Journal of Research in Reading, ISSN 0141-0423
Volume 43, Issue 3, 2020, pp 290–308
DOI:10.1111/1467-9817.12304

First and second language vocabulary affect early second language reading comprehension development

Liza J. van den Bosch

Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

Eliane Segers

Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

Ludo Verhoeven

Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

The aim of this study was to examine variation in early reading comprehension development for second language (L2) readers compared with first language (L1) readers and to investigate the impact of vocabulary knowledge in their first and second language. Participants were 75 Dutch monolingual children (L1 readers) and 71 Turkish–Dutch bilingual children (L2 readers), aged between 6 and 8 years old at the start of the study. In a longitudinal design, three waves of data were collected across second and third grades. The L2 readers had lower reading comprehension scores than the L1 readers on average, 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 compared with L1 readers. The L2 subgroup with high vocabulary in both languages even caught up with their monolingual peers in third grade. 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 and education.

Keywords: reading comprehension, early literacy, bilingualism, second language, vocabulary

Highlights

What is already known about this topic?

- Children learning to read in their second language (L2 readers) generally fall behind their monolingual peers (L1 readers) with respect to reading comprehension.
- Vocabulary knowledge is important for reading comprehension performance and development in monolingual as well as bilingual children.

What this paper adds

- This is one of the few longitudinal studies focusing on L2 reading comprehension development in the early stages and assessing children's vocabulary knowledge in both their first and second language.
- The reading comprehension performance gap between L1 readers and L2 readers decreased across the lower primary grades; larger growth evidenced for L2 readers.
- Among the 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.

Implications for theory, policy or practices

- Bilingual children's vocabulary in both L1 and L2 can be considered the key factor to close the reading comprehension performance gap between L1 and L2 readers.
- For bilingual children with an immigrant background, it is important to promote their language proficiency in general not by definition in the language of instruction.
- To obtain a complete picture of their full competence, L2 readers' first language proficiency should be taken into account in research as well as in education.

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 L1 vocabulary knowledge 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 (Gottardo & Mueller, 2009; Lervåg & Aukrust, 2010), also L1 vocabulary knowledge may be of influence depending on children's L1 proficiency (Cummins, 1979). 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 (van den Bosch, Segers, & Verhoeven, 2019) 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 with monolingual children (L1 readers).

Reading comprehension

The importance of reading comprehension is widely acknowledged, and multiple theories on reading comprehension exist in the literature. The Simple View of Reading (Gough & Tunmer, 1986) 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, McMaster, Van den Broek, Kendeou, & Rapp, 2009).

The Lexical Quality Hypothesis strongly emphasises the role of vocabulary knowledge (Perfetti, 2007), stating 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 emphasised in the Reading Systems Framework (Perfetti & Stafura, 2014), which aims to capture the complexity of reading comprehension in an overarching theory.

Linguistic diversity

Due to migration and internationalisation, 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. Previous studies have shown that L2 readers do not differ from L1 readers in terms of decoding skills (Melby-Lervåg & Lervåg, 2014), whereas other studies have found an advantage for L2 readers (Bialystok, Luk, & Kwan, 2005; Droop & Verhoeven, 2003). Yet, 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 in reading comprehension 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, Maluch, Rjosk, Stanat, & Merkens, 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 with 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 grades 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 using a comparative rather than an additive 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 with L1 readers. For the present study, we collected longitudinal data for Dutch monolingual children (L1 readers) and Turkish–Dutch bilingual children (L2 readers) across the second and third grades 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 with L1 readers?
2 To what extent can we explain variation in early reading comprehension development among L2 readers (as compared with L1 readers) on the basis of individual differences in L1 and L2 vocabulary knowledge?

First, we expected that a performance gap in reading comprehension would exist across the second and third grades; 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 versus low levels of vocabulary knowledge in both their first and second language, which 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

A total of 13 primary schools located in lower-middle class neighbourhoods in urban areas in the Netherlands participated in the current study. 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 maximise the homogeneity of the L2 group, because 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 towards 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. Throughout the 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% male) and 71 L2 readers (60.6% male). 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 reported not to read in Turkish (77.5%).

As background variables, we compared the two language groups on nonverbal reasoning skills and socio-economic status (SES) (Table 1). The L1 readers and L2 readers did not differ in nonverbal 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 4-point scale (1 = primary education, 2 = lower secondary/vocational education, 3 = higher secondary education/intermediate vocational education, and 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 grades with an interval of approximately 6 months: (T1) start second grade, (T2) end second grade and (T3) mid third grade. Time (T1, T2 and 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 and T3), Dutch vocabulary (T1), Turkish vocabulary (T1) and language exposure at home (T1).

Table 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[a]	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[a]	137.9 (26.8)	84–210	115.3 (21.5)	70–190	5.68	<.001	0.93
T2[a]	141.9 (23.0)	97–201	121.2 (20.6)	78–176	5.71	<.001	0.95
T3[a]	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[a]	2.8 (0.7)	1.5–4.0	2.3 (0.9)	1.0–4.0	3.41	.001	0.62

Notes: DEC WD = decoding words, DEC PS = decoding pseudo words, VOC NL = Dutch vocabulary, VOC TU = Turkish vocabulary, RC = reading comprehension, NVR = nonverbal reasoning, SES = socio-economic status.
[a]Mean difference is significant based on critical *p*-values from linear step-up procedure (Benjamini & Hochberg, 1995) to correct for multiple testing.

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* and *gluifel*). For each list, children were instructed to read aloud the (pseudo)words as quickly and accurately as possible within 1 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 (Dunn & Dunn, 2007) was used (see van den Bosch et al., 2019 for further details). All 96 items were administered, and a child's score was the number of items answered correctly (Cronbach's α = .860, 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 and parents together) on a 5-point scale (0 = only Dutch, 1 = mostly Dutch, 2 = equal, 3 = mostly Turkish and 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, 2014). This standardised 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 tests administered halfway second grade (Cito M4), end second grade (Cito E4) and halfway third grade (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, because 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 minutes 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 national pupil monitoring system.

Data analyses

To answer the first research question, a mixed-design ANOVA was performed with time (T1, T2 and T3) as within-subject factor, language background (L1 vs L2) as

between-subject 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 3 *SDs* 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 with median split, which is an often used top-down procedure where individuals are categorised 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-group 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 versus L1 readers), we now selected language group as a between-subject factor to compare the L2 subgroups as determined by the cluster analysis with the L1 readers.

Results

Descriptive statistics

The descriptive statistics for the L1 readers and L2 readers are displayed in Table 1. To examine between-group differences, independent sample *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 L1 readers and L2 readers are shown in Table 2. Reading comprehension performance across time points showed a weak, positive correlation with decoding skills (particularly for L2 readers) 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 2. Pearson correlations with reading comprehension (T1–T3) for L1 readers below diagonal (n = 75) and L2 readers above diagonal (n = 71)

	DEC WD	DEC PS	VOC NL	VOC TU	RC T1	RC T2	RC T3
DEC WD	–	.95***	−.08	.14	.15	.35**	.26*
DEC PS	.95***	–	−.08	.16	.16	.31**	.27*
VOC NL	.06	.03	–	.04	.50***	.46***	.38**
VOC TU	n/a	n/a	n/a	–	.07	.06	.18
RC T1	.19	.18	.39**	n/a	–		.73***
RC T2	.23*	.22	.46***	n/a	.59***	–	.70***
RC T3	.20	.20	.42***	n/a	.66***	.73***	–

Notes: 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 versus 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 × 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 (Figure 1). Within-subject contrast testing showed that the interaction with language background existed for T1 versus T3, $F(1, 142) = 11.97$, $p = .001$, $\eta_p^2 = .078$, and for T2 versus T3, $F(1, 142) = 8.60$, $p = .004$, $\eta_p^2 = .057$, but not for T1 versus T2, $F(1, 142) = 0.31$, $p = .470$, $\eta_p^2 = .004$. Consequently, this indicates that 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, even though the L2 readers

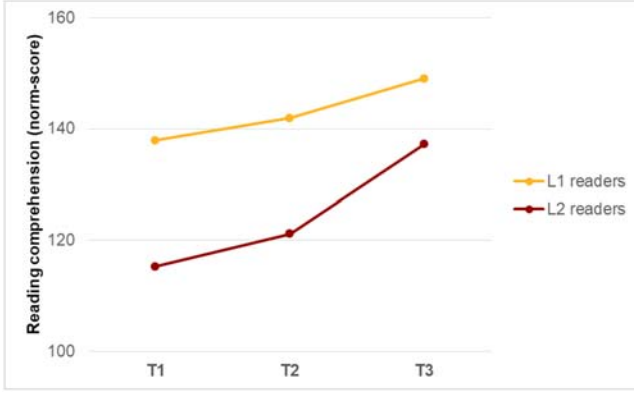

Figure 1. Growth in reading comprehension from T1 (start G2) to T2 (end G2) to T3 (mid G3) split by language background (L1 readers vs L2 readers). [Colour figure can be viewed at wileyonlinelibrary.com]

scored lower at all time points as can be concluded from the simple main effects reported in Table 1.

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 2, there was no significant correlation between the two vocabulary measures (r = .043, p = .721). This is also reflected in the scatterplot (Figure 2). Based on the scree plot and the amount of explained variance, we found support for a four-cluster solution that explained 70.7% of the variance in Dutch vocabulary (L2) and 65.1% of the variance in Turkish vocabulary (L1), with an average silhouette width of 0.34. Based on the scatterplot (Figure 2) and the mean *z*-scores for Dutch and Turkish vocabulary (Table 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%).

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 11 different schools. Second, the LowL2-LowL1 subgroup included only five children. These were extreme cases in terms of vocabulary; they scored more than 2 *SDs* below the mean on L2 vocabulary (n = 2), or they scored more than 1 *SD* (n = 1) or even 2 *SDs* (n = 2) below the mean on both L2 and L1 vocabulary (Figure 2). Furthermore, their SES was considerably lower compared with the other three L2 subgroups, and they had the lowest mean score for nonverbal reasoning (Table 3). Given the small sample size

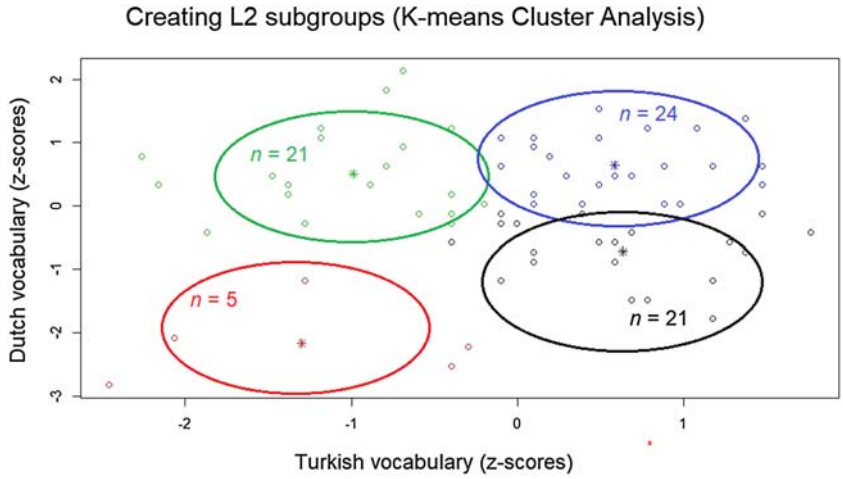

Figure 2. Scatterplot for the four-cluster solution: (1) LowL2-LowL1 (red), (2) LowL2-HighL1 (black), (3) HighL2-LowL1 (green) and (4) HighL2-HighL1 (blue). [Colour figure can be viewed at wileyonlinelibrary.com]

Table 3. Descriptive statistics split by 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)

Notes: 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 = nonverbal reasoning, SES = socio-economic status.
**$p < .01$.
***$p < .001$.

and their deviant profile, these five children were excluded from the analyses, but for descriptive purposes, they are still included in the remaining tables and figures.

Table 3 shows 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 nonverbal reasoning, $F(2, 63) = 1.94$, $p = .152$, $\eta_p^2 = .058$. Yet significant differences between the three L2

subgroups were found for reading comprehension, $F(2, 63) = 5.84$, $p = .005$, $\eta_p^2 = .156$, vocabulary knowledge (L1: $F(2, 63) = 55.44$, $p < .001$, $\eta_p^2 = .638$, L2: $F(2, 63) = 40.32$, $p < .001$, $\eta_p^2 = .561$) and language exposure ($F(2, 63) = 10.74$, $p < .001$, $\eta_p^2 = .254$). 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 with children in the HighL2-LowL1 subgroup ($p < .001$). This particular difference was also reflected in their main home language and language preference (Table 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.

Early reading comprehension development: L2 subgroups versus L1 readers

The results of the mixed-design ANOVA showed that there was a significant main effect of language group, $F(3, 135) = 9.03$, $p < .001$, $\eta_p^2 = .171$, and a significant main effect of time, $F(1.99, 268.99) = 6.15$, $p = .002$, $\eta_p^2 = .044$, on reading comprehension. In addition, there was a significant Time × Language group interaction effect, $F(5.98, 268.99) = 2.45$, $p = .025$, $\eta_p^2 = .052$, indicating that the degree of growth in reading comprehension differed across language groups (Figure 3). Within-subject contrast showed that the interaction existed for T1 versus T3, $F(3, 135) = 4.57$, $p = .004$, $\eta_p^2 = .092$, and for T2 versus T3, $F(3, 135) = 2.91$, $p = .037$, $\eta_p^2 = .061$, but not for T1 versus T2, $F(3, 135) = 0.44$, $p = .724$, $\eta_p^2 = .010$. 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, $F(1.96,$

Table 4. Frequencies 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%)

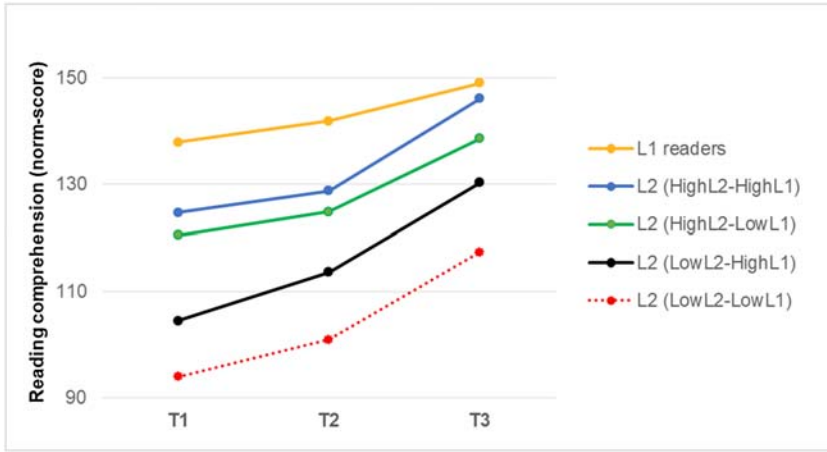

Figure 3. Growth in reading comprehension from T1 (start G2) to T2 (end G2) to T3 (mid G3) split by language group (L1 readers vs L2 subgroups). [Colour figure can be viewed at wileyonlinelibrary.com]

186.02) = 3.32, p = .039, η_p^2 = .034, but not for the HighL2-LowL1 subgroup, F(2, 184) = 1.24, p = .292, η_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 with L1 readers.

As an effect size partial eta squared (η_p^2) was reported earlier, 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 (Table 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 5. Quantifying mean differences (Cohen's d) between L2 subgroups and L1 readers for reading comprehension (RC) at T1 (start G2), T2 (end G2) and T3 (mid G3)

	LowL2-LowL1 versus L1 readers	LowL2-HighL1 versus L1 readers	HighL2-LowL1 versus L1 readers	HighL2-HighL1 versus 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 with 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 grades to their monolingual peers. 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 become more complex and include more academic language (Snow, 2010) 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 grades.

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 L2 vocabulary and low L1 vocabulary; and (4) children with high vocabulary in both their languages. This categorisation 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, which 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 categorisation 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 with 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 nonverbal 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 the 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 with 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 the Cummins (1979) interdependency hypothesis 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, Malda, Mesman, & Van IJzendoorn, 2016).

However, some caution is required when interpreting the latter finding as cross-language transfer, because 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, 2018).

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 determined by not only 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, 2017), 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

To conclude, the present study can be seen as an important step forward in explaining variation in early reading comprehension development of L1 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 a previous study (van den Bosch et al., 2019), the current findings also demonstrate that categorising 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 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 Royal Statistical Society*, 57, 289–300.

Bialystok, E. & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, 122, 67–73. https://doi.org/10.1016/j.cognition.2011.08.003.

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

Bialystok, E., Luk, G., Peets, K.F. & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(4), 525–531. https://doi.org/10.1017/S1366728909990423.

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. https://doi.org/10.1037/0022-0663.96.1.31.

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 [Reading Comprehension 3.0 for second grade]*. Arnhem, the Netherlands: Cito B.V.

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

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222–251. https://doi.org/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. https://doi.org/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 Netherlands' *other languages: 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. https://doi.org/10.1111/jrir.12003.

Field, A. (2009). *Discovering statistics using SPSS*. (Third edn). 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. https://doi.org/10.1007/s10648-011-9175-6.

Garcia, D., MacDonald, S. & Archer, T. (2015). Two different approaches to the affective profiles model: Median splits (variable-oriented) and cluster analysis (person-oriented). *PeerJ*, 3, e1380. https://doi.org/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. https://doi.org/10.1037/a0014320.

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

Jeon, E.H. & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160–212. https://doi.org/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(5), 612–620. https://doi.org/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. https://doi.org/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. https://doi.org/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(3), 701–711. https://doi.org/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(2), 409–433. https://doi.org/10.1037/a0033890.

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), e12433. https://doi.org/10.1111/desc.12433.

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

Perfetti, C. & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18, 22–37. https://doi.org/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. https://doi.org/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. https://doi.org/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(2), 405–424. https://doi.org/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. https://doi.org/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(9), 1381–1405. https://doi.org/10.1007/s11145-015-9575-y.

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

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. https://doi.org/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). Berlin, Germany: Mouton de

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. https://doi.org/10.1080/10888438.2018.1509864.

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

Verhoeven, L. & Keuning, J. (2017). The nature of developmental dyslexia in a transparent orthography. *Scientific Studies of Reading*, 22(1), 1–17. https://doi.org/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. https://doi.org/10.1016/j.jneuroling.2019.02.001.

Verhoeven, L., Voeten, M. & Vermeer, A. (2018). Beyond the simple view of early first and second language reading: The impact of lexical quality. *Journal of Neurolinguistics*, 50, 28–36. https://doi.org/10.1016/j.jneuroling.2018.03.002.

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

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics*, 27, 741–747. https://doi.org/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. https://doi.org/10.1598/RRQ.44.3.4.

Liza Jane van den Bosch obtained her PhD at the Behavioural Science Institute at Radboud University (Nijmegen, the Netherlands) under the supervision of Prof. Eliane Segers and Prof. Ludo Verhoeven. Her research focuses on reading comprehension in beginning readers. She currently works as a lecturer at Leiden University (Leiden, the Netherlands).

Eliane Segers is a professor in 'Learning and Technology' at the Behavioural Science Institute at Radboud University (Nijmegen, the Netherlands). She is also professor in 'Reading and Digital Media' at Twente University (Enschede, the Netherlands).

Ludo Verhoeven is an extraordinary professor in 'Variation in Communication and Literacy' at the Behavioural Science Institute at Radboud University (Nijmegen, the Netherlands). He is also professor in 'Language Innovation and Language Policy' at the University of Curaçao.

Received 4 March 2019; revised version received 7 February 2020.

Address for correspondence: Miss Liza J. van den Bosch, Behavioural Science Institute, Radboud University, Montessorilaan 3, PO Box 9104, 6500 HE Nijmegen, The Netherlands. E-mail: l.vandenbosch@pwo.ru.nl