THE JOURNAL OF
CHILD PSYCHOLOGY AND PSYCHIATRY
ACAMH
Journal of Child Psychology and Psychiatry 51:5 (2010), pp 612–620 doi:10.1111/j.1469-7610.2009.02185.x

Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners

Arne Lervåg and Vibeke Grøver Aukrust
University of Oslo, Norway

Background: This study examines the role of decoding and vocabulary skills as longitudinal predictors of reading comprehension in young first (L1) and second (L2) language learners. **Methods:** Two-group latent growth models were used to assess differences in growth and predictions of growth between the 198 L1 and 90 L2 language learners. **Results:** L1 learners had better initial reading comprehension skills and faster growth in these skills over time. Individual differences in decoding and vocabulary predicted initial reading comprehension skills, but only vocabulary predicted the subsequent growth of reading comprehension skills. Vocabulary seemed to be a stronger predictor of growth in reading comprehension among the L2 learners than among the L1 learners. **Conclusions:** Vocabulary appears to be a critical predictor of the early development of reading comprehension skills in both L1 and L2 learners. The limitations in vocabulary skills in the L2 learners seemed sufficient to explain their lag in developing reading comprehension skills, and this suggests that oral vocabulary training should be given a high priority in this group. **Keywords:** Reading, language, comprehension, educational psychology. **Abbreviations:** DEP: Danish Ability Scales; L1: first language learners; L2: second language learners; NARA II: Neale Analyses of Reading Abilities II; WISC III: Wechsler Intelligence Scale for Children III. WRMT-PC: Woodcock Reading Mastery Test-R, Passage Comprehension.

This study investigates the growth of reading comprehension in young first (L1) and second (L2) language learners in a relatively consistent orthography (Norwegian). The main aim is to further our understanding of oral language and decoding as possible causal influences on the development of early reading comprehension skills. According to The Simple View of Reading, variability in reading comprehension in L1 and L2 learners is determined by decoding and linguistic comprehension skills and their interaction (Hoover & Gough, 1990). Decoding is the mapping of print to sound and linguistic comprehension refers to the processes involved in understanding words and the texts they compose. As L2 learners on average have a poorer command of the language they are learning, their linguistic comprehension will suffer and they will therefore be at risk of reading comprehension difficulties. Adequate reading comprehension skills are crucial for virtually all aspects of formal education as well as for full participation in society.

A number of concurrent studies have shown that both decoding and oral language skills predict reading comprehension at different points in development and that these skills differ between L1 and L2 learners. Longitudinal studies, on the other hand, are scarce but are essential for understanding how reading comprehension skills develop. Longitudinal studies can tell us about the rate of growth, whether rates of growth differ between L1 and L2 children and whether variations in decoding and oral language skills can predict differences in growth rate. Such understanding will, in turn, be important for planning interventions to help prevent the development of reading comprehension problems in L2 learners. Theoretically, it seems likely that having a limited vocabulary in the language in which children are being taught will lead to severe problems in understanding texts. A reader may be able to tolerate some unknown words and still be able to infer the meaning of a text with the help of text-level constraints. However, if too many words are unknown it is likely that comprehension will break down altogether (Carver, 1994). This implies that L2 learners will have comprehension difficulties if they have to read the same books and texts as their L1 counterparts who have better vocabulary knowledge. If vocabulary knowledge represents a bottleneck in reading comprehension, then variations in vocabulary knowledge may predict rates in the growth of reading comprehension. It is also implicit in The Simple View of Reading that decoding will play a more important role in limiting reading comprehension in younger children, while in older children, for whom decoding is more automatized, linguistic comprehension will be a more important factor (Hoover & Gough, 1990). This suggests that the gap in reading comprehension between L1 and L2 learners will increase as decoding skills become better developed and vocabulary skills come to play a

Conflict of interest statement: No conflicts declared.

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

more dominant role in determining individual differences in comprehension. The current study seeks to explore these issues by following 288 (198 L1 and 90 L2) beginning readers over a period of 18 months.

The children in the current study are learning to read the relatively consistent Norwegian orthography. As decoding skills are acquired faster in consistent (e.g., Norwegian, Dutch, German, Spanish) than in inconsistent (e.g., English) orthographies (Seymour, Aro, & Erskine, 2003), reading comprehension skills should also be acquired faster. However, as most studies focus on English language learners, more studies are warranted in consistent orthographies.

Most of the cross-sectional studies that have examined reading and oral language skills in L1 and L2 learners show that L2 learners have equivalent decoding (for reviews see August & Shanahan, 2006; Snow, Burns, & Griffin, 1998), but poorer oral language skills than L1 learners (e.g., Carlisle, Beeman, Davis, & Spharim, 1999; Proctor, Carlo, August, & Snow, 2005). This is seen in inconsistent as well as in consistent orthographies (e.g., Verhoeven, 2000). However, effects of orthographic consistency can be found in studies investigating reading comprehension skills: Some studies suggest that L2 English learners' reading comprehension skills do not differ from L1 learners in the first two years of school (Manis, Lindsey, & Bailey, 2004), but after that the L1 learners start to fall behind (Nakamoto, Lindsey, & Manis, 2007). In more consistent orthographies the differences appear to be already present in second grade (Verhoeven, 2000). Both decoding and oral language skills have been shown to be concurrent predictors of reading comprehension in several studies. These studies support The Simple View of Reading in that decoding predicts more variance in reading comprehension in earlier (e.g., Manis et al., 2004) than in later grades (e.g., Proctor et al., 2005), while the opposite pattern is found for oral language skills.

The few longitudinal studies that have compared the growth of reading comprehension skills in L1 and L2 learners find mixed results. Verhoeven (2000) found that the gap between L1 and L2 Dutch language learners remained the same across first and second grades on one reading comprehension measure while it decreased on another. In Droop and Verhoeven (2003) the gap decreased, increased and stayed the same on three different measures in third and fourth grade. Finally, Nakamoto et al. (2007) found that the gap in reading comprehension skills started to increase between L2 English learners and a national standardization from third grade on. However, as both Droop and Verhoeven (2003) and Verhoeven (2000) note, ceiling effect might have caused the decrease between the two groups found on some of the measures. When it comes to predicting the development of reading comprehension skills, Droop and Verhoven (2003) found that both oral language and decoding (weak) skills in third grade predicted later reading comprehension (after controlling for earlier levels of reading comprehension skills) in both L1 and L2 Dutch language learners. However, this pattern was not found in Verhoeven (2000) or Nakamoto et al. (2007). In Verhoeven (2000), oral language and decoding did not predict later reading comprehension beyond earlier reading comprehension skills in L1 or L2 language learners. Further, Nakamoto et al. (2007) found that oral language skills among Spanish-speaking English learners were negatively related to the growth in reading comprehension – an opposite pattern from that of Droop and Verhoeven (2003). In summary, existing longitudinal studies give an inconsistent picture and more research is needed to establish the patterns of growth in reading comprehension skills in L1 and L2 children.

Second language learners' L1 decoding and oral skills might contribute to their L2 reading comprehension skills. However, the evidence on this issue is mixed: While some studies have found no significant relationship between measures of these L1 skills, beyond the impact of corresponding L2 predictors (Bialystok, McBride-Chang, & Luk, 2005; Manis et al., 2004; Lindsey, Manis, & Bailey, 2003), other studies have found that at least a small part of the variance in L2 reading comprehension can be accounted for by L1 decoding and vocabulary beyond their L2 counterparts (Carlisle et al., 1999; Proctor, August, Carlo, & Snow, 2006).

The current study seeks to increase our understanding of the possible causal influences of oral language and decoding skills on the growth of reading comprehension. Our study has three major aims: 1) to confirm that L2 readers have deficits in vocabulary, but not decoding skills at the beginning of the study; 2) to clarify the form and rate of growth in reading comprehension in L1 and L2 learners; 3) to clarify the extent of any possible causal relationships between decoding and vocabulary and the growth of reading comprehension. Because differences in educational background and/or different types of reading tests may have contributed to the heterogeneity in outcomes between earlier studies (e.g., Droop & Verhoeven, 2003) we have, as a control, included a measure of maternal education and used two very different tests of reading comprehension (Neale Analyses of Reading Abilities II and Woodcock Reading Mastery Revised Passage Comprehension) that avoid both floor and ceiling effects.

Method

Participants

Two hundred and eighty-eight Norwegian second-grade children (134 girls, 154 boys) were recruited four months after formal reading instruction had started. Of these, 198 children (93 girls, 105 boys) had Norwegian

as their first language (L1) and 90 (41 girls, 49 boys) had Urdu as their first language and Norwegian as their second language (L2). None of the children had any reported learning disabilities at the beginning of the study. The average age was 7 years and 6 months for the L1 and L2 samples (*SD* = 3.50 months and 3.73 months respectively). The two samples were recruited from the same schools in eastern and southern parts of Oslo and informed consent was obtained from the parents. The instructional language at school was Norwegian and both groups were being taught to read in Norwegian. Seventy-five percent of the L2 parents reported that their children could not read more than a few words in Urdu at the end of the study. Among those who were reported to read at the sentence (15%) or at the text level (10%), 79% had learned it at home. The L2 parents also reported that they spoke both Norwegian and Urdu with their children (but slightly more Urdu) and that their children spoke more Norwegian than Urdu with both their siblings and their friends. Some 96% and 91% of the fathers and the mothers (respectively) of the Norwegian L1 children were employed and the corresponding figures were 90% and 40% for the L2 sample. It should also be noted that two-thirds of the Urdu-speaking population in Norway are marrying transnationally, most usually to people from Pakistan (Statistics Norway, 2009). This probably strengthens Urdu as a spoken language at home even if most of the children in the study can be considered to be second/third-generation immigrants. In the L1 sample attrition was .5%, 7.6% and 8.6% at Times 2, 3 and 4 respectively and in the L2 sample the corresponding attrition rates were 3.3%, 6.7% and 8.9%. The main cause of attrition was children moving out of a school's catchment area.

Design and procedure

The children were tested on four occasions over a period of 18 months (6-month intervals). All testing was done individually (except for Raven which was done in small groups) in school and the tests were given in a fixed order to all participants. As the current study is part of a larger study that addresses a wide range of research questions, only variables that specifically relate to the current research questions are presented here.

Tests and materials

At Time 1, 10 different tests were used to measure reading comprehension, word decoding, vocabulary breadth, vocabulary definitions and maternal education; only the reading comprehension tests were re-administered at Times 2–4.

Reading comprehension was measured by a Norwegian translation of the Woodcock Reading Mastery Test-R, Passage Comprehension (WRMT-PC; Woodcock, 1989) and the Neale Analysis of Reading Ability II (NARA II; Neale, 1997). The WRMT-PC consists of 67 short passages of increasing difficulty where one word is missing. The children read the passages silently and provided the missing word orally. All children started with the first item and discontinued after six consecutive incorrect responses. The NARA II consisted of 6 stories of increasing difficulty. The children were asked to read each story aloud and answer 4 questions from the first story and 8 questions from each of the other 5 stories. All questions were asked by the test administrator and the test was discontinued after the number of decoding errors specified in the manual was reached.

Word decoding was measured by a Norwegian translation of the Test of Word Reading Efficiency (TOWRE) forms A and B (Torgesen, Wagner, & Rashotte, 1999). The children read as many words as they could in 45 seconds from a list of 104 words

Vocabulary breadth was measured by a Norwegian translation of the first 144 words of the Peabody Picture Vocabulary Test III (PPVT) forms A and B (Dunn & Dunn, 1997). Testing was done according to the test manual except that all children started at Set 3 (ages 6–7). In addition, an Urdu translation of the British Picture Vocabulary Scale II (BPVS II; Dunn, Dunn, Whetton, & Burley, 1997) was administered for the L2 sample only. Here all children started at Item 1 in Set 1. The BPVS II was back-translated by an independent translator. A translation of the BPVS II was used instead of the translation of the PPVT to prevent direct learning effects by using an identical test in another language.

Vocabulary definition was measured by the Vocabulary test from the Norwegian translation of Wechsler Intelligence Scale for Children III (WISC III; Wechler, 2003) and by the Word definition test from the Danish Ability Scales (DEP; Elliott, 1996).

Nonverbal abilities were measured by Raven Standard Progressive Matrixes sets A, B, and C. Standard group administration was followed (Raven, Raven, & Court, 1988).

Maternal education was measured by asking the mothers to rate their educational level in the following categories: no formal education (score = 0), 1–6 years (score = 1), 7–9 years/junior high school (score = 2), 10–12 years/senior high school (score = 3), 1–4 years of collage/university (score = 4) and 5 years of more at a university (score = 5). Eighty-two percent of the L2 sample and 81 percent of the L2 sample completed this questionnaire.

Results

Differences between L1/L2 learners

Means, standard deviations, Cronbach's alphas and Cohen's *d* for the independent variables are shown in Table 1. As can be seen, L1 learners had higher scores than the L2 learners on all independent variables (vocabulary, non-verbal abilities and maternal education) except for those measuring decoding. Further, the reliability for most variables was good. The correlations between all variables can be seen in electronic Appendix 1.

Growth of reading comprehension

The mean (observed) growth curves of the two reading comprehension tasks are plotted in Figure 1. Both measures had good reliabilities at all time

Table 1 Means, standard deviations, Cronbach's alpha and Cohen's d for all independent variables

	L1 Norwegian			L2 Urdu			Diff. L1–L2
Independent variable	*M*	*SD*	α	*M*	*SD*	α	*d*
Age in months Time 1	90	3.50	–	90	3.73	–	.00
PPVT III Form A	99.16	14.25	.96	62.16	15.23	.94	2.54
PPVT III Form B	97.79	17.09	.95	62.28	14.40	.94	2.18
WISC III Vocabulary	18.55	4.19	.72	13.19	3.46	.67	1.35
DEP Definitions	12.89	2.90	.71	8.24	3.00	.71	1.59
TOWRE A	27.67	15.05	.96[a]	28.31	16.30	.97[a]	-.04
TOWRE B	26.07	17.06	.96[a]	26.23	16.52	.97[a]	-.01
Mother's Education	3.82	.77	–	2.89	1.24	–	.99
BPVS II Urdu	–	–	–	37.77	18.25	.96	–
Raven non-verbal abilities	20.25	5.96	.87	14.23	5.34	.85	1.04

Note: Maximum for PPVT III A and B = 144; Maximum for WISC III Vocabulary = 62; Maximum for DEP Definitions = 39; Maximum for Mothers Education = 5; Maximum for BPVS II = 144; Maximum for Raven non-verbal abilities = 36; [a] correlation between TOWRE A and TOWRE B.

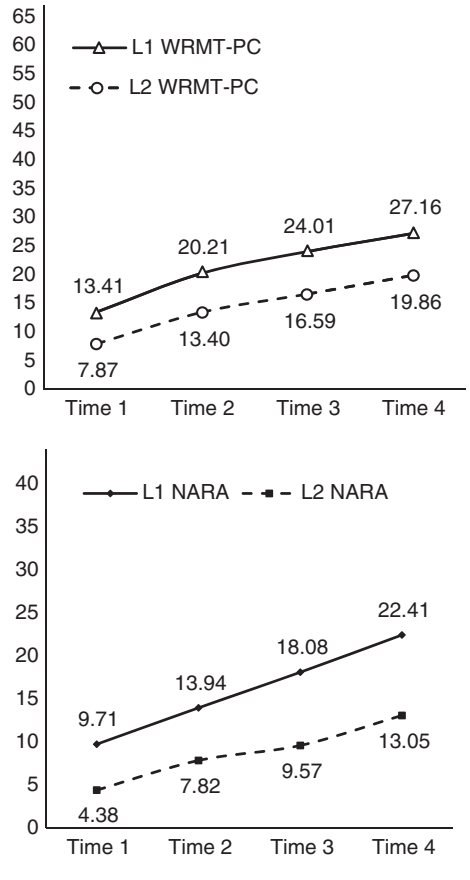

Figure 1 Mean Growth of the Reading Comprehension Skills in Both L1 and L2 Learners

points both in the L1 sample (α range .87–.92) and in the L2 sample (α range .77–.88). As we are interested in predicting and comparing the starting level and the growth rates of reading comprehension in the L1 and L2 children, two-group growth models were estimated. In these growth models we estimated both intercepts (the level of reading comprehension at Time 1) and slopes (the rate of growth between Time 1 and Time 4). However, as indicated by the shape of the growth curves in Figure 1, the growth pattern was different for the two reading comprehension tests. This meant that we had to estimate rates of growth separately for the two different measures (Ferrer, Balluerka, & Widaman, 2008).

The growth in WRMT-PC scores was best represented by nonlinear growth curves (the growth deviated from a straight line so the factor loadings of the two middle time points were freely estimated) that were equal over groups, $\Delta\chi^2$ (2) = 2.04, $p > .10$. The growth on the NARA test was best represented by a linear growth function in the L1 group and by a nonlinear function in the L2 group. A significant difference between this model and a model with both curves constrained to be linear, $\Delta\chi^2$ (2) = 9.73, $p < .01$, confirmed that the growth trajectories were not equal in the two groups. This makes it sensible to further compare the groups on both intercepts (starting level) and slopes (growth) on the WRMT-PC test, but only on the intercepts on the NARA test (since the growth is defined in different ways here).

The results of this comparison are shown in Table 2. The L1 learners had higher starting scores on both the reading comprehension tests and faster growth than the L2 learners on the WRMT-PC.[1] It should also be noted that even if different shapes of the NARA slopes between the L1 and L2 learners made it nonsensical to compare the full growth rate on this measure, a simpler model (testing whether the simple difference scores between Time 1 and Time 4 were equal over groups) showed clearly that the gap between the two groups increased between Time 1 and Time 4 on this measure as well, $\Delta\chi^2$ (1) = 26.44, $p <. 01$.

[1] There was a greater variability around the intercept means in the L1 group than in the L2 group on both tests, suggesting that the L1 learners with the most developed reading comprehension skills scored higher than the most skilled L2 learners.

Table 2 Parameter estimates of growth in reading comprehension and differences between the L1 and L2 samples

	Reading comprehension measure					
	WRMT-PC			NARA II		
	L1 sample	L2 sample	$\Delta\chi^2$ (1)	L1 sample	L2 sample	$\Delta\chi^2$ (1)
Intercept mean	13.476**	7.842**	47.892**	9.696**	4.385**	79.766**
Slope mean	4.520**	3.923**	5.396*	4.188**	2.927**	–
Intercept variance	51.754**	21.612**	10.710**	31.701**	8.990**	25.905**
Slope variance	1.207	2.568**	1.626	2.240**	1.667**	–
Corr. intercept–slope [a]	–.363**	–.452**	.177	–.126	.074	–
Time 1 res. variance	.212**	.101	–	.113*	.207	–
Time 2 res. variance	.218**	.505**	–	.273**	.403**	–
Time 3 res. variance	.205**	.306**	–	.251**	.367**	–
Time 4 res. variance	.203**	.342**	–	.174**	.268**	–
Model fit	χ^2 (8) = 13.32, p < .101, CFI = .992, TLI = .989, RMSEA = .068 (90% CI = .000–.130), SRMR = .054			χ^2 (8) = 8.38, p < .397, CFI = .999, TLI =.999, RMSEA = .018 (90% CI = .000–.100), SRMR = .049		

Note: All parameter values are unstandardized except for the correlation between intercept and slope and the residual variance of the observed variables; [a] = the correlation between initial status (intercept) and later growth (slope); *p < .05; **p < .01.

Predicting the growth of reading comprehension

To predict variations in reading comprehension two multi-group models were estimated (one for NARA and one for WRMT-PC). Here, both intercepts (initial skills) and slopes (later growth) of the reading comprehension measures were predicted from mothers' educational level, nonverbal-abilities, and vocabulary and decoding skills measured at Time 1 (composite scores were used for the four vocabulary and the two decoding measures). The four predictors were then entered hierarchically, with vocabulary and decoding skills being entered alternately at steps 3 and 4. These models were estimated using Cholesky factoring (de Jong, 1999) and fitted the data well for both the NARA, χ^2 (24) = 33.01, p < .104, CFI = .992, TLI = .981, RMSEA = .051 (90% CI = .000–09), SRMR = .039, and the WRMT-PC, χ^2 (24) = 40.63, p < .02, CFI = .989, TLI = .966, RMSEA = .069 (90% CI = .029–105), SRMR = .049. An illustration of the model can be seen in Figure 2 and the results of these hierarchical regressions are shown in Table 3.

As can be seen, decoding and vocabulary predicted the starting level (intercept) of the two reading comprehension measures after accounting for the effects of all other variables in both the L1 and the L2 groups. When entered at the last step, vocabulary predicted the growth (slope) on the NARA test in both groups, but only among the L2 learners on the WRMT-PC. In addition, decoding was a negative predictor of the growth in WRMT-PC. However, closer analyses revealed that this last prediction was confounded with the fact that poor decoders (or non-decoders) to a certain degree

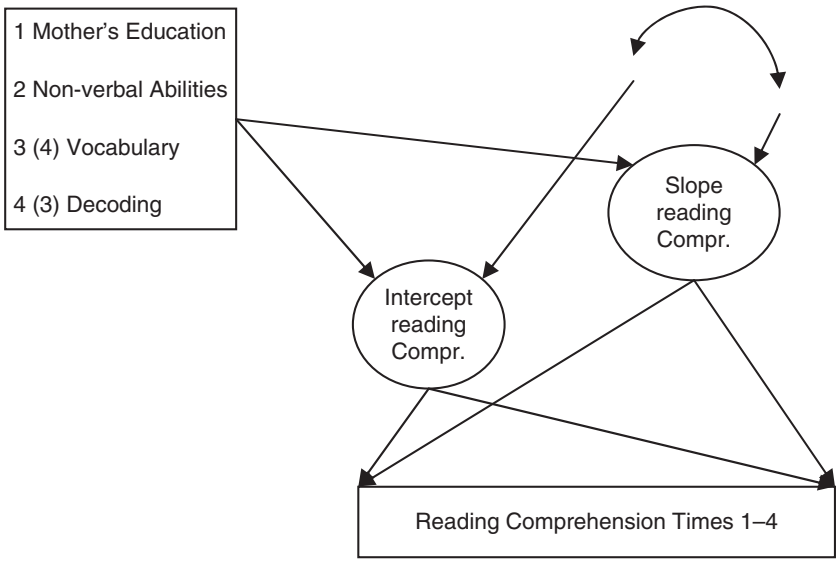

Figure 2 The Hierarchical Regression Models Predicting Beginning Reading Comprehension (intercept) and Further Growth (slope) in Reading Comprehension from Mother's Educational Level, Raven IQ, Word Decoding and Vocabulary in both L1 and L2 Children

Table 3 Incremental R^2 of the hierarchical regression models predicting beginning reading comprehension (intercept) and further growth (slope) in reading comprehension from mother's educational level, Raven IQ, word decoding and vocabulary in both L1 and L2 children

	NARA				WRMT-PC			
	L1		L2		L1		L2	
	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
1 Mother's education	.002	.054*	.010	.127*	.036*	.037[a]	.000	.114*
2 Raven non-verbal abilities	.228**	.030[a]	.036	.124**	.260**	.028[a]	.146**	.002
3 Word decoding	.262**	.009	.555**	.020[a]	389.**	.081[a]	.511**	.022[a]
3 Vocabulary	.278**	.061*	.303**	.083*	.168**	.004	.163**	.167**
4 Word decoding	.171**	.002	.401**	.047[a]	.306**	.092*[a]	.407**	.063[a]
4 Vocabulary	.187**	.054*	.148**	.109**	.085**	.015	.059**	.209**
R^2	.677	.147	.749	.380	.771	.161	.717	.347
Corr. intercept–slope	−.256		−.192		−.201		−.549*	

Note: [a]= negative beta weight; * = $p < .05$; ** = $p < .01$.

caught up with their peers who were better decoders to begin with.

Finally, vocabulary in Urdu only made a small (2%) and marginally significant ($p = .053$) unique contribution to initial reading comprehension on the NARA test. This was estimated with a single-group version of the model shown in Figure 2, with Urdu vocabulary skills included at the last (fifth) step. However, vocabulary in Urdu was not a predictor of the WRMT-PC intercept or the growth of comprehension skills as measured by either test. Thus, there is little support for the idea that L1 vocabulary is a causal influence on the development of L2 reading comprehension skills.

Are group differences in reading comprehension mediated via vocabulary and decoding skills?

Further analyses revealed that differences between the L1 and L2 learners, on both initial and later assessments (growth) of reading comprehension, could be fully accounted for by initial differences in vocabulary (but not decoding) skills. Thus, there was no difference between a model conditioned on (predicted from) one predictor at time, where the intercept and slope means were fixed to be equal, and a corresponding model where the means were freely estimated. This was the case for vocabulary, $\Delta\chi^2$ (2) = 3.82, $p > .10$, but not for decoding, $\Delta\chi^2$ (2) = 10.40, $p < .01$, on the WRMT-PC test. Also, the difference between the L1 and L2 learners' growth in reading comprehension (WRMT-PC) could not be attributed to the initial differences in initial reading comprehension $\Delta\chi^2$ (1) = 5.40, $p < .01$. The difference on the NARA intercept was fully mediated by vocabulary, $\Delta\chi^2$ (1) = .32, $p > .10$, but not by decoding, $\Delta\chi^2$ (1) = 14.77, $p < .01$.

Discussion

Our longitudinal study has revealed a number of critical findings concerning the similarities and differences between the growth in reading comprehension between L1 and L2 learners. First, the L1 learners had better initial reading comprehension skills and faster growth in these skills over time. Both of these differences could be fully attributed to initial differences in vocabulary skills. No differences were found between the groups on decoding skills. Second, beginning reading comprehension skills were predicted by vocabulary and decoding skills in both L1 and L2 learners. Third, vocabulary predicted further growth in reading comprehension among both the L1 and L2 learners on the NARA test (but only among the L2 learners on the WRMT-PC test). Finally, vocabulary in Urdu made a small and only marginally significant independent contribution to the prediction of initial reading comprehension on the NARA test.

The finding that L2 learners had poorer reading comprehension skills than the L1 learners from the earliest stages of learning to read is consistent with the Dutch studies (Droop & Verhoeven, 2003; Verhoeven, 2000) and in conflict with several studies of English language learners (Manis et al., 2004; Nakamoto et al., 2007). The reason for this discrepancy might reflect the consistency of the languages that have to be learned combined with the role decoding skills play in reading comprehension at this early stage. Decoding seems to be the dominant component of beginning reading comprehension and this component takes more time to learn in inconsistent orthographies (like English) than in consistent orthographies (like Norwegian and Dutch). Further, as decoding does not seem to interact with language, the finding that L1 and L2 learners do not differ on beginning reading comprehension in inconsistent orthographies might be because decoding is isomorphic with reading comprehension for a longer period in these orthographies. Later on, when decoding is mastered at a sufficient level, it is the semantic language component that is causing the variance between groups and, as semantic language skills differ between L1 and L2 learners,

reading comprehension will also differ. This interpretation is consistent with the findings of Nakamoto et al. (2007) who found that English language learners started to fall behind their L1 counterparts on reading comprehension from third grade. It is also consistent with the finding of the current study that the L2 learners performed more poorly than the L1 learners on vocabulary and that this difference in vocabulary skills fully explained the mean difference between the two groups in both initial reading comprehension and its later growth.

Our finding that the L2 learners showed slower growth in their reading comprehension skills than their L1 counterparts is also in agreement with the English studies showing that L2 learners tend to fall behind in reading comprehension when their decoding skills reach a certain level (e.g., Nakamoto et al., 2007). Thus, as decoding skills become more proficient the differences in vocabulary skills between L1 and L2 learners become increasingly important for reading comprehension. As a result, L2 learners will, in this period, have a slower growth in reading comprehension than L1 learners. Additionally, having a rich vocabulary might facilitate the learning of new words through reading. Thus, having a rich vocabulary to begin with might lead to an even richer vocabulary later on which, in turn, leads to a further increase in rate of growth in reading comprehension, beyond the effects caused by decoding becoming less important over time. This is consistent with the finding that vocabulary predicted the growth in reading comprehension.

It is also important to note that decoding and vocabulary seems to predict the two reading comprehension tests somewhat differently (Table 3). Closer analyses, where reading comprehension (both intercept and slope) was predicted from either decoding or vocabulary, confirmed that the WRMT was significantly more dependent on decoding, $\Delta\chi^2$ (4) = 20.54, p <. 01 (but not vocabulary, $\Delta\chi^2$ (4) = 5.88, p >. 10) than the NARA. This was so for the L1 learners, $\Delta\chi^2$ (2) = 9.72, p <. 01, but not for the L2 learners, $\Delta\chi^2$ (2) = 1.12, p >. 10. The fact that short passage cloze tests like the WRMT taps more decoding skills than longer text-based open-ended tests like the NARA is important knowledge for practitioners and in accordance with earlier findings (Keenan, Betjemann, & Olson, 2008; Nation & Snowling, 1997). Short passage cloze tests might be more dependent on decoding skills because failing to decode a single crucial word might have more severe consequences here than in longer open-ended tests where there are more words to lean on (Keenan et al., 2008).

The finding that vocabulary skills were a more important predictor of the growth of reading comprehension in L2 than in L1 learners, together with the fact that the children's initial vocabulary skills could explain all the difference in reading comprehension between the groups (both initial and later growth), suggests that L2 learners should be given help in developing their L2 oral vocabulary skills from an early age. Early intervention studies indicate that vocabulary skills may be supported through vocabulary-focused instruction to improve reading comprehension. For example, Bowyer-Crane et al. (2008) found that children with weak oral language skills at school entry who received an oral language intervention showed an advantage on measures of vocabulary at the end of the intervention period as well as 5 months later. In another study Aukrust (2007) found that the qualities of vocabulary use by teachers to children in preschool predicted L2 vocabulary in first grade, with control for concurrent preschool vocabulary skills. The L2 learners in the current study had poorer non-verbal abilities than the L1 learners and non-verbal abilities did predict the development of growth on the NARA for the L2 learners (Table 3). Both these facts might contribute to moderate the effects of a language intervention program for these L2 learners. Therefore, while some evidence exists concerning the value of specific instructional approaches to vocabulary learning, future research is needed to address the value of more comprehensive approaches to promoting vocabulary learning in both L1 and L2 learners. This research should also include nonverbal abilities and see whether these skills are able to moderate the L2 learner's response to language treatment.

Supplementary material

The following supplementary material is available for this article:

Appendix 1. Correlations between all variables for the L1 and the L2 samples above and below the diagonal respectively (Word document)

This material is available as part of the online article from:

http://www.blackwell-synergy.com/doi/abs/10.1111/j.1469-7610.2009.02185.x

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Correspondence to

Arne Lervåg, Institute for Educational Research, University of Oslo, Norway. P.O. Box 1092 Blindern, NO-0317 Oslo, Norway; Tel: +47 92448054; Fax +47 2285425; Email: a.o.lervag@ped.uio.no

Key points

- It is well known that first (L1) language learners have better reading comprehension skills than second (L2) language learners.
- How the development of reading comprehension skills differs between L1 and L2 learners and what predicts this development is less well known.
- Vocabulary appears to be a critical predictor of the early development of reading comprehension skills in both L1 and L2 learners.
- The limitations in vocabulary skills in the L2 learners seem sufficient to explain their lag in developing reading comprehension skills. Initial differences in vocabulary skills explain all the difference between L1 and L2 learners on both initial and later growth in reading comprehension
- Clinically, it is important to give L2 learners rich opportunities to learn the language of instruction from an early age.

References

August, D., & Shanahan, T (Eds.). (2006). *Developing literacy in second-language learners*. Report on the national literacy panel on language-minority children and youth. Mahwah, NJ: Lawrence Erlbaum.

Aukrust, V. G. (2007). Young children acquiring second language vocabulary in preschool group-time: Does amount, diversity, and discourse complexity of teacher talk matter? *Journal of Research in Childhood Education*, *22*, 17–38.

Bialystok, E., McBride-Chang, C., & Luk, G. (2005). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology*, *97*, 580–590.

Bowyer-Crane, C., Snowling, M., Duff, F.J., Fieldsend, E., Carroll, J.M., Miles, J., Götz, K., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry*, *49*, 422–432.

Carlisle, J.F., Beeman, M.M., Davis, L.H., & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics*, *20*, 459–478.

Carver, R.P. (1994). Percentage of unknown vocabulary words in texts as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior*, *26*, 413–437.

de Jong, P.F. (1999). Hierarchical regression analysis in structural equation modeling. *Structural Equation Modeling*, *6*, 198–211.

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

Dunn, L., & Dunn, L. (1997). *The Peabody Picture Vocabulary Test-Third Edition*. Circle Pines, MN: American Guidance Service.

Dunn, L.M., Dunn, L.M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale, version II*. Windsor: NFER-Nelson.

Elliott, C.D. (1996). *Dansk Evneprøve* [Danish Ability Scales]. Copenhagen: Dansk Psykologisk Forlag.

Ferrer, E., Balluerka, N., & Widaman, K.F. (2008). Factorial invariance and the specification of second-order latent growth models. *Methodology*, *4*, 22–26.

Hoover, W.A., & Gough, P.B. (1990). The simple view of reading. *Reading and Writing*, *2*, 127–169.

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

Lindsey, K., Manis, F.R., & Bailey, C.E. (2003). Prediction of first-grade reading in Spanish- speaking English-language learners. *Journal of Educational Psychology*, *95*, 482–494.

Manis, F.R., Lindsey, K.A., & Bailey, C.E. (2004). Development of reading in grades K-2 in Spanish-speaking English-language learners. *Learning Disabilities Research and Practice*, *19*, 214–224.

Nakamoto, J., Lindsey, K., & Manis, F.R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing*, *20*, 691–719.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity of current measures of reading skills. *British Journal of Educational Psychology*, *67*, 359–370.

Neale, N.D. (1997). *Neale Analysis of Reading Ability – Second Revised British Edition (NARA II)*. Windsor, UK: NFER-Nelson.

Proctor, C.P., August, D., Carlo, M., & Snow, C. (2006). The intriguing role of Spanish vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology*, *98*, 159–169.

Proctor, C.P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*, 246–256.

Raven, J., Raven, J.C., & Court, J.H. (1988). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Section 3: The Standard Progressive Matrices. Oxford: Oxford Psychologists Press.

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, *94*, 143–174.

Snow, C.E., Burns, S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Statistics Norway. (2009). *Who do immigrants in Norway marry?* Retrieved July 14, 2009 from http://www.ssb.no/vis/english/magazine/art-2006-10-13-02-en.html

Torgesen, J., Wagner, R., & Rachotte, C. (1999). *Test of Word-Reading Efficiency*. Austin, TX: PRO-ED.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*, 313–330.

Wechler, D. (2003). *Administration manual* (Norwegian version). San Antonio, TX and Assessio Norway: The Psychological Corporation.

Woodcock, R.W. (1989). *Woodcock Reading Mastery Tests Revised*. Circle Pines, MN: American Guidance Service.

Manuscript accepted 21 August 2009