



Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Research paper

Beyond the simple view of early first and second language reading: The impact of lexical quality

Ludo Verhoeven^{a,*}, Marinus Voeten^a, Anne Vermeer^b^a Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands^b Communication and Information Sciences, University of Tilburg, The Netherlands

A B S T R A C T

According to the simple view of reading (SVR), reading comprehension is the product of word decoding and listening comprehension. Against this background, we examined the additional role of early lexical quality in the prediction of reading comprehension, either directly or indirectly via word decoding or listening comprehension. Following a longitudinal design, 566 children learning to read Dutch as L1 and 463 children learning to read Dutch as L2 in the Netherlands were tested on indicators lexical quality (LQ) in kindergarten (speech decoding, morphological knowledge and vocabulary); word decoding and listening comprehension in first grade; and then reading comprehension in second grade. The results showed L2 learners to consistently lag behind L1 readers on all measures except for word decoding. Both word decoding and listening comprehension predicted later reading comprehension for not only L1 but also L2 learners. However, later reading comprehension was also directly predicted by the children's early morphological and vocabulary knowledge, on the one hand, and indirectly by speech decoding and morphological knowledge via word decoding and indirectly by morphological and vocabulary knowledge via listening comprehension. These results show the beginning reading achievement of both L1 and L2 learners to be largely predicted by the quality of their early lexicons.

Given the fact that second language (L2) learners are confronted with the task of learning to read in a language that they have yet to master, many of them experience serious reading difficulties. Beginning L2 readers generally show minor problems with word decoding but their reading comprehension often stays substantially behind their monolingual peers (see Verhoeven, 2017). In the Simple View of Reading (SVR) as originally put forth by Hoover and Gough (1990), reading comprehension is conceived as purely the product of word decoding and listening comprehension. Considerable empirical evidence has further shown the SVR to characterize both L1 reading development (Florit & Cain, 2011) and L2 reading development (Verhoeven and van Leeuwe, 2012). Learning to read in a second language (L2) is a particularly challenging task for children who have had only limited L2 exposure and therefore only limited listening comprehension to facilitate their starting reading. More recently, indicators of basic phonological, morphological, and semantic abilities (i.e., lexical quality) have also been shown to be predictive of both beginning L1 (Perfetti, 2007) and beginning L2 reading comprehension (Droop & Verhoeven, 2003). However, exactly how lexical quality relates to the other SVR predictors of reading comprehension is as yet unclear. In the present study, we therefore examined the impact of early phonological, morphological, and semantic abilities in kindergarten on the prediction of beginning L1 and L2 reading comprehension both directly and indirectly via children's word decoding and/or listening comprehension in first and second grade.

According to the Simple View of Reading (SVR) as proposed by Hoover and Gough (1990), reading comprehension can be

* Corresponding author. Radboud University Nijmegen, Behavioral Science Institute, P.O. Box 9044, 6500 KD, Nijmegen, The Netherlands.
E-mail address: l.verhoeven@pwo.ru.nl (L. Verhoeven).

conceived as the product of word decoding and listening comprehension. To begin with, it is claimed that listening comprehension or, in other words, the linguistic processes involved in the comprehension of oral language strongly constrain the process of reading comprehension. There is widespread empirical evidence for the role of early listening comprehension in early reading comprehension (see Kendeou, Savage, & Van den Broek, 2009; Perfetti, Landi, & Oakhill, 2005; Verhoeven & Perfetti, 2008). Within the SVR, a major role is also attributed to word decoding or, in other words, the fast and accurate retrieval of the phonological code for written word forms. And, indeed, automated word recognition has been shown to free mental resources for closer consideration of the meaning of a text and thereby to facilitate reading comprehension (see National Reading Panel, 2000).

During the past decades, the SVR has provided a useful theoretical framework for understanding the variation observed in learning to read and particularly the development of L1 versus L2 reading. Younger and poorer readers experience more problems than older and better readers with both listening and reading comprehension (e.g., Cain & Oakhill, 1998). In other studies, the role of word decoding in the development of children's reading comprehension has been shown to initially be large for beginning readers but decline as reading becomes more proficient and the role of listening comprehension becomes more prominent (e.g., Bast & Reitsma, 1998; Chen & Vellutino, 1997; Tunmer & Hoover, 1993). In a similar vein, L2 reading comprehension can be impeded for various reasons — including the possibility of delayed L2 word decoding (cf. Verhoeven, 2010). In the long run, however, the word decoding skills of L1 and L2 learners have been found to become equal (Siegel, 2003; Verhoeven, 2000). L2 speech decoding has been found to predict the early literacy skills of L2 learners (Janssen, Segers, McQueen, & Verhoeven, 2017), and as they learn to master the essentials of the target language, including the phonology and orthography, cross-language transfer may further facilitate their word decoding and thereby the development of their reading comprehension. L2 overlap with L1 phonology (Lindsey, Manis, & Bailey, 2003; Lopez & Greenfield, 2004) and orthography (Deacon, Chen, Luo, & Ramirez, 2011; Deacon, Wade-Woolley, & Kirby, 2009) has indeed been found to promote L2 word decoding skills.

Limited oral language proficiency can place the development of reading comprehension at risk, also for L2 learners (see Genesee, Lindholm-Leary, Saunders, & Christian, 2006). Recently, however, Verhoeven and van Leeuwe (2012) found equal levels of word decoding for the two groups of learners while both the listening comprehension and reading comprehension of the L2 learners still lagged behind that of the L1 learners. The predictive associations between word decoding, listening comprehension, and reading comprehension were found to be highly comparable for the two groups of learners, however, which suggests that the impact of word decoding on reading comprehension decreases while the impact of listening comprehension on reading comprehension increases to the same extent during the development of two such groups of learners. Similar findings were reported in other studies on second language learners as well (cf. Farnia & Geva, 2013; Bonifacci & Tobias, 2017).

Recent research has highlighted the importance of various aspects of lexical quality - semantic, phonological and orthographic aspects of words - for successful reading comprehension (cf. Perfetti & Stafura, 2014; Perfetti, 2007). Importantly, the learning of orthographic representations by word decoding is highly dependent on the quality of semantic and phonological representations of words in the mental lexicon (Perfetti, 1992). Knowledge of word meanings or, in other words, receptive vocabulary knowledge has indeed been found to be critical for adequate word decoding and reading comprehension (cf. Verhoeven, van Leeuwe, & Vermeer, 2012). Estimates of vocabulary knowledge further reveal large individual differences across both L1 and L2 learners (Vermeer, 2001). When Droop and Verhoeven (2003) studied the development of the oral language skills, word decoding, and reading comprehension of L1 and L2 learners in the third and fourth grades of Dutch elementary school, they found, in particular, vocabulary knowledge to play a more prominent role in the explanation of the observed variation in the reading comprehension of the L2 learners. Proctor, Carlo, August, and Snow (2005) similarly found vocabulary to play a more predictive role than word decoding in the reading comprehension of Spanish-speaking learners of English as a second language. L2 vocabulary knowledge can thus be seen to be a crucial predictor of L2 reading comprehension.

It is important to note that not only the number of lexical entries but also the precision of the information contained in lexical entries can shape the process of learning to read. According to the lexical quality hypothesis (Perfetti & Hart, 2001), the reading skill of both children and adults is supported not only by the sheer number of words in the mental lexicon but also by the precision of the phonological and morphological information stored in the mental lexicon. It is now widely assumed that phonological knowledge, no matter how implicit, is critical for the development of oral language but also reading. Phonological connections that depend on critical speech distinctions indeed play a central role in the emergence of literacy. The extent to which children's speech decoding ability addresses a fully specified sound structure of words in the mental lexicon has indeed been shown to influence phonemic awareness and subsequent reading ability in both L1 learners (Elbro, 1996; Goch, McQueen, & Verhoeven, 2014) and L2 learners (Janssen, Segers, McQueen, & Verhoeven, 2015). In addition, the quality of morphological knowledge, including information on word roots, syntactic inflections, and derivational affixes, stored in the mental lexicon has been found to influence reading comprehension both directly (Deacon, Kieffer, & Laroche, 2014) and indirectly via word decoding (see Reichle & Perfetti, 2003).

From research to date, it can be concluded that the simple view of reading (SVR) provides an adequate framework for understanding the development of not only L1 reading comprehension but also L2 reading comprehension. The explanatory power of the SVR can nevertheless be expanded with the addition of measures of early lexical quality. Measures related to phonological distinctions in speech decoding, morphological knowledge, and receptive vocabulary can help predict both L1 and L2 word decoding, listening comprehension, and later reading comprehension. In the present study, we therefore investigated longitudinally the role of such lexical quality measures in the prediction of children's L1 or L2 reading comprehension either directly or indirectly via their word decoding and listening comprehension.

In the Netherlands, children enter elementary school at the age of four. After two years of kindergarten (or the equivalent of one year of preschool and one year of kindergarten), formal reading instruction is initiated. We therefore examined the longitudinal associations between 1) various indicators of lexical quality (LQ) in kindergarten, namely, children's speech decoding, morphological

knowledge, and receptive vocabulary; 2) word decoding and listening comprehension in first grade; and 3) reading comprehension in second grade. This was done in the Netherlands for 566 children learning to read Dutch as L1 and 463 children learning to read Dutch as L2. Specifically, we sought to determine if the scores of L1 and L2 learners differ for components of the SVR and LQ. In addition, the effects between LQ predictor and SVR criterion variables were examined in L1 and L2 using multiple group path modeling. By doing this, we aimed to answer the following research questions.

1. To what extent are there differences in kindergarten lexical quality skills (speech decoding, morphological knowledge, receptive vocabulary), first grade word decoding and listening comprehension skills and second grade reading comprehension skills of L1 and L2 learners?
2. What is the impact of kindergarten lexical quality measures on early L1 and L2 reading development?

Our first prediction was that in comparison to their L1 peers, L2 learners would lag behind on measures of LQ, listening comprehension, and reading comprehension but not on word decoding. Our second prediction was that the LQ measures collected at the end of kindergarten would similarly affect the measures of SVR later collected among the L1 and L2 learners. Finally, for both the L1 and L2 learners, it was hypothesized that they would predict their later reading comprehension both directly and indirectly via their word decoding and listening comprehension.

1. Method

1.1. Participants

The participants in this study were 1029 students from 76 Dutch elementary schools (507 boys, 522 girls). Longitudinal data were collected in three waves: at the end of kindergarten, the end of grade one, and the end of grade two. The sample was generated by first stratifying schools across different parts of the country, followed by a random sampling of schools; the kindergarteners of these schools were selected for the sample. All parents were asked for passive consent, which was given in more than ninety percent of the cases. The selected group of children used in the main analyses consisted of 566 L1 learners and 463 L2 learners who were tested on speech decoding, morphological knowledge and receptive vocabulary by the end of kindergarten and on listening comprehension in first grade. A subsample of 60 schools with 460 L1 learners and 347 L2 learners followed a standardized national assessment program for word decoding in first grade of which 17 schools with 143 L1 and 84 L2 learners also used one and the same standardized reading comprehension assessment in second grade (Cito, 1992). The percentage of missing values was around 1.5% for the variables measured in kindergarten, 14% (listening comprehension) to 33% (word decoding) in grade 1, and 81% for reading comprehension in grade 2. The large percentages of missings for word decoding and reading comprehension were due to the fact that not all sample schools did take part in the national assessment program from which the test scores were taken.

The group of L2 learners included students of a Mediterranean, mainly Turkish or Moroccan, origin (56%) as well as students from Suriname or the Dutch Antilles (16.5%), and other countries (18.5%), in addition to children from mixed marriages (9%). The Turkish and Moroccan children were part of the third generation of immigrants who came to the Netherlands to work at industrial sites, whereas the children from Suriname and the Dutch Antilles were descendants of families who came to the Netherlands after decolonization of their home countries. For the students in the L1 group, the home language was Dutch; for the students in the L2 group, besides Dutch another language was spoken at home, or no Dutch at all (32%). The educational level of the parents of the students in the two groups also differed: of the parents in the L1 group, 42% had completed lower-level education, 29% middle-level education and 29% higher-level education whereas in the L2 group the proportions were 73%, 17% and 10%, respectively.

1.2. Measures

1.2.1. Word decoding

Word Decoding was measured using the standardized Three Minutes Test (Verhoeven, 1991), consisting of three cards with words of different Dutch orthographic complexity. The first card presents a list of 150 high-frequency monosyllabic words of the type consonant-vowel-consonant (e.g., *dak/dak*/'roof'); the second card presents 150 high-frequency monosyllabic words with consonant clusters at the beginning or end of the word (e.g., *straat/strat*/'street'); and the third card presents 120 less frequent polysyllabic words of increasing difficulty (e.g., *boter/botθr*/'butter', *aardappel/ardapθl*/'potato'). The test is administered individually and the child is asked to read as many words as possible in 1 min for each card. The child's score for each list is the number of words read correctly. The Cronbach's alphas of .87, .92, and 0.91 attest to the reliability of the three measures (cf. Verhoeven & van Leeuwe, 2009).

1.2.2. Reading comprehension

A standardized test was administered for each grade level and required the children to read a series of short written narrative texts and answer 30 multiple-choice questions about the texts (Cito, 1992). The items address both explicit and implicit meaning relations between the sentences in the individual texts. There is no time limit for the test. And given that the items constitute a unidimensional scale, scale scores could be computed for each child. A Cronbach's alpha of .86 was reported.

1.2.3. Listening comprehension

Part of the Revised Dutch Language Proficiency Test (cf. Verhoeven & Vermeer, 2001), which is a standardized discrete-point test to assess the Dutch oral language proficiency of 4- to 10-year-olds was used to assess L1 and L2 listening comprehension. Six short stories are read aloud to the individual child and, following each story, the child answers four questions for a total of 24 test items. A Cronbach's alpha of .84 was reported.

1.2.4. Lexical quality measures

The quality of the children's Dutch lexicon was evaluated using three tasks from the Revised Dutch Language Proficiency Test, which is a standardized discrete-point test to assess the Dutch oral language proficiency of 4 to 10-year-olds (cf. Verhoeven & Vermeer, 2001). The number of items constituting each task also represents the maximum possible score for each task.

- Speech decoding task requiring 50 pairs of Dutch phonemes to be distinguished from each other.
- Morphological knowledge task (24 items) requiring the child to form the plural of nouns and the past tense of verbs for words presented orally along with the support of a picture (“This is one *key*, these are two ...”).
- Receptive vocabulary task (96 items) requiring the child to point to the one picture representing an orally presented word out of four pictures.

All of the lexical quality tests produced sufficiently reliable outcomes, with Cronbach's alphas of .96, .92 and 0.97.

1.3. Procedure

Multiple-group path analyses were conducted for the native Dutch (L1) and ethnic minority students (L2) separately using Mplus 7 (Muthén & Muthén, 1998–2012). Students with no data whatsoever on one or more of the variables selected for analysis ($N = 213$) or with a missing value on a) all three of the exogenous variables at the end of kindergarten ($N = 35$) or b) all three of the dependent measures at the ends of grades one and two ($N = 138$) were omitted from the main analyses. As mentioned, for both groups, a large number of missing values occurred on the grade-1 Word Decoding test and the grade-2 Reading Comprehension test. These two tests were administered by schools as part of a standardized, national testing program, but not all of the schools included in our study participated in the program. We used full information maximum likelihood (FIML) estimation to deal with the missing values. This requires the assumption that the missing values are missing at random (MAR), but not completely at random (MCAR) (cf. Enders, 2010). Missing values were not related to the purpose of the present study, but they appeared not missing completely at random. Students with a missing value on a variable generally scored on average lower on other variables than students with a valid score did. But the FIML procedure does only require MAR, which is much less demanding than MCAR. The MAR assumption is not testable (Enders, 2010, p. 17).

A path model was constructed (see the final model in Fig. 1) to test the direct and indirect effects of kindergarten Lexical Quality (i.e., speech decoding, morphological knowledge, and receptive vocabulary) on grade-1 Word Decoding (WD), grade-1 Listening Comprehension (LC), and grade-2 Reading Comprehension (RC). In the proposed model, RC was hypothesized to be predicted by WD and LC in Grade 1 and by Speech Decoding (SD), Morphological Knowledge (MK), and Receptive Vocabulary (RV) in kindergarten. Furthermore, WD was hypothesized to be predicted by SD and MK, while LC was hypothesized to be predicted by MK and RV.

We expected the model to be the same for both groups of students and therefore started from the same model with all parameters freely estimated for the two groups of students. Maximum likelihood estimation was used (MLR estimator in Mplus), taking into account the clustering of students in schools. The goodness of fit for the proposed model was evaluated using the chi-square statistic, the Root Mean Square Error of Approximation (RMSEA), and the Comparative Fit Index (CFI). Conventional rules of thumb for the RMSEA state that a value below 0.08 represents an acceptable fit and a value below 0.05 a good fit (e.g., Little, 2013, p. 109). A CFI above 0.95 indicates a good fit (Little, 2013, p. 115). For the comparison of nested models (i.e., testing for regression differences between the two groups of students), we applied the Satorra-Bentler Scaled Chi-Square (Satorra & Bentler, 2011). In addition, we used the Bayesian Information Criterion (BIC) to compare the fit of the models (Raftery, 1995). For this measure, the relatively best model is that with the lowest BIC value.

2. Results

2.1. Descriptive statistics

Table 1 shows the means and standard deviations of all study variables, separately for the two groups of students. As could be expected, the L2 students scored lower on average than the L1 students on all measures; the mean difference between the two groups was statistically significant for all measures except grade-1 Word Decoding (WD). The largest group differences manifested themselves for kindergarten Receptive Vocabulary (RV) and kindergarten Morphological Knowledge (MK) with the mean differences being larger than the standard deviation (see effect size d in Table 1). For grade-1 Listening Comprehension (LC), the mean difference approached the standard deviation. The standard deviations were generally higher for the L2 group than for the L1 group, with the exception of grade-2 Reading Comprehension (RC).

Table 2 shows the correlations between the study variables for the L1 and L2 groups separately. All of the correlations between the predictor variables and RC were moderately positive and statistically significant. The correlations were also found to be highly similar

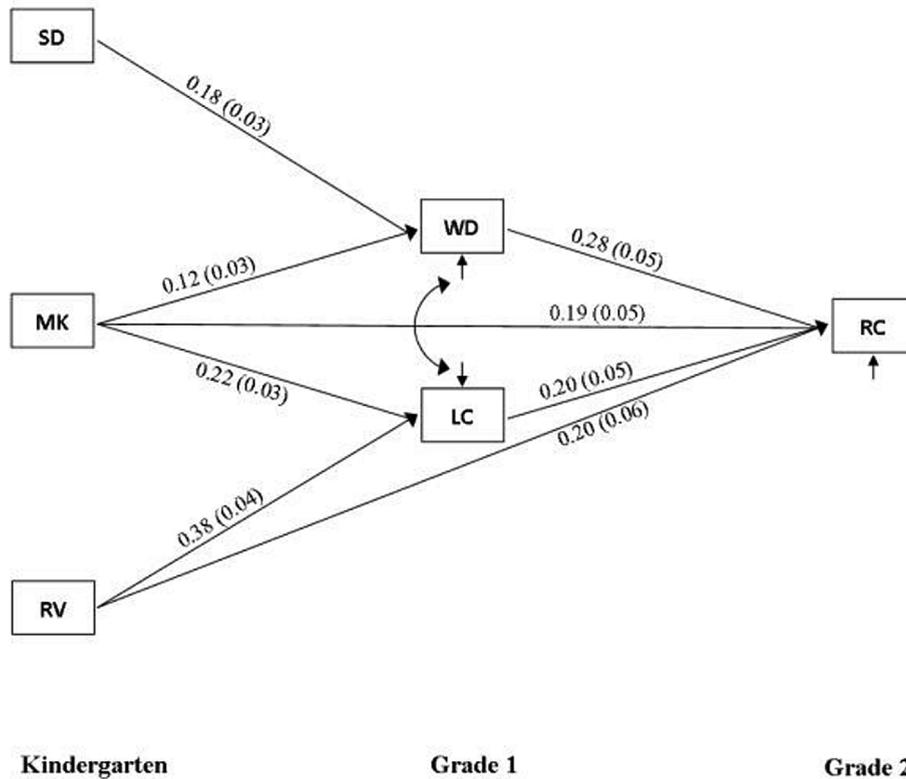


Fig. 1. Final path model (Model 4) with standardized parameter estimates and their standard errors (within parentheses). Correlations between kindergarten measures, residual variance in grade 1, and residual variances in grades 1 and 2 were specific for each group and omitted from the figure. The standardized parameter estimates apply to the L1 group; for L2 these estimates were highly similar but not exactly equal because of group differences in variances. The unstandardized regression coefficients were fixed to be equal for L1 and L2 students. SD = Speech Decoding, MK = Morphological Knowledge, RV = Receptive Vocabulary, WD = word decoding, LC = Listening Comprehension, RC = Reading Comprehension.

Table 1
Means and standard deviations L1 and L2 groups of students.

	L1 Dutch			L2 Dutch			Mean Difference	
	N	Mean	SD	N	Mean	SD	Z ^b	d
Speech Decoding (SD), kindergarten	639 ^a	46.20	4.842	543	43.98	6.609	5.34***	0.39
Receptive Vocabulary (RV), kindergarten	642	66.17	13.074	543	43.01	15.611	16.26***	1.62
Morphological Knowledge (MK), kindergarten	639	15.83	4.319	542	9.09	5.565	14.56***	1.37
Word Decoding (WD), grade 1	460	67.79	38.199	347	66.56	37.830	0.79	0.03
Listening Comprehension (LC), grade 1	564	19.62	3.687	468	15.80	4.907	9.72***	0.89
Reading Comprehension (RC), grade 2	143	100.99	6.809	84	98.08	6.143	3.78***	0.44

***p < .001.

^a All available data used.

^b Mean differences tested with correction of standard errors for clustering of students by schools; Z = Mean Difference/Standard Error, d = Mean Difference/Pooled SD.

for the L1 and L2 groups of students.

2.2. Testing of path models

The model was built up in three steps: (i) SVR variables only for grades 1 and 2, (ii) adding early LQ variables to predict the two grade-1 variables, allowing only indirect effects of LQ on reading comprehension, and (iii) adding direct effects of early LQ on reading comprehension. For estimating the first model a sample of 1045 students was available, while the other models including the LQ variables could be estimated on a sample of 1029 students. We started out from a path model predicting grade-2 reading comprehension from the two grade-1 variables (Model 1), to test whether the data for both groups of students were in agreement with the

Table 2
Correlations between Measures for L1 (below diagonal) and L2 (above diagonal).

	SD	RV	MK	WD	LC	RC
Speech Decoding (SD), kindergarten		.30	.30	.21	.28	.33
Receptive Vocabulary (RV), kindergarten	.35		.62	.10	.48	.42
Morphological Knowledge (MK), kindergarten	.31	.52		.16	.45	.45
Word Decoding (WD), grade 1	.20	.19	.19		.12	.32
Listening Comprehension (LC), grade 1	.21	.49	.40	.16		.47
Reading Comprehension (RC), grade 2	.25	.46	.43	.34	.43	

All correlations were statistically significant ($p < .05$) except that between Receptive Vocabulary and Word Decoding for the L2 Group.

SVR and whether the regression coefficients of grade-1 word decoding and grade-1 listening comprehension were the same across groups. This hypothesis was confirmed, $\chi^2(2, N = 1045) = 1.501, p = .472$. Both predictors showed a statistically significant ($p < .001$) positive relationship with reading comprehension. Moreover, word decoding and listening comprehension appeared almost equally important predictors of grade-2 reading comprehension (based on t-values and standardized regression coefficients, in the L1-group $\beta = 0.32$ for WD and 0.33 for LC, while in the L2-group $\beta = 0.33$ for WD and 0.46 for LC¹).

Next, we extended Model 1 with the three lexical quality variables observed in kindergarten as predictors of LC or WD in first grade and only indirect effects on RC in second grade. In this Model 2, again RC at the end of grade 2 was expected to be predicted by WD and LC at the end of grade 1. WD and LC in turn were expected to be influenced by the lexical quality variables; WD was related to SD and MK, and LC was related to RV and MK. Thus SD influenced RC only indirectly through WD. MK influenced RC indirectly through both WD and LC. And RV influenced RC only through LC. This model was fitted to the data for the two groups of students simultaneously, allowing all hypothesized relationships to differ for the two groups. As can be seen from the goodness-of-fit statistics in Table 3, Model 2 did not provide a good fit for the data. The chi-square statistic was particularly high. This lack of fit suggested that the LQ variables had not only indirect but also direct effects on RC. Therefore, we constructed Model 3 including direct effects of LQ variables.

Model 3 is the same as Model 2 but with two regression relationships added. MK and RV in kindergarten were now allowed to have a direct effect on grade-2 RC in addition to their indirect effects on RC via grade-1 LC or WD. Again, the model was estimated simultaneously for the L1 and L2 groups of students without restrictions on the parameters. The estimated model fit the data satisfactorily, as shown in Table 3. The chi-square statistic was no longer statistically significant ($p > .05$). And the RMSEA and CFI also indicated a good fit. Model 3 provided evidence of direct effects of early LQ on reading comprehension. We did not hypothesize a direct effect of speech discrimination on reading comprehension, since speech discrimination was supposed to work at the word level and not beyond. Nevertheless, we checked a possible direct effect of SD on RC. As expected this effect was not statistically significant, $b = 0.11, p = .58$ for L1 students and $b = 0.12, p = .16$ for L2 students). Therefore such a direct effect was left out of the model.

We next tested whether the regression (i.e., predictive) relations within Model 3 could be considered the same for the two groups of students. Model 4 is the same as Model 3 but now with the restriction that all of the regression coefficients be the same across the two groups. Model 4, however, still allows the correlations and residual variances to be different for the two groups. The more parsimonious Model 4 appeared to provide a good fit, $\chi^2(14) = 17.931, p = .210$, RMSEA = 0.023 (90% confidence interval 0.000 - 0.051). The fit statistics in Table 3 further show Model 4 to provide the best fit in terms of RMSEA, CFI, and BIC. When we compared the fit of Models 3 and 4 using the Satorra-Bentler Scaled Difference Chi-square Test (Satorra & Bentler, 2011), which is a statistical test of equality of regression coefficients across groups, no evidence for significant group differences was found; the test result was $\chi^2(8) = 5.931, p = .655$. The prediction of grade-1 WD and grade-1 LC but also of grade-2 RC was the same for the L1 and L2 groups of students.

Model 4 is depicted in Fig. 1 together with the standardized parameter estimates (regression coefficients) and their standard errors. As expected, both SD and MK at the end of kindergarten predicted WD at the end of grade 1. The R-squares were rather low, however (0.06 for the L1 group and .10 for the L2 group). RV and MK at the end of kindergarten predicted LC at the end of grade 1. The R-squares amounted to 0.28 and 0.26 for the L1 and L2 groups of students, respectively. Finally, the two grade-1 variables of WD and LC in addition to the two kindergarten variables of RV and MK predicted RC at the end of grade 2. All four of these predictor variables showed statistically significant regression coefficients ($p < .001$ except for RV ($p < .01$)). Per group, the standardized regression coefficients were about equal for the four predictors (for L1: 0.19 to 0.27; for L2: 0.24 to 0.27). The R-squares greatly differed for the two groups despite the requirement that the regression coefficients for all of the predictors be equal for the two groups. For the L1 students, the R-square for reading comprehension was .35; for the L2 students, it was 0.51.

Finally, we studied mean group differences by looking at the intercepts for the two groups of students estimated in the final model (Model 4). There were three dependent variables and thus three intercepts to be estimated for each group of students: the intercepts for grade-1 WD, grade-1 LC, and grade-2 RC. As shown in Table 1, these three variables had larger means for the L1 group compared to the L2 group and these differences were statistically significant for both LC and RC. But similar differences already existed for the predictor measures in kindergarten. The question that thus arose was whether the group differences in grades 1 and 2 could be

¹ Unstandardized regression coefficients were equal across groups; nevertheless standardized regression coefficients still may differ because of differing variances across groups.

Table 3
Goodness-of-Fit Statistics for the Path Models with SVR and LQ variables.

Model ^a	χ^2	df	<i>p</i>	RMSEA	CFI	BIC ^b
Model 2 with only indirect effects of LQ	44.210	10	.000	.082	.911	14992.5
Model 3 with direct and indirect effects of LQ	10.733	6	.097	.039	.988	14975.0
Model 4 with equality of coefficients across groups	17.931	14	.210	.023	.990	14951.0

^a Model 1 not included because it has no degrees of freedom.

^b Sample-size adjusted Bayesian Information Criterion.

explained by the earlier group differences. This was investigated by testing for equality of group intercepts.

Table 4 shows the intercepts estimated from Model 3 and the Wald statistics resulting from equating the intercepts for each of the three variables across groups. This analysis resembles an analysis of covariance, which is a test of group means on a variable after control for the effects of one or more other variables. For all of the dependent variables, Table 4 shows a higher intercept for the L2 students than for the L1 students. This difference in the group intercepts was statistically significant for WD ($p = .042$) and RC ($p = .015$) but not for LC. These results show that the mean group differences observed in grade-1 listening comprehension (see Table 1) could be explained by earlier differences between the groups. Differences in RV and MK in kindergarten predict listening comprehension in grade 1 (see Fig. 1). Word decoding showed no mean differences between groups (see Table 1), but after controlling for the LQ predictors in kindergarten, a small mean difference in favor of L2 students did arise. As Table 1 shows, L2 students were on average lagging behind L1 students in reading comprehension. From the intercepts in Table 4, however, L2 students seem on average to outperform L1 students after controlling for the SVR and early LQ predictors. Thus, mean differences between L1 and L2 in grade-2 RC may be explainable from earlier differences in SVR and LQ predictors.

3. Discussion

Several conclusions can be drawn from the present results. First, substantial differences between first and second language learners were found for all language and literacy skills with the exception of word decoding in the present study. For the variables of listening comprehension, reading comprehension, and speech decoding, the differences can be qualified as moderate. For morphological knowledge and receptive vocabulary, the differences can be qualified as large. It is striking that not only the listening comprehension and reading comprehension scores of the L2 students but also their scores on all three measures of early lexical quality lagged behind those of the L1 students. L2 learners must therefore build their orthographic insights on relatively weak L2 phonological, morphological, and semantic knowledge. The finding of no significant differences between the groups on word decoding or the basics of reading shows that L2 learners are able to attain near native levels despite having to learn to read in a foreign language. This finding is in line with the findings of previous studies (Lesaux & Siegel, 2003; Siegel, 2003; Verhoeven, 2010), and it suggests that cross-language transfer may help second-language learners but only to the extent of overlap between L1 and L2 phonologies and orthographies (cf. Janssen et al., 2015).

The L1 and L2 data presented here lend support to the simple view of reading but also take our understanding of learning to read beyond this view. For both the L1 and L2 learners, we found moderate to high correlations between reading comprehension, on the one hand, and word decoding and listening comprehension, on the other hand. In addition to this, our path model shows the early lexical quality measures of morphological knowledge and receptive vocabulary — even when measured two years prior — to directly predict later reading comprehension. This is over and beyond the indirect effects of these kindergarten variables on reading comprehension in grade 2 via word decoding and listening comprehension in grade 1. In other words, the quality of the mental lexicon as early as in kindergarten adds substantially to the prediction of beginning reading comprehension. The fact that the same predictive model provided an adequate fit for the data from both the L1 and L2 learners further makes it clear that learning to read depends highly on the quality of children's mental lexicons (see also Verhoeven & Van Leeuwe, 2012).

To summarize: going beyond the simple view of reading, the present study shows the quality of children's early lexical knowledge to significantly affect their later reading comprehension. In a model combining measures from the simple view of reading with measures of early lexical quality, it is clear that not only the *number* of lexical entries but also the *detail* of the phonological and morphological information represented in these entries contributes to the success of children's later word decoding and even their later reading comprehension. Interestingly, our result that children's knowledge of inflectional morphology contributes to their reading is fully commensurate with earlier research outcomes such as by Pliatsikas and Marinis (2013) who showed that both native English and L2 English students apply rule-based regular and irregular past-tense morphological processing in a self-paced reading

Table 4
Intercepts for two groups of students in the final model (model 4).

	L1 Dutch	L2 Dutch	Wald χ^2 (1)	<i>p</i>
Word Decoding	−17.02	−9.32	4.14	.042
Listening Comprehension	9.15	9.27	0.09	.760
Reading Comprehension	78.17	80.78	5.88	.015

task. Our results are in line with previous research outcomes (Perfetti, 2007; Richter, Isberner, Naumann, & Neeb, 2013) but also reinforced by the present finding that the *same* predictors prevail for *both* beginning L1 and L2 readers. Although the lexical quality and listening comprehension of beginning L2 readers generally lags behind the lexical quality and listening comprehension of beginning L1 readers, the reading comprehension for both groups appears to build upon the same set of linguistic predictors (see Fraser, Massey-Garrison, & Geva, 2016; Verhoeven, 2010; Wolter & Helms-Park, 2016).

Some potential limitations apply to the present study. To begin with, due to a variety of practical issues beyond our control, there was considerable missing data, especially in our final criterion variable, reading comprehension. We took care of this problem by using FIML estimation, but we cannot be sure that the data satisfied the assumptions of this procedure; we could only assume that the missing data were missing at random, though not completely at random. Another possible limitation is that the analyses in the present study were conducted without taking the socio-economic backgrounds of the L1 and L2 students into consideration. Adding the parental educational background to the models as a predictor of grade-2 reading comprehension significantly improved predictability. But it hardly changed the parameter estimates of the models; thus conclusions remained the same. Also, possibly important sociolinguistic measures, such as the nature and extent of language input, were not taken into consideration. Replication with even larger groups, perhaps using a more rigorous research design, is called for. And finally, documentation of the validity of the collected measures is also called for. For most of the variables in our study, we had only one measure. In future studies, multiple measures derived from the simple view of reading and the lexical quality hypothesis should be used.

To close, the present findings have some important practical implications. First, the finding that lexical quality measures assessed in kindergarten clearly and consistently predict beginning reading comprehension two years later suggests that children's lexical knowledge should be assessed from an early age and efforts to stimulate the development of this knowledge are needed. Intense vocabulary training, for example, should certainly be offered in preschool and kindergarten — vocabulary training that emphasizes both information on word forms and word meanings in a rich contextual environment (Biemiller & Boote, 2006). Finally, to help children achieve deeper levels of text comprehension (i.e., reading competence), we can recommend engagement in early collaborative storybook reading activities with a focus on both lexical information and its unification in the text to not only personal knowledge but also other people's knowledge (cf. Snow, Burns, & Griffin, 1998; Sénéchal & LeFevre, 2002). The implementation of a comprehensive model of reading development incorporating not only word decoding and listening comprehension but also lexical quality for first and second language learners can thus be warranted.

References

- Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology, 34*, 1373–1399.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*(1), 44–62.
- Bonifacci, P., & Tobias, T. (2017). The simple view of reading in bilingual language-minority children acquiring a highly transparent second language. *Scientific Studies of Reading, 21*, 109–119.
- Cain, K., & Oakhill, J. (1998). Comprehension skill and inference-making ability: Issues of causality. In C. Hulme, & R. Joshi (Eds.). *Reading and spelling: Development and disorders* (pp. 329–342). Mahwah, NJ: Lawrence Erlbaum Associates.
- Chen, R. S., & Vellutino, F. R. (1997). Prediction of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research, 29*, 1–24.
- Cito (1992). *Standardized test for reading comprehension in the primary grades*. Arnhem: Central Institute for Educational Measurement.
- Deacon, S. H., Chen, X., Luo, Y., & Ramirez, G. (2011). Beyond language borders: Orthographic processing and word reading in Spanish-English bilinguals. *Journal of Research in Reading, 3*, 1–17.
- Deacon, S. H., Kieffer, M. J., & Laroche, A. (2014). The relation between morphological awareness and reading comprehension: Evidence from mediation and longitudinal models. *Scientific Studies of Reading, 18*(6), 432–451.
- Deacon, S. H., Wade-Woolley, L., & Kirby, J. R. (2009). Flexibility in young second-language learners: Examining the language specificity of orthographic processing. *Journal of Research in Reading, 32*, 215–229.
- Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*, 78–103.
- Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing: An Interdisciplinary Journal, 8*, 453–485.
- Enders, C. K. (2010). *Applied missing data analysis*. New York/London: The Guilford Press.
- Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading, 36*(4), 389–421.
- Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review, 23*(4), 553–576.
- Fraser, C., Massey-Garrison, A., & Geva, E. (2016). Reading problems. In X. Chen, V. Dronić, & R. Helms-Park (Eds.). *Reading in a second language* (pp. 236–281). New York, NY: Routledge.
- Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2006). *Educating English language learners: A synthesis of research evidence*. New York, NY: Cambridge University Press.
- Goch, M. M., van McQueen, J. M., & Verhoeven, L. T. W. (2014). Learning phonologically specific new words fosters rhyme awareness in Dutch preliterate children. *Scientific Studies of Reading, 18*(3), 155–172.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127–160.
- Janssen, C., Segers, P. C. J., McQueen, J. M., & Verhoeven, L. (2015). Lexical specificity training effects in second language learners. *Language Learning, 65*(2), 358–389.
- Janssen, C., Segers, E., McQueen, J. M., & Verhoeven, L. (2017). Transfer from implicit to explicit phonological abilities in first and second language learners. *Bilingualism: Language and Cognition, 20*, 795–812.
- Kendeou, P., Savage, R., & Van den Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology, 79*, 353–370.
- Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39*, 1005–1019.
- Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-learners. *Journal of Educational Psychology, 95*, 482–494.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York/London: The Guilford Press.
- Lopez, L. M., & Greenfield, D. B. (2004). The cross-language transfer of phonological skills of Hispanic Head Start children. *Bilingual Research Journal, 28*, 1–18.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: The National Institute of Child Health and Human Development.
- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.). *Reading acquisition* (pp. 145–174). Hillsdale,

- NJ: Lawrence Erlbaum Associates.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357–383.
- Perfetti, C. A., & Hart, L. (2001). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.). *Precursors of functional literacy* (pp. 189–214). Amsterdam/Philadelphia: John Benjamins.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, & C. Hulme (Eds.). *The science of reading: A handbook* (pp. 227–247). Oxford: Basil Blackwell.
- Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18, 22–37.
- Pliatsikas, C., & Marinis, T. (2013). Processing of regular and irregular past tense morphology in highly proficient second language learners of English: A self-paced reading study. *Applied Psycholinguistics*, 34(5), 943–970.
- Proctor, C. P., Carlo, M., August, D., & Snow, C. E. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, 97, 246–256.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163.
- Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, 7(3), 219–237.
- Richter, T., Isberner, M. B., Naumann, J., & Neeb, Y. (2013). Lexical quality and reading comprehension in primary school children. *Scientific Studies of Reading*, 17, 415–434.
- Satorra, A., & Bentler, P. (2011). *A scaled difference chi-square test statistic for moment structure analysis*. Department of statistics, UCLA. UCLA: Department of statistics. UCLA. Retrieved from <https://escholarship.org/uc/item/23c604tb>.
- Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, 73, 445–460.
- Siegel, L. (2003). Bilingualism and reading. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.). *Precursors of functional literacy* (pp. 287–302). Amsterdam, The Netherlands: John Benjamins.
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington: National Academy Press.
- Tunmer, W., & Hoover, W. (1993). Components of variance models of language-related factors in reading disability: A conceptual overview. In R. J. Joshi, & C. K. Leong (Eds.). *Reading disabilities: Diagnosis and component processes* (pp. 135–173). Dordrecht, The Netherlands: Kluwer.
- Verhoeven, L. (1991). *Drie-minuten-toets Word decoding test*. Arnhem: Cito).
- Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, 4, 313–330.
- Verhoeven, L. (2010). Second language reading acquisition. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Eds.). *Handbook of reading research* (pp. 661–683). New York: Taylor & Francis.
- Verhoeven, L. (2017). Learning to read in a second language. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.). *Theories of reading development* (pp. 215–234). Amsterdam: John Benjamins.
- Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing: An Interdisciplinary Journal*, 25(8), 1805–1818.
- Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. *Applied Cognitive Psychology*, 22, 293–301.
- Verhoeven, L., & van Leeuwe, J. (2009). Modeling the growth of word decoding skills: Evidence from Dutch. *Scientific Studies of Reading*, 13, 205–223.
- Verhoeven, L., & Vermeer, A. (2001). *Taaltoets alle kinderen [standardized language test for all children]*. Arnhem: Cito).
- 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., & Helms-Park, R. (2016). The role of lexical knowledge in second language reading. In X. Chen, V. Dronić, & R. Helms-Park (Eds.). *Reading in a second language* (pp. 133–158). New York, NY: Routledge.