 shows all measurement moments on a timeline of kindergarten to Grade 3.

2.4. Analytic approach

Differences on precursor measures between groups were analysed in multivariate analyses of variance, controlling for the possible moderation of SES. Nondirectional hypotheses, and therefore two-sided test statistics, were maintained for all analyses. Repeated measures of word decoding efficiency and reading comprehension of both groups were analysed with Repeated Measures Analysis of Variance, controlling for the possible moderation of SES. Analyses of Variance were conducted using SPSS (2015, IBM, SPSS 24).

LISREL path modelling (Jöreskog & Sörbom, 1996) was used for examining the structural relations in longitudinal development of the independent constructs of word decoding and reading comprehension (in simple path models) and their integrated development (in a cross-lagged panel model), determined by their predictors. In the cross-lagged panel model, the variables at one time point could theoretically share common ‘cause’ not explained by the specified predictors. Therefore, the residual variances of constructs measured at one time point were freely estimated. Variables later in time were expected not to influence variables earlier in time. After the conceptualised relations, plausible relations with relevant improvements of the $\chi^2$ ($MI \geq 3.84$ refers to significance improvement of the $\chi^2$ at $\alpha = 0.05$) have been considered using Modification Indices of LISREL. Only significant paths (at $\alpha < 0.05$) were preserved in the final models (see Little, 2013). The kindergarten precursors were allowed to correlate, and residual variances among word decoding and reading comprehension measures at one measurement moment were allowed to associate. The fit of the models was evaluated using a chi-square ($\chi^2$) test, and the relative chi-square ($\chi^2_{df}$), calculated as the ratio of the chi-square with the degrees of freedom. The relative chi-square should be lower than 3 to be considered good fit (Carmines & McIver, 1981). The Root Mean Square Error of Approximation (RMSEA) was additionally evaluated, because of its robustness for large sample influences. The critical value for RMSEA was set on < 0.06 to be considered good fit (Tabachnick & Fidell, 2014).
using LISREL. We first computed the models with an equality restriction for the complete model, meaning that all the coefficients in a model were assumed to be equal for the L1 and the L2 group. Then, we checked for each coefficient whether allowing the coefficient to vary between groups would significantly increase the model fit, using Modification Indices of both groups. If model fit did not increase, the constraint was retained in the final model. If free estimation of a coefficient did increase model fit, it was freely estimated in the final model. Differences between the fit of nested models were assessed by means of a $\chi^2$ difference test (Kline, 2011).

3. Results

3.1. Preliminary considerations

Prior to the analyses, the missing values in this longitudinal dataset were analysed. There were no missing values in the Kindergarten cognitive and linguistic measurements. In Grades 1 and 2, < 2% of the values were missing for the word decoding measures. Missing values in word decoding were considered missing completely at random (MCAR; Little's missing completely at random test for word decoding, values in word decoding were considered missing completely at random values were missing for the word decoding measurements. Missing nitive and linguistic measurements. In Grades 1 and 2, < 2% of the were analysed. There were no missing values in the Kindergarten cog-

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Subsample</th>
<th>M (SD)</th>
<th>Group comparison</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid naming</td>
<td>L1-learner (n = 613)</td>
<td>40.81 (8.43)</td>
<td>6.84</td>
<td>.009</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner (n = 109)</td>
<td>36.50 (11.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial phoneme isolation</td>
<td>L1-learner</td>
<td>8.59 (1.82)</td>
<td>0.01</td>
<td>920</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>8.03 (2.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoneme segmentation</td>
<td>L1-learner</td>
<td>4.83 (2.60)</td>
<td>5.15</td>
<td>.024</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>3.34 (2.47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grapheme-to-phoneme Knowledge</td>
<td>L1-learner</td>
<td>19.61 (7.44)</td>
<td>0.15</td>
<td>.697</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>17.97 (7.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>L1-learner</td>
<td>14.73 (3.74)</td>
<td>91.52</td>
<td>&lt; .001</td>
<td>.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>7.60 (4.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short term memory verbal</td>
<td>L1-learner</td>
<td>15.35 (3.17)</td>
<td>3.78</td>
<td>.052</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>13.84 (3.72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short term memory visual</td>
<td>L1-learner</td>
<td>8.41 (2.95)</td>
<td>0.012</td>
<td>.913</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>8.12 (2.75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

**p < .01.

***p < .001.

3.2. Differences between L2 learners and L1 learners in literacy measures

Our first research questions addressed the extent to which L2 learners differed from L1 learners in kindergarten cognitive and linguistic skills, word decoding development, and reading comprehension development.

We first investigated whether L2 learners and L1 learners scored differently on kindergarten measures of rapid naming, initial phoneme isolation, phoneme segmentation, grapheme-to-phoneme knowledge, vocabulary, and visual and verbal short term memory (see Table 1 for descriptive statistics). A MANOVA was used, controlling for the possible moderation effect of the socioeconomic status of the home environment (SES, operationalized as the educational level of the main caregiver). Although quite robust analyses were used, outcomes should be interpreted with some caution, since equal covariances could not be assumed (Box’s test $p < .001$). Using Wilk’s Lambda, a main effect of Group on the kindergarten measures was found, $F (7, 708) = 13.70$, $p < .001$, $\eta^2 = .119$. No main effect of SES was found, $F (21, 2033.54) = 1.513$, $p = .063$, $\eta^2 = .015$ and there was no interaction of Group $\times$ SES, $F (21, 2033.54) = 0.635$, $p = .896$, $\eta^2 = .006$. This means that the effect of Group (L1 or L2 learner) on the kindergarten measures was not influenced by the SES of the home environment. To further specify the exact differences per kindergarten measure, separate univariate ANOVAs were conducted (see Table 1). The L2 learners scored below the L1 learners on rapid naming, phoneme segmentation, and vocabulary. In line with the conclusions from the overall multivariate analyses, no interaction effects were found for the univariate tests. This means that the group differences found for rapid naming, phoneme
segmentation and vocabulary were not moderated by the educational level of the main caregiver (SES).

To analyse group differences in word decoding development, we conducted a Repeated Measures Analysis of Variance with Time (WDM1-WDE2) as a within-subjects factor and Group (L1 learners, L2 learners) as a between-subjects factor. We again controlled for the possible moderation effect of SES. Table 2 shows the means and standard deviations of the standardized word decoding measurements. Mauchly’s test of sphericity was violated; therefore, multivariate test results (Wilks’Lambda) were reported. This was appropriate, since the standard deviations of the standardized word decoding measurements.

Table 2 shows the means and standard deviations of word decoding efficiency and reading comprehension.

### Table 2

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>WDM Grade 1</td>
<td>L1-learner (n = 613)</td>
<td>52.56</td>
</tr>
<tr>
<td></td>
<td>L2-learner (n = 109)</td>
<td>54.65</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>52.87</td>
</tr>
<tr>
<td>WDE Grade 1</td>
<td>L1-learner</td>
<td>113.63</td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>114.39</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>113.75</td>
</tr>
<tr>
<td>WDM Grade 2</td>
<td>L1-learner</td>
<td>166.66</td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>169.17</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>167.04</td>
</tr>
<tr>
<td>WDE Grade 2</td>
<td>L1-learner</td>
<td>190.86</td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>195.69</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>191.59</td>
</tr>
<tr>
<td>RCM Grade 2</td>
<td>L1-learner</td>
<td>133.45</td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>118.59</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>131.21</td>
</tr>
<tr>
<td>RCE Grade 2</td>
<td>L1-learner</td>
<td>140.47</td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>126.60</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>138.38</td>
</tr>
<tr>
<td>RCM Grade 3</td>
<td>L1-learner</td>
<td>154.93</td>
</tr>
<tr>
<td></td>
<td>L2-learner</td>
<td>139.57</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>152.61</td>
</tr>
</tbody>
</table>

Note. WDM = Word Decoding Middle of the year; WDE = Word Decoding End of the year; RCM = Reading Comprehension Middle of the year; RCE = Reading Comprehension End of the year.

3.3. Prediction of word decoding and reading comprehension development

The second research question concerned the integrated development of word decoding and reading comprehension, and the predictive values of kindergarten measures of cognitive and linguistic skills. To address this research question, first, the autoregressive developmental path of word decoding in Grade 1 and Grade 2 was evaluated (Model 1a). Second, the predictive values of the kindergarten measures on word decoding were analysed (Model 1b). Subsequently, the developmental path of reading comprehension in Grade 2 and Grade 3 was analysed (Model 2a), and the models of word decoding and reading comprehension were combined and integrated (Model 2b). It was then studied how the integrated model was determined by kindergarten precursors, resulting in one integrated model of early literacy development from kindergarten precursors up to reading comprehension in Grade 3 (Model 2c). All models were analysed using LISREL path modelling statistics following the building steps from a conceptual approach section. Furthermore, it was questioned how the developmental models differed between the groups. For each model, multiple group comparison statistics were used to study the developmental differences between L1 learners and L2 learners in depth (the procedure used is explained in the analytical approach section).

In Table 3, both L2 learners’ (upper right diagonal) and L1 learners’ (lower left diagonal) correlations between precursor measures, word decoding efficiency and reading comprehension were presented. Both groups showed that kindergarten measures were correlated moderately, indicating that related but independent skills were measured. The kindergarten measures were correlated with the outcome measures of both reading comprehension and word decoding. Vocabulary was not correlated with the word decoding measures, but showed significant correlations with reading comprehension. Correlations between grapheme-phoneme knowledge and word decoding seemed higher for L2 learners compared to those for L1 learners. In contrast to the L1 learners, no correlation was found between verbal short term memory and word decoding for L2 learners. Word decoding and reading comprehension were correlated to a similar extent in both groups. No remarkable differences between groups were found in the within construct correlations of the repeated measures of the outcome variables, although the within construct relations of reading comprehension tend to be stronger in the L2 learner group. Word decoding and reading comprehension were correlated in both groups.

In a simple path model, the autoregressive relations of word decoding were specified from halfway Grade 1 towards the end of Grade 2. Standardized coefficients of the model are presented in Fig. 2. The standardized coefficients can be considered autoregressive stability coefficients, showing high autoregression during word decoding development. No within group modification indices were suggested, indicating that there were no differences between the L2 learners and the L1 learners in the developmental word decoding path. The model fits the data adequately, $\chi^2(3) = 9.97$, $p = .030$, RMSEA = 0.074. The $\chi^2$ was determined for 19.11% by the L1 group and 80.98% by the L2 group.

To analyse the prediction of word decoding development by kindergarten measures of cognitive and linguistic skills, the kindergarten measures were added to the simplex model. These kindergarten measures were allowed to covary. First, the direct paths towards Grade 1 word decoding were estimated. Second, modification indices in LISREL suggested direct paths from rapid naming to word decoding measures at the end of Grade 1 and halfway Grade 2 over and above the indirect contribution via the measurement moment halfway Grade 1. Phonemic awareness and active vocabulary had no independent contributions to the prediction model. They were excluded in the final model, since only standard deviations of the standardized reading comprehension measurements.
significantly if the predictive path of grapheme-phoneme knowledge to word decoding was allowed to vary between the groups. The model showed that the contribution of grapheme-to-phoneme knowledge on word decoding was stronger for the L2 learners as compared to the L1 learners. Except from the path for grapheme-phoneme knowledge, all other kindergarten predictive paths could be considered similar across groups. The model fitted the data very well, $\chi^2 (46) = 56.25, p = .143$, RMSEA = 0.025. The chi-square was determined for 62.55% by the L1 group and for 37.45% by the L2 group. See Fig. 3 for the resulting model.

The autoregressive relations of reading comprehension of halfway Grade 2 towards halfway Grade 3 were specified in a simplex path model. Modification indices suggested a direct contribution from halfway Grade 2 to Grade 3, in addition to the structural relation from the end of Grade 2 to Grade 3. To test for group differences in this simplex reading comprehension model, we first computed a model in which all coefficients were constrained between groups. Next, we followed the building steps as elaborated on in the analytic approach section. The model fit significantly increased if the coefficient from the end of Grade 2 to halfway Grade 3 was allowed to vary between the groups. For that specific structural relation, the model showed a stronger autoregressive effect for the L2 learner group as compared to the L1 learner group. No further within-group modifications were suggested, showing that further coefficients in the model could be considered similar across the groups. The model fits to the data very well, $\chi^2 (20) = 68.8, p = .644$, RMSEA $< .001$ (see Fig. 4). The $\chi^2$ was determined for 13.95% by the L1 group and 86.05% by the L2 group.

In an integrated model, the word decoding development was added to investigate reading comprehension development in Grade 2 and Grade 3 in interaction with word decoding in Grade 1 and Grade 2 (see Fig. 5). In a first step, the concurrent measures between the constructs in Grade 2 were allowed to covary. This covariation was not significant at the end of Grade 2, so only the covariation halfway Grade 2 retained. In the second step, the first order cross-lagged relationships between the constructs were analysed. Only lags forward in time were considered. The model showed a strong cross-lagged relation from word decoding at the end of Grade 1 to reading comprehension halfway Grade 2, indicating a strong contribution of word decoding skills to reading comprehension skills. Additionally, a reciprocal cross-lagged relationship was found between word decoding and reading comprehension from halfway Grade 2 to the end of Grade 2. Although the subsequent cross-lagged effects contributed significantly to the complete model, they were of a small magnitude. No further cross-lagged relationships were found. To test for group differences, we first computed a model in which all coefficients were constrained between groups. Next, we followed the building steps as elaborated on in the analytic approach section. The model fit significantly increased if the coefficient from reading comprehension end Grade 2 to halfway Grade 3 was allowed to vary between the groups. That means that the autoregressive effect from reading comprehension at the end of Grade 2 to halfway Grade 3 was stronger for L2 learners as compared to the L1 learner group. The fit of the integrated model was adequate, $\chi^2 (29) = 82.44, p < .001$, RMSEA = 0.072.
In a final step, kindergarten precursors were added to the integrated literacy model. Word decoding development was best predicted by grapheme-phoneme knowledge and (developmentally) by rapid naming, and to a lesser extent by visual and verbal short term memory. Reading comprehension was best predicted by active vocabulary, visual short term memory, and phonemic awareness. At the end of Grade 2, there was a small independent contribution of verbal short term memory and vocabulary. The group comparison model showed that the contribution of grapheme-to-phoneme knowledge on word decoding was stronger for the L2 learners as compared to the L1 learners.

The resulting model revealed quite similar developmental paths for both groups. Except from the grapheme-phoneme knowledge, all other kindergarten predictive paths could be considered similar between the groups. Except from the autoregression coefficient in Grade 3, all autoregression coefficients and all cross-legged paths between word decoding and reading comprehension could be considered similar between the groups. The model fits to the data adequately, $\chi^2(132) = 181.81, p = .003, \text{RMSEA} = 0.033$. The $\chi^2$ was determined for 50.33% by the L1 group and 49.67% by the L2 group. See Fig. 6 for the resulting model.

4. Discussion

4.1. Summary of findings

The present longitudinal study in L1 and L2 learners of Dutch investigated cognitive and linguistic precursors of early word decoding and reading comprehension from kindergarten to Grade 3. In kindergarten, second language learners scored below first language learners on active vocabulary, rapid naming of objects, and phoneme segmentation. No differences were found in grapheme-to-phoneme knowledge, phoneme isolation, nonverbal reasoning, and memory skills. Whereas the L1 and L2 groups scored similarly on word decoding across the different grades, L2 learners lagged behind in reading comprehension at all measurement moments. The discrepancy for reading comprehension prevailed over time, but did not increase. Strong autoregressive effects were found for word decoding and reading comprehension in L1 as well as L2 learners, and word decoding was a strong predictor of reading comprehension for both groups. Kindergarten cognitive and linguistic skills impacted word decoding and reading comprehension similarly in both groups, except that the prediction of word decoding development by grapheme-to-phoneme knowledge was stronger for L2 learners as compared to L1 learners.

4.2. Differences between L1 and L2 learners

Our first research question concerned to what extent L2 learners differed from L1 learners in kindergarten cognitive and linguistic skills, early word decoding development, and early reading comprehension. With regard to the comparisons of the full set of kindergarten characteristics, no differences were found for tasks measuring grapheme-to-phoneme knowledge, initial phoneme isolation, short-term memory or nonverbal reasoning. This was as expected from previous literature (e.g., Geva et al., 2000), and has now been confirmed in a representative large scaled sample of the primary school population. However, L2 learners scored below L1 learners on tasks measuring rapid naming, phoneme segmentation, and vocabulary. This is consistent with the hypothesis that L2 learners do not differ from L1 learners on grapheme-to-phoneme knowledge and memory skills, but do lag behind in vocabulary and rapid naming of familiar objects. Vocabulary and rapid naming profit from well-developed language skills (Babayigit, 2014; Burgoyne et al., 2011). Although similar group levels for initial phoneme isolation skills were found, results indicated lower levels for L2 learners on the relatively more complex phonological awareness task of phoneme segmentation. This indicates that L2 learners are behind in developing more complex phonological awareness at the end of kindergarten. It could be speculated that these results strengthen the suggestion of Verhoeven (2000) that incomplete auditory discrimination of phonemes in the second language might hamper the correct segmentation of orally provided words. More generally, this segmentation task requires readers to manipulate and remember phonemes in existing Dutch words, so L1 learners might have an advantage via their higher general vocabulary level (Lesaux & Siegel, 2003). With respect to the comparison of word decoding development, the results indicated that L2 learners did not differ from L1 learners on word decoding performance, and developmental structural relations were similar. This is in line with previous studies showing that L2 learners perform similarly or even better than L1 learners on word decoding (e.g., Babayigit, 2014; Raudszus et al., 2018; Verhoeven & van Leeuwe, 2012). In contrast to the similar word decoding skills, the current data revealed a performance gap between L1 and L2 learners in terms of reading comprehension skills, which was already apparent in the earliest stages of reading development. Lower reading comprehension skills in L2 learners were found in several other studies (see the meta-analysis by Melby-Lervåg & Lervåg, 2014), but was now established during the very early phases of reading comprehension development. To sum up, L2 learners lag behind in kindergarten rapid naming, phoneme segmentation, and vocabulary skills, as well as in reading comprehension in Grades 2 and 3 but not in word decoding in Grades 1 and 2. The delays in kindergarten precursors does not seem to impact their early word decoding development, but does impact reading comprehension development in later grades.

4.3. Development of word decoding and reading comprehension

Our second research question addressed how early word decoding and early reading comprehension developed, and how both were determined by kindergarten cognitive and linguistic skills. Contributions were compared between L1 and L2 learners. Results on the early word decoding development indicated strong autoregressive effects, which was similar for both groups. High stability in early word decoding development was also found in previous studies (Authors, 2017; Verhoeven & van Leeuwe, 2008) and was now also found explicitly for L2 learners. Furthermore, development of reading comprehension was
also largely stable over time. Good reading comprehenders at the start of Grade 2 are likely to remain good reading comprehenders over time. High stability in reading comprehension was also found by, for example, Lerkkanen et al. (2004), and De Jong and van der Leij (2002). This stability was now found for both L1 and L2 learners. In the transfer from Grade 2 to Grade 3, stability was even higher for L2 as compared to L1 learners. High stability in L2 learners was found in previous studies too (Droop & Verhoeven, 2003; Netten, Droop, & Verhoeven, 2011) but was not yet established in initial phases of development. The current findings reveal that the stable reading comprehension development of L2 learners is similar to the stability of L1 learners already from the very beginning of reading comprehension. The finding shows that from the very beginning, the predictability of the subsequent reading comprehension performances by autoregression is similar for both L1 and L2 learners.

The integrated model of word decoding and reading comprehension describes the strong developmental prediction of reading comprehension by word decoding in Grade 1. This prediction was as expected from the literature on the Simple View of Reading (Gough & Tunmer, 1986), and has now been found to emerge from a developmental perspective throughout the initial primary grades. The integrated path model was similar for both L1 and L2 learners, indicating that groups did not differ in the predictability of reading comprehension by word decoding, despite the fact that L2 learners were delayed in reading comprehension skills. The current findings expand the Simple View of Reading (Gough & Tunmer, 1986) by showing that the influence of word decoding on reading comprehension is developmental in nature during the early grades of elementary school, and that this prediction path is similar in L1 and L2 learners.

4.4. Precursors of word decoding

We further investigated how the integrated model of word decoding and reading comprehension was determined by kindergarten cognitive and linguistic skills. For both groups, word decoding was best predicted by grapheme-to-phoneme knowledge, rapid naming of objects, and short-term memory skills. These predictors have also been established in previous research on predictors of word decoding development (e.g., Al Otaiba & Fuchs, 2002; Authors, 2017; Kirby et al., 2008; Landerl et al., 2013; Melby-Lervåg et al., 2012; Moll et al., 2014) and were now shown to be similar in terms of first and second language reading. Over and above the direct contribution of rapid naming to the prediction of word decoding halfway Grade 1 and the indirect contribution through the longitudinal model, additional direct contributions of rapid naming to word decoding at the end of Grade 1 and halfway Grade 2 were found. This suggests that rapid naming continues to contribute to word decoding development across time. Although similar prediction patterns occur, the prediction of word decoding development by grapheme-to-phoneme knowledge was stronger for L2 learners as compared to L1 learners. This might relate to the finding of Segalowitz, Segalowitz, and Wood (1998) that automatic orthographic processes are later to develop in L2 learners. Therefore, the nonlexical route of reading (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) might be used in addition to the lexical route for a longer period during early development. Whereas L1 learners use the lexical (orthographic) route more, for L2 learners the support by good phonological recoding skills (using grapheme knowledge) might be more needed.

4.5. Precursors of reading comprehension

Regarding the kindergarten prediction of reading comprehension in Grades 2 and 3, vocabulary, phoneme segmentation, and short term memory skills contributed to reading comprehension development, over and above the strong developmental prediction by word decoding. The current finding of phoneme segmentation as a direct predictor of later reading comprehension is in line with previous findings that phonological awareness in kindergarten was an independent predictor of reading comprehension (Parrila, Kirby, & McQuarrie, 2004). However, not all studies found a relation between phonological awareness and reading comprehension (e.g., Oakhill & Cain, 2012). Further research on the prediction of reading comprehension by phonological awareness skills is warranted. Additional longitudinal contributions of visual short term memory and vocabulary to reading comprehension at the end of Grade 2 were found, over and above the indirect contribution that was established for the first measurement moment halfway Grade 2. These additional longitudinal contributions could be interpreted as developmental prediction effects, although it should be mentioned that the effects were small. No independent direct contribution of rapid naming to the prediction of reading comprehension was found, whereas rapid naming was found to be highly predictive of word decoding. Possibly, the speeded nature of RAN and word decoding efficiency explain this correlation (Kirby et al., 2008).
The predictive patterns from kindergarten to reading comprehension in later grades were similar for both groups, despite the fact that L2 learners scored lower than L1 learners on oral language skills and reading comprehension. Therefore, findings of the current fine grained longitudinal design complement similar findings in the literature (Babajigjig, 2014; Burgoine et al., 2011; Cain et al., 2000; Lerrkanen et al., 2004; Melby-Lervåg & Lervåg, 2014; Lervåg & Aukrust, 2010).

4.6. Limitations and future research

The following limitations should be taken into account, and addressed in future research on this topic. In addition to previous studies in the field, the early phases of reading comprehension were examined and related to the full range of kindergarten precursors of lexical quality and of memory skills, and to word decoding skills in Grades 1 and 2, in a truly longitudinal and fine-grained design. However, although the current focus was on studying the full range of precursor skills related to lexical quality and memory skills, the current study could have been enriched by broader (more constructs) as well as more in depth (more aspects within one construct) measures. For example, the battery could have been expanded by precursor measures of higher order skills such as grammatical skills, giving a more complete picture of children’s level of lexical quality. Also, although orthographic elements were tested with the grapheme-to-phoneme task, a task measuring spelling skills would have enriched the insights in orthographic skills. In addition, according to Perfetti (2007), the connection of the lexical quality with reading comprehension also requires a certain quality of knowledge about word forms (i.e., grammar). For example, Botting, Simkin, and Conti-Ramsden (2006) found grammatical skills to be predictive of later reading comprehension of children with specific language impairments. It could be reasoned that inclusion of a measure of grammar skills would further explain variance in the current prediction models. Relevant to the current study, also, a more comprehensive insight in oral language proficiency (e.g., also including expressive syntax and listening comprehension) would have increased the impact of our conclusions (e.g., LARRC, 2015). For more in depth measuring of the constructs used, it would have been useful to create latent variables consisting of multiple skills. A combined (standardized) measure of receptive and expressive vocabulary would have comprised a more robust measure of the construct. Nevertheless, the current study offers an active vocabulary measure with a high reliability score to provide insight in the vocabulary of the children in the end of kindergarten. For practical and ethical reasons (viz., concise test battery) of the larger longitudinal study of which the current study is a part, oral language was not measured more extensively.

Second, it should be noted that the current study focused on prediction from kindergarten cognitive and linguistic skills measured at one moment in time. However, longitudinal (repeated) data of the cognitive and linguistic skills might have been contributive. More specifically, to give a developmental perspective of the Simple View of Reading (Gough & Tunmer, 1986), it would be interesting to see the prediction of reading comprehension by word decoding over time combined with the prediction by vocabulary over time.

Third, the study was conducted using a representative sample of the Dutch school going population, both in size and composition. Although representative, it is valuable to be aware of the disadvantages of such a large scaled heterogeneous longitudinal study. We especially focused on L2 learners, who have been labeled in the literature to be at risk for later reading comprehension problems. With respect to the subsample of second language learners, it was not always clear how frequent and in what contexts first or second languages were used in individuals. In addition, background languages and home language environments were widely variable. It would be useful to investigate the influence of the frequency and quality of language input in the home environment on reading development. In addition, with regard to the variety of first languages in the current sample, it should be noticed that characteristics of the specific orthographies might have their influence on the development of literacy skills in the second language of Dutch (Koda, 1996, but also see Akamatsu, 2002). For practical values and for transfer of the findings towards interventions in the daily classroom, this heterogeneous subsample was applicable, since the heterogeneity is a good reflection of Dutch multi-cultural society.

Fourth, the current study was conducted in the context of Dutch orthography. Although this transparent orthography seems particularly suitable for studying the interrelation between word decoding, reading comprehension, and kindergarten precursors in first and second language learners, it could be hypothesized that the characteristics of this specific orthography have impacted the developmental trajectories of learning to read as well as the transfer between first and second language learning. Future research should be conducted to see whether these results replicate in other (more opaque) orthographies too.

4.7. Implications for practice

The results of the current study have some clear practical implications. Many schools already use kindergarten screening of phonological awareness and grapheme knowledge to arrive at an early identification of children at risk for later word reading problems. Indeed, these precursors have been evidenced to be predictive of word decoding development. However, the current study indicated that it cannot simply be assumed that the same precursors are the best predictors for reading comprehension. Kindergarten literacy skills including the full range of precursors of lexical quality and of memory skills were evidenced to add to the identification of children at risk for literacy problems. Specifically, vocabulary, memory skills, and word decoding skills are the best predictors for later reading comprehension problems, and turn out to be valuable for insight into later reading development. Reading comprehension two years later was predicted by kindergarten cognitive and linguistic skills, even over and above the contribution of word decoding skills. Along with a growing body of research findings, we therefore advise special attention to children who lag behind in kindergarten cognitive and linguistic skills such as vocabulary, since it turns out to be difficult to catch up later during literacy development (Verhoeven, 2000). Especially, we emphasize the need for early enrichment for L2 learners to reduce or prevent later reading comprehension difficulties in L2 learners. Early enrichment could involve Early Childhood Education and Care (ECEC; Leseman & Slot, 2014; OECD, 2018), which provides L2 learners with specific opportunities to encounter, use, and practice the second language. This is enhanced by storybook reading, rich and meaningful play contexts using oral language, and implicit and explicit cognitive and linguistically-based learning experiences. Also, if possible, parents should be involved in the learning experiences in the ECEC program, to stimulate further practice of language skills at home (e.g., storybook reading). The impact of the delay in L2 learners in terms of kindergarten vocabulary, rapid naming, and phoneme segmentation might be overlooked, since this delay does not immediately impact the early word decoding development in Grade 1. However, results of the current study clearly show that it does impact reading comprehension development in later grades. This, in turn, might also affect general academic performances, suggested by the secondary school Programme for Internal Student Assessment (PISA), in which L2 learners still tend to achieve below the national average on a variety of academic domains (Stanat & Chistensen, 2006). The finding of similar stability in autoregressive effects for both L1 and L2 reading comprehension, implies that similar approaches may be used for L2 learners in terms of reading comprehension instruction. Rather, there should be early differentiation in instruction based on actual performances on reading comprehension in general. Poor comprehenders in Grade 2 should be signaled early and immediately get extra support, to avoid lagging behind in later grades. The finding of interrelatedness between word decoding and reading comprehension in both L1 and L2 learners, implies that both classroom instruction and individual
intervention should emphasize the integration of skills that are involved in word decoding and reading comprehension.

4.8. Conclusion

To conclude, the current study highlights the comparable developmental trajectories of early word decoding and reading comprehension in first and second language learners. The groups perform similar on kindergarten grapheme-to-phoneme knowledge, phoneme isolation, nonverbal reasoning, memory skills, and on later word decoding development. However, the groups do differ in semantically related kindergarten skills of RAN, phonological awareness, and vocabulary, and on reading comprehension scores. Both groups build upon kindergarten predictors for both word decoding and reading comprehension.

References


Authors. (2017).


Geva, E., & Farnia, P. (2012). Developmental changes in the nature of language profi-


The case of Arabic/Hebrew (L1); English (L2) readers. Reading and Writing, 16(8), 717–736. https://doi.org/10.1007/s11145-010-9220-9.