Journal of Research in Reading
UKLA The United Kingdom Literacy Association
Journal of Research in Reading, ISSN 0141-0423
Volume 40, Issue S1, 2017, pp S107–S124
DOI:10.1111/1467-9817.12078

The impact of L2 German on component processes of reading

Pauline Schröter and Sascha Schroeder
Max Planck Institute for Human Development, Berlin, Germany

Background: In Germany, there is a substantial gap in reading literacy between monolingual children and their L2-speaking peers. Nevertheless, it is still unclear where these performance differences are rooted.
Methods: We investigated children of grades 5, 6 and 7 with comparable socio-economic status, who completed a battery of tests assessing their linguistic and executive functioning skills as well as their reading performance on the letter, word, sentence and text level.
Results: Whereas L1 speakers showed better linguistic skills, there was no difference between groups in executive functioning. After controlling for individual differences on each level of reading, groups only differed in text comprehension. This effect, however, disappeared when participants' linguistic skills were additionally controlled.
Conclusions: In sum, results show that reading problems in L2 speakers cannot be attributed to deficits on specific component processes of reading, but to a lack of linguistic skills, which negatively affects reading comprehension.

Highlights

What is already known about this topic

- L2 students lack behind their native speaking peers on reading literacy achievement even after controlling for differences in SES
- The process of reading involves low-level and high-level components, which each have to be mastered for successful reading comprehension
- L2 speakers are often found to differ from L1 speakers in linguistic as well as executive functioning skills, which both are known to be relevant for the process of reading

What this paper adds

- The gap in reading comprehension between L1 and L2 speakers can be ascribed to differences in their text representation and model refinement skills
- Differences in reading performance are moderated by linguistic skills, which are generally weaker in but not confined to L2 speakers
- High proficiency in a second language alone does no lead to improved executive functioning skills, which is why findings on bilingual advantages cannot be generalised.

Implications for theory, policy or practice

- Interventions to bridge the gap in reading literacy between L1 and L2 students should target at their vocabulary knowledge and listening comprehension skills
- Educational policy should expand its focus from improving L2 reading acquisition towards promoting poor readers' development of their linguistic skills
- For the interpretation of research findings on bilingualism special emphasis should be placed on sample selection criteria

It is commonly accepted that individuals who have to start reading in their second language (L2) face a more difficult task than native speakers. Findings from international studies such as the Program for International Student Assessment (PISA) show that in many countries, L2 children make up a big part in the proportion of students that fail to reach a level of reading proficiency assumed to be necessary in modern society. In Germany, the gap in literacy achievement between L1 and L2 speakers has repeatedly been found to belong to the most pronounced amongst all OECD countries (OECD, 2001; OECD, 2010). As a result, L2 reading acquisition has become a central issue of German educational policy, and a variety of actions have been implemented to promote L2 students' reading skills. However, while there is no dearth of research on the question which socio-cultural background factors associated with L2 speakers lead to lower reading literacy, hardly any study focused on performance differences in more detail with regard to the process of reading. In fact, understanding written language is the result of a number of different component processes, ascribed to the levels of letter, word, sentence and text processing, and it is still unclear where performance deficits for L2 speakers are rooted. In contrast, it is well established that L1 and L2 speakers often differ in their linguistic and cognitive abilities, which, in turn, have been shown to impact the process of reading. The present study aims to disentangle potential causes for the performance gap between L1 and L2 German speakers by assessing their proficiency on specific component processes of reading, while controlling for socio-cultural background factors that are associated with lower reading literacy. At the same time, the impact of reading-relevant skills known to vary in their degree of proficiency between groups, namely linguistic and executive functioning skills, are investigated. The goal is to better understand the nature of reading problems in L2 speakers, so that measures to improve their reading skills can be targeted more efficiently in the future.

Reading comprehension in L2 German speakers

In PISA 2009, the gap in reading performance between L2 German-speaking children and their native peers equalled more than 1 year of formal schooling (OECD, 2010). Students' socio-economic status (SES), that is, their parents' level of education and the investments they are able to make towards their children's education, is often suggested as a major factor of explanation (e.g. Goldenberg et al., 2006). However, even after controlling for SES, which was operationalised within PISA through structural characteristics of SES as measured by the International Socio-Economic Index (Ganzeboom et al., 1992), the difference between groups continued to be statistically significant (e.g. Baumert & Schümer, 2001). L2 students also reported to have access to fewer cultural resources (e.g. books) and participate in fewer communicative practices (e.g. talking to their parents about a book they read), which are referred to as process characteristics of SES. Analyses revealed that,

as opposed to structural characteristics, SES process characteristics were significantly related to their lower reading literacy (Müller & Stanat, 2006). When controlling for these variables, Müller and Stanat found that it was the fact whether or not students spoke German at home that affected their reading performance. They concluded that it was the lack of opportunities for L2 students to develop their German skills that accounted for their lower scores on the reading literacy scale. In other words, performance discrepancies in reading between L1 and L2 speakers seem to originate in differences directly related to language processing. In order to successfully target measures at these differences, however, it is necessary to look at the process of reading in more detail.

Items measuring reading literacy are designed in a way that students need to demonstrate high-level text comprehension skills. Within PISA, for instance, reading literacy is defined as "an individual's capacity to understand, use, reflect on and engage with written texts" (OECD, 2010, p. 14). To achieve this, however, a number of component processes have to be mastered (e.g. Perfetti et al., 1996). In a first step, the visual features of print need to be translated into individual letters and syllables that make up a word – a process that is often referred to as orthographic processing (Grainger, 2008). A very robust observation in research on word recognition is that letters are easier to be identified when they are part of a word rather than of an unpronounceable letter string (Baron & Thurston, 1973). This so-called word superiority effect is because of top-down feedback from the mental lexicon, which speeds up the processing of words. In the next step, associated semantic and phonological representations are retrieved from the lexicon – a process called lexical access (Coltheart et al., 2001). After a word is successfully recognised, it needs to be organised and integrated with other words of a sentence in order to compute meaning from their interrelations – a complex process that is described as parsing (Mitchell, 1994). Finally, inferences that relate each sentence to previous information given in the text as well as to the reader's world knowledge need to be generated in order to construct a coherent situation model of the text (Zwaan & Radvansky, 1998). Some of these processes are obligatory (low-level components such as orthographic processing and lexical access), whereas others are optional (high-level components such as parsing and the construction of a situation model, McKoon & Ratcliff, 1992). However, although all of them are highly automatised, there are great differences between readers in their efficacy and the amount of cognitive capacity required for their execution (Perfetti, 1985). From this, it follows that differences in reading comprehension between L1 and L2 speakers could originate in any of these component processes. Accordingly, it is crucial to investigate students' reading proficiency beyond comparing them solely on their reading comprehension skills.

Determinants of L2 reading

A frequently reported difference between L1 and L2 speakers shown to be relevant for the process of reading is their vocabulary knowledge. Research on bilingualism has demonstrated that individuals who speak two languages have smaller vocabularies in each of their languages relative to monolingual speakers of either language. Assumingly, lexicalised concepts are distributed across a bilingual's two languages, such that some words are known in one language, some in the other and only some in both (Oller, 2005). It is widely accepted that a smaller vocabulary in L2 may cause difficulties in students' reading acquisition in an L2 environment (Verhoeven, 2000). Results of a study by Limbird (2007), who compared determinants of reading comprehension in L1 and L2 German-speaking

elementary school students, suggested that the impact of vocabulary on reading was even more pronounced for L2 than for L1 speakers. Naturally, a second determinant that is likely to differ between groups are the skills associated with the usage of words, such as making inferences about the relationship between two notions. Thus, it is not surprising that tasks reflecting applied vocabulary knowledge, such as many verbal intelligence subtests, serve as strong predictors for reading comprehension in L1 (e.g. Oakhill et al., 2003) as well as in L2 (e.g. Koda, 1989). A third ability that is frequently elaborated on with regard to reading comprehension in L2 speakers is listening comprehension. According to the *simple view of reading* (Hoover & Gough, 1990), there is a multiplicative link between orthographic processing and listening comprehension skills, which are both assumed to impact reading comprehension. While orthographic processing is more important in the beginning of reading acquisition, listening comprehension becomes more significant in later grades, when decoding is automatised. Conducting a longitudinal study with second and fourth graders, Babayiğit and Stainthorp (2011) found that listening comprehension skills emerged as the most powerful predictor for later reading comprehension performance. In a study with Dutch elementary school students, Droop and Verhoeven (2003) did not only demonstrate that reading comprehension was strongly affected by listening skills but also that the effect was more pronounced in L2 compared with L1 speakers. Taken together, L1 and L2 students have been shown to differ in vocabulary knowledge, verbal intelligence and listening comprehension, which, in turn, are known to affect reading comprehension. In the following, these variables are referred to as participants' linguistic skills.

Another reading-relevant difference that could arise between L1 and L2 speakers is related to cognitive processes that are involved when using two languages instead of one. Bilinguals have repeatedly been found to outperform monolinguals on their executive functioning skills (e.g. Bialystok, 2001) – a set of cognitive control mechanisms that regulate human thought and action (Miyake et al., 2000). After a vast body of research has demonstrated that the two languages of a bilingual are both active when one of them is being used, this is often attributed to a bilingual's need to select one language while inhibiting the other (Green, 1998). Advantages for bilinguals over monolinguals on executive functioning have been shown on a number of different paradigms and populations (Bialystok, 2001). However, most research reporting bilingual benefits has been conducted with individuals who have acquired two languages from birth and use them equally on a daily basis. Especially after some recent studies have failed to replicate these findings (for a review, refer to Paap, 2014), it is unclear whether a bilingual advantage can be found in L2 speakers, who – even if they speak both of their languages fluently – are not balanced in their manner of acquisition or use of both languages. The three executive core functions frequently postulated in literature (Miyake et al., 2000) are the shifting between tasks or mental sets ('shifting'), the updating and monitoring of working memory representations ('updating') and the inhibition of prepotent responses ('inhibition'). After Just and Carpenter (1992) argued that individual differences in the efficiency of language comprehension were related to differences in working memory, a large number of studies have investigated the impact of cognitive skills on the different levels of reading. Fedorenko et al. (2004), for instance, demonstrated that sentence processing and arithmetic processing overlapped in the use of working memory resources, indicating that reading on the sentence level does not solely rely on verbally related skills. Linking the Stroop Color–Word Test (SCWT; Stroop, 1935), which measures shifting and inhibition processes, to a number of personal skills, Johnson et al. (2003) found that correlations between the SCWT and

reading and spelling measures were particularly high with regards to reading on the word level. Comparing young and older adults on their letter discrimination performance, Thapar et al. (2003) observed slower and less accurate responses in older adults, which they attributed to the overall decline of cognitive capacities with advancing age. In sum, there is plenty of evidence that reading performance does not exclusively depend on linguistic skills. After Babayiğit and Stainthorp (2011) found a contribution of working memory to reading comprehension, they highlighted the need to conduct further research on the impact of executive functioning skills on reading. In a study identifying predictors for reading acquisition in L1 and L2 German speakers, Duzy et al. (2013) could show that a test on nonverbal intelligence, encompassing executive skills, was the best predictor for orthographic processing speed in L2 but not in L1 speakers. From this, it follows that the gap between L1 and L2 students' reading performance might be driven by differences in executive functioning. Hence, in addition to participants' linguistic abilities, we assessed their proficiency on shifting, updating and inhibition, which in the following are referred to as executive skills.

The present study

Expanding research on the differences in German reading comprehension reported for L2 students compared with their native peers (e.g. Baumert & Schümer, 2001), the present study was conducted with two goals in mind. Our first goal was to compare groups of L1 and L2 German-speaking students of comparable SES, focusing on their linguistic and executive skills as well as on their reading performance. Based on the results of Baumert and Schümer (2001), we measured process characteristics of SES and additionally assessed students' reading behaviour to account for potential interrelations. Component processes of reading were ascribed to the levels of (1) letter, (2) word, (3) sentence and (4) text processing. By assessing participants' performance separately for each level, we wanted to identify the specific locus of reading differences between groups. Our second goal was to investigate whether group differences found on specific levels were actually because of participants being L1 or L2 speakers (L2 status) or rather a result of the disparity between L1 and L2 speakers in their linguistic or executive functioning skills. If the latter is the case, performance differences should occur as a function of participants' linguistic or executive functioning skills and irrespective of L2 status. To test this hypothesis, we performed a series of hierarchical regression analyses for each component process while controlling for the impact of previous components and reading-relevant skills.

Method

Participants

Native speakers were recruited from a pool of volunteers using the database of the Max Planck Institute for Human Development Berlin. This database is used as an in-house tool to collect personal data from people living in the local area, who are interested in scientific research and had registered to be contacted for participation in upcoming experiments. In order to recruit L2 German speakers, we contacted institutions known to be frequented by a high percentage of people with a native language other than German, such as bilingual schools or churches from different nationalities. In sum, 98 children attending grades 5, 6 or 7 (52% female, $M = 12$ years, $SD = 0.9$ years) participated in the study. All participants

had normal or corrected-to-normal vision and received financial reimbursement for their participation. L2 status was assessed using a language background questionnaire, including questions on participants' native language(s), home language(s), time of L2 acquisition and use of L2. Native speakers of German, who were not fluent in any other language, were classified as L1 speakers ($n = 64$), whereas individuals with a different native tongue, who spoke German fluently, were defined as L2 speakers ($n = 34$). All L2 speakers reported to have acquired German before the age of 6 years and to use both languages on a daily basis. Their native languages covered a broad range of orthographies, accounting in sum for 11 different L1s. Regarding the use of their languages at home, 79% of L2 speakers indicated to use both languages, whereas 21% reported to only use their native language at home. Overall sample characteristics are given in Table 1.

To ensure that there were no SES-related differences between groups, participants completed a questionnaire on their socio-cultural background (adapted from PISA 2009), including ratings on a 5-point Likert scale. We measured process characteristics of SES by assessing children's cultural resources (i.e. how much time children spend with their parents going to the movies, museums, concerts, operas, theatres or public readings) and communicative practices (i.e. how often children talk to their parents about books, school, friends, politics, daily life or personal problems). To account for interrelations, we additionally measured reading behaviour by assessing children's reading pleasure (i.e. how much time they spend reading for pleasure on a daily basis) and reading motivation (i.e. how much they engaged in reading activities). As expected, groups did not differ in any of these measures (all ts < 2).

Measures

We administered a battery of tests to assess children's linguistic and executive skills, as well as their reading ability ascribed to each level. All tasks were visually presented on a computer screen, while participants used the keyboard to respond. In the following, measures are described for each construct separately.

Table 1. Sample characteristics of groups.

	German L1		German L2	
	M	*SD*	*M*	*SD*
N	64	—	34	—
Gender (in % female)	58	—	41	—
Age (in years)	12	0.84	12	1.00
SES process characteristics				
Cultural resources[a]	2.06	0.42	1.88	0.45
Communicative practices[a]	2.94	0.50	2.77	0.57
Reading behaviour				
Reading pleasure[a]	2.66	1.17	2.71	1.34
Reading motivation[a]	3.17	0.61	2.96	0.65

Note:
[a]Measured on a 5-point Likert scale.

a. Linguistic skills

Vocabulary knowledge. Vocabulary was measured using form A of a vocabulary knowledge test for German (CFT 20-R WS; Weiß, 1998), which comprises 30 written multiple-choice items. Each item comprised one target word in isolation (e.g. 'chaos'), and participants had to select a word with the same or a similar meaning from a set of five answer alternatives (e.g. '(a) confusion, (b) damage, (c) currency, (d) anger or (e) demonstration'). The number of correctly answered items served as the final score. Internal consistency was sufficient (Cronbach's $\alpha = .76$).

Verbal intelligence. We assessed verbal intelligence with form A of the verbal analogies subtest from a standardised test battery for German students of grades 4 to 12 (KFT 4-12+ R; Heller & Perleth, 2000). Participants were presented a written pair of words in isolation (e.g. 'fire' and 'hot'), followed by a target word (e.g. 'ice') and a set of five answer alternatives (e.g. 'pudding', 'melted', 'fridge', 'cold', 'white'). They were asked to select the word that together with the target word would reflect best the analogy represented by the word pair given. There were 20 items, and the sum of correct answers was used as the final score, which showed sufficient reliability ($\alpha = .78$).

Listening comprehension. We used a shortened version of a listening comprehension task developed for the German National Assessment Study on Languages at Secondary Level I (Köller et al., 2010). Participants were asked to listen to three short stories via headphones, one at a time, and to subsequently answer a set of written multiple-choice comprehension questions by using the keyboard. There were 44 items, and the sum of correct answers was used as the final score. The reliability of the test was sufficient ($\alpha = .6$).

b. Executive functioning skills

Shifting. Adapted from the Test of Everyday Attention for Children (Manly et al., 2001), we used an Opposite Worlds task to measure participants' ability to switch between mental sets. Children were presented with a list of 24 items showing the digits 1 and 2 in a randomised order. In the 'same world' condition, they were asked to read out loud the digits as quickly and accurately as possible. In the 'opposite world' condition, they had to say the opposite for each digit ('1' for 2 and '2' for 1). Four blocks were run in the order of same world, opposite world, same world, opposite world. The time taken to complete each block was recorded. The time difference between the first 'same' and 'opposite' blocks was used as the final score. As a measure of reliability, we computed the split-half correlation (odd-even-split, adjusted by Spearman–Brown), which was very high ($r = .89$).

Updating. Participants' updating skills were measured using a German version of the working memory task developed by Daneman and Carpenter (1980). Children were asked to read a set of unrelated sentences and judge whether they were true or false. Each sentence was five to eight words long and ended on a different word. Sentences were arranged in three sets of three, four, five and six sentences, increasing in set size, and displayed for 8000 ms each. At the end of a set, students had to recall the last word of each sentence and enter those words into a window on the computer screen. Out of a maximum of 64, the sum of correctly recalled words was used as the final test score, which was highly reliable ($\alpha = .84$).

Inhibition. We measured participants' ability to inhibit irrelevant information by employing a classic Stroop task (Heathcote et al., 1991). Stimuli were presented in either red-coloured or green-coloured upper case letters, and there were three presentation conditions: RED displayed in red or GREEN displayed in green (congruent), RED displayed in green or GREEN displayed in red (incongruent) and XXX or XXXXX in red or green (neutral). Participants were asked to decide as quickly and accurately as possible whether a stimulus was presented in green or red while ignoring its form. Each trial started with a fixation cross, followed by the stimulus after 1000 ms, which remained on screen until a response was given. Overall, participants completed 48 trials and the difference in mean correct RT between the neutral and incongruent condition was used as the final test score. Reliability, measured using the split-half correlation, was very high ($r = .9$).

c. Reading skills

Letter-based reading skills

Letter identification. We asked participants to identify double versus single letters. After a practice phase, they were presented 100 unpronounceable arrays of three or four letters (e.g. 'PPTZ') and had to decide as quickly and accurately as possible whether the target included a letter once or twice. In each trial, a fixation cross was displayed for 500 ms, followed by a target that remained on screen until a response was given. Mean correct RT was used as the final score. As expected, the split-half correlation was very high ($r = .96$).

Letter search: string. In a letter search task, participants were shown a single letter and asked to decide as quickly and accurately as possible whether a subsequently presented unpronounceable string of letters (e.g. 'cadzr') included the letter or not. Each trail started with a fixation cross for 500 ms, followed by the presentation of a single letter for 1000 ms, which then was replaced by a string of five letters until a response was given. The task comprised a practice phase and 72 trials, of which half included the respective letter and half did not. Mean correct RT served as the final score, which was highly reliable ($r = .75$).

Word-based reading skills

Letter search: word. In order to assess participants' speed to process words as a whole, we employed the same task as described in the preceding texts, keeping everything constant but the targets, which were German words of five letters (e.g. 'Pferd') instead of letter strings. Participants were presented a single letter for 1000 ms, which then was replaced by a word in isolation. Again, mean correct RT was used as the final score, which again was very high ($r = .89$).

Lexical decision. To measure lexical access, we used a standard lexical decision task, in which participants were presented a letter string in isolation and asked to decide as quickly and accurately as possible whether the target represents a real word or not. After participants completed a practice phase, 60 German words (e.g. 'krug') and 60 nonwords (e.g. 'bost'), ranging in length from four to six letters, were presented in a randomised order. Each trial began with a fixation cross for 500 ms, after which the target appeared and remained until a response was given. Mean correct RT served as the final score ($r = .93$).

Sentence-based reading skills

Sentence processing. We assessed participants' ability to process and understand sentences using the 'sentence' subtest of a standardised reading test for German students of grades 1 to 6 (ELFE 1-6; Lenhard & Schneider, 2006). Because this subtest is known to show strong ceiling effects in higher grades, following other language tests (e.g. ProDi-L, Richter et al., 2012), we measured children's response speed. Participants were asked to complete 23 sentences as quickly and accurately as possible by choosing one out of five alternative words to fill in the blank. Response options differed in either their syntactic or semantic fit. Sentences were presented one at a time until a response was given. As expected, accuracy was very high (L1: $M = 96\%$, $SD = 5.7\%$; L2: $M = 95\%$, $SD = 13.7\%$), and participants' total time to complete the task was used as the final score ($\alpha = .96$).

Reading fluency. To measure participants' ability to integrate words within a sentence, we assessed their sentence reading fluency for connected text, using an excerpt from a popular German children's book (Cornelia Funke: 'Potilla'). The text consisted of 19 coherent sentences (216 words), varying in their degree of complexity. Participants were asked to read out loud as much of the text as possible until after 1 min the text disappeared. The total number of correctly read words was used as the final score. As there was only one value per person, reliability could not be calculated for this task.

Text-based reading skills

Comprehension speed. Participants' comprehension speed was assessed using the 'text' subtest of the ELFE test, which includes 20 stories consisting of two to six sentences. Again, response speed but not accuracy was measured. Each story was followed by a statement that had to be completed by selecting one out of four alternatives. As expected, accuracy was again very high (L1: $M = 93\%$, $SD = 7.6\%$; L2: $M = 89\%$, $SD = 6.3\%$), and we used participants' overall time to complete the task as the final score, which was highly reliable ($\alpha = .99$).

Comprehension accuracy. We used a shortened version of a reading comprehension test for German students of grades 7 to 9 (Verständiges Lesen 7–9; Anger et al., 1972). Participants were asked to read three texts of about 150 words, which were excerpts of longer stories with moderate difficulty. After they had read a text, eight comprehension questions along with four answer alternatives were shown one at a time and remained on screen until a response was given by using the keyboard. Out of a maximum of 24, the sum of correct answers was used as the final score ($\alpha = .79$).

Procedure

Children participated individually in a single self-paced lab session that lasted approximately 2 h. All tasks were presented on 15-in thin film transistor screens of IBM-compatible laptops, which recorded data automatically, while participants used the keyboard to respond. The order of tasks was lexical decision, letter search: string, letter search: word, vocabulary knowledge, sentence processing, letter identification, updating, inhibition, comprehension speed, reading fluency, verbal intelligence, shifting, comprehension accuracy and listening comprehension.

Results and discussion

Linguistic and executive functioning skills

Descriptive statistics of linguistic and executive skills for groups are given in Table 2. As expected, L1 speakers outperformed L2 speakers on all measures of linguistic skills. They scored significantly higher on the vocabulary task, $F(1, 96) = 5.91$, MSE = 101.13, $p < .05$, as well as on listening comprehension, $F(1, 92) = 5.37$, MSE = 100.53, $p < .05$). The performance gap on the verbal intelligence task also approached significance, $F(1, 96) = 2.5$, MSE = 42.49, $p = .1$). To ensure sufficient width of the construct, we *z*-transformed their scores and averaged them into a composite score reflecting participants' overall linguistic skills in German. Here and for all following composite scores, higher scores represent better performance. Overall, L1 speakers showed better linguistic skills compared with their L2-speaking peers (L1: $M = .16$, $SE = .11$; L2: $M = -.26$, $SE = .19$), $F(1, 92) = 4.82$, MSE = 3.51, $p < .05$.

In contrast to linguistic skills, there was no significant difference between groups in any of the tasks measuring executive functioning (all Fs < 1.2). Similar to linguistic skills, we averaged *z*-scores of the individual tasks into a composite score for participants' executive skills, which did not differ between groups (L1: $M = .01$, $SE = .08$; L2: $M = -.09$, $SE = .14$), $F(1, 82) = 0.52$, MSE = 0.20, $p = .48$.

Reading skills

Descriptive statistics for all reading measures are given in Table 3. Groups neither differed on the letter identification task nor on any of the two versions of the letter search task or the lexical decision task (all Fs < 1). However, L1 students performed better on the sentence processing task, which approached significance, $F(1, 95) = 3.03$, MSE = 13.27, $p = .08$. They clearly outperformed L2 students on the reading fluency task, $F(1, 92) = 6.67$, MSE = 7589, $p < .05$. Likewise, they exceeded their peers' performance on the

Table 2. Descriptive statistics of linguistic and executive skills between groups.

	German L1		German L2	
	M	*SE*	*M*	*SE*
Linguistic skills				
Vocabulary[a]	22.78	0.51	20.65	0.88
Verbal intelligence[a]	10.27	0.52	8.88	0.88
Listening comprehension[a]	30.62	0.54	28.42	0.95
Executive skills				
Shifting[b]	148.37	7.48	145.48	12.75
Updating[a]	36.93	1.07	35.03	1.81
Inhibition[b]	17.98	9.84	21.76	16.91

Notes:
[a]Raw scores.
[b]RT in ms.

Table 3. Descriptive statistics of measures for different levels of reading between groups.

	German L1		German L2	
	M	*SE*	*M*	*SE*
Letter-based skills				
Letter identification[a]	758.84	17.81	774.19	30.53
Letter search: string[a]	1028.26	31.06	1002.03	53.51
Word-based skills				
Letter search: word[a]	948.79	27.88	952.32	48.03
Lexical decision[a]	1154.26	51.42	1224.66	88.13
Sentence-based skills				
Sentence processing[b]	5.98	0.26	6.76	0.45
Reading fluency[c]	135.89	4.32	117.06	7.29
Text-based skills				
Comprehension speed[b]	16.99	0.84	20.77	1.46
Comprehension accuracy[c]	11.70	0.47	9.41	0.81

Notes:
[a]RT in ms.
[b]RT in min.
[c]Raw scores.

comprehension speed task, $F(1, 92) = 6.75$, MSE = 297.9, $p < .05$, as well as on the comprehension accuracy task, $F(1, 94) = 7.95$, MSE = 112.55, $p < .05$.

z-scores for the individual tasks were averaged into a composite score for each level. There was no difference between groups in their skills on the letter and the word level (both Fs < 1). On the sentence level, L1 speakers scored significantly higher than their L2-speaking peers (L1: $M = .15$, $SE = .12$; L2: $M = -.30$, $SE = .20$), $F(1, 92) = 5.05$, MSE = 4.34, $p < .05$. This performance gap was even more pronounced on the text level (L1: $M = .18$, $SE = .09$; L2: $M = -.36$, $SE = .17$), $F(1, 91) = 10.77$, MSE = 6.09, $p < .01$. Taken together, while groups did not differ on low-level component processes of reading, L2 speakers lagged behind their peers in their execution of high-level processes.

The impact of L2 status and reading-relevant skills on the process of reading

To investigate the nature of these high-level reading differences between L1 and L2 speakers, we performed a series of hierarchical regression analyses while controlling for participants' SES. The inter-correlations between all levels of reading, L2 status, linguistic and executive skills are given in Table 4. In a base model (M1), we investigated the inter-relationship between the different levels of reading. In the second model (M2), we examined the effect of L2 status on the different reading levels by entering it as a further variable into the model. In the third model, we additionally included participants' linguistic and executive skills (M3) to investigate whether the effects of L2 status remained significant even after controlling for these reading-relevant skills. Table 5 presents the regression coefficients, standard errors and *t*-values for each model separately.

Table 4. Inter-correlations between levels of reading, L2 status, linguistic and executive skills.

	1	2	3	4	5	6
1. Letter	—					
2. Word	0.85	—				
3. Sentence	0.42	0.60	—			
4. Text	0.25	0.41	0.74	—		
5. L2[a]	0.03	−0.02	−0.20	−0.28	—	
6. Ling. skills	0.14	0.25	0.52	0.73	−0.30	—
7. Exec. skills	0.32	0.27	0.32	0.25	−0.09	0.25

Note:
[a]Dummy-coded with L1 as reference category.

The relationship between component processes of reading

Results of M1 showed that participants' performance on each level of reading was best explained by their performance on the preceding level. Word-based reading skills were strongly affected by letter-based skills ($R^2 = .79$), while participants' sentence-level reading skills, in turn, were best predicted by their performance on the word-level ($R^2 = .50$). Text comprehension increased with better skills on the sentence-level ($R^2 = .58$), while the impact of letter-level and word-level performance was insignificant. In other words, each level of reading was best predicted by performance on the preceding level, while all second-order effects were not significant. Merely on the sentence level, there was a small negative effect of letter-based skills on participants' performance on the sentence level, which might be attributable to a suppressor constellation. To check for potential differential effects, we repeated all analyses by using one of the two subcomponents of each score separately as the regression criterion. Nevertheless, the pattern of results remained the same. Overall, results of M1 demonstrated that the base model was able to adequately describe our data.

The relationship between L2 status and reading

In M2, participants' L2 status was added to the base model. In contrast to mere comparisons between groups on each level of reading (as described under 'reading skills'), we now additionally controlled for their skills on all preceding levels. All effects of M1 remained significant, that is, participants' performance on one level continued to be highly predicted by their skills on the previous one. On the letter, word and sentence level, there was no main effect of L2 status (all ts < 1). Merely on the text level, L2 status still had a significant effect on participants' performance ($t = -2.4$). This means that the performance gap between L1 and L2 speakers on reading literacy is solely attributable to differences in processes ascribed to the text level.

The relationship between reading-relevant skills, L2 status and reading

In M3, we additionally controlled for participants' linguistic and executive skills. Like in M2, their performance on one level of reading remained to be a strong predictor for their performance on the next. Results of M3 showed significant main effects for linguistic skills on all levels of reading (all ts ≥ 3.2) but the letter level ($t = 1.0$). This indicates that linguistic

Table 5. Regression coefficients, standard errors and *t*-values for models M1, M2 and M3.

		M1			M2			M3		
Level	Variable	*β*	*SE*	*t*	*β*	*SE*	*t*	*β*	*SE*	*t*
Letter	L2[a]				0.03	0.21	0.19	0.12	0.22	0.50
	Ling. skills							0.13	0.13	1.03
	Exec. skills							0.40	0.17	2.31*
Word	Letter	0.77	0.05	14.77***	0.77	0.05	14.76***	0.74	0.05	13.64***
	L2[a]				−0.09	0.10	−0.85	−0.01	0.10	−0.05
	Ling. skills							0.16	0.06	2.61*
	Exec. skills							−0.05	0.08	−0.71
Sentence	Word	0.97	0.15	6.44***	0.96	0.15	6.35***	0.71	0.15	4.68***
	Letter	−0.34	0.14	−2.48*	−0.33	0.14	−2.39*	−0.24	0.13	−1.80
	L2[a]				−0.18	0.14	−1.23	−0.04	0.14	−0.28
	Ling. skills							0.26	0.08	3.21**
	Exec. skills							0.20	0.11	1.84
Text	Sentence	0.73	0.10	7.52***	0.69	0.09	7.32***	0.47	0.09	5.08***
	Word	0.08	0.16	0.48	0.08	0.16	0.53	−0.01	0.14	−0.01
	Letter	−0.16	0.12	−1.28	−0.15	0.12	−1.22	−0.01	0.11	−0.14
	L2[a]				−0.31	0.13	−2.44*	−0.10	0.11	−0.89
	Ling. skills							0.40	0.07	5.90***
	Exec. skills							−0.07	0.09	−0.84

Notes:
[a]Dummy-coded with L1 as reference category.
*$p < .05$.
**$p < .01$.
***$p < .001$.

skills have a strong impact on performance on almost all levels of reading. Most importantly, however, the effect of L2 status on the text level disappeared ($t = -0.9$). This suggests that differences in reading comprehension are not because of participants' L2 status per se but vary as a function of their linguistic skills. In other words, participants with poor linguistic skills will lag behind on reading comprehension irrespective of whether they speak German as their L1 or L2. Executive skills, on the other hand, showed a main effect on the letter level ($t = 2.3$), with no significant effects on the word, sentence and text level (all $ts < 1$). This indicates that solely on the most basic of all reading levels, executive skills have an impact. Beyond that level, differences in reading performance are determined by differences in linguistic skills. Taken together, results of M3 indicate that effects of L2 status on reading comprehension are moderated by reading-relevant skills. When controlling for participants' linguistic and executive functioning skills, L2 status had no significant impact on their performance on any of the levels involved in the process of reading.

Conclusion

The present study's aim was to investigate the nature of reading problems in L2 compared with L1 German-speaking students by identifying performance differences on specific levels of reading while controlling for the impact of reading-relevant skills. Sixty-four L1 and 34 L2 speakers of comparable SES completed a battery of tests assessing their linguistic and executive functioning skills as well as their reading performance on the letter, word, sentence and text level. In sum, findings suggest that the gap in reading literacy between L1 and L2 speakers is located on the text level and thus can be ascribed to differences in their text representation and model refinement skills. These differences, however, are moderated by linguistic skills, which are weaker in L2 compared with L1 speakers.

Groups particularly differed in their vocabulary knowledge, which is in line with most studies that have investigated language-related differences between L1 and L2 speakers (Oller, 2005). We could further show that these differences generalise to other aspects of L2 use, such as verbal reasoning and listening comprehension. Given that group differences occurred even after controlling for process characteristics of SES, we assume that this is because of L2 speakers' fewer opportunities to improve their L2 proficiency. This view is in accordance with Müller and Stanat (2006), who concluded that it is the fact whether or not students speak German at home that affects their reading performance. Thus, the amount of L2 speakers' use of German seems to highly contribute to their acquisition of reading literacy. In contrast, we did not find any difference between groups in their performance on any of the executive functioning tasks. This finding adds to the growing evidence that the bilingual advantage in executive functioning does not generalise to all individuals who are fluent in two languages (Paap, 2014). Most studies reporting a benefit for bilinguals investigated participants that had been exposed to two languages from birth and used both of them on a daily basis (Bialystok, 2001). Given that L2 students represent the group of bilinguals that is targeted by most educational interventions (Müller & Stanat, 2006), this points to the importance of sample selection criteria within research on the impact of speaking more than one language.

Regarding their skills on the specific levels of reading, L2 speakers did not differ from native speakers in their performance on the letter and word level. This is particularly interesting because many text comprehension difficulties are assumed to be caused by

inefficient low-level reading skills (Perfetti, 1985). Low-level reading skills, in turn, have been shown to strongly relate to the overall time children spend reading (Mol & Bus, 2011). Thus, given that our groups did not differ in the time they spend reading neither at school nor at home, the lack of significant differences on the lower levels might not be that surprising. In addition, we investigated rather advanced readers that had already mastered the initial stages of reading. It is assumed that low-level skills saturate earlier during development and become less important over time (Hoover & Gough, 1990). With respect to high-level reading skills, groups did not differ in their performance on the sentence level but on the text level. This indicates that L2 speakers might be more akin to 'poor comprehenders', who score lower on text comprehension tasks in spite of sufficient parsing abilities. Thus, it would be interesting to further investigate the impact of skills that have been shown to influence text comprehension over and above decoding and parsing, such as specific inference and integrations skills, students' knowledge about text structures and comprehension monitoring (Oakhill et al., 2003). Many of these subskills are likely to be influenced by students' linguistic abilities, that is, vocabulary knowledge and verbal reasoning, which is why they become generally more important in later stages of reading development.

Given that L1 and L2 students differed in their linguistic skills, it is not surprising that the L2 reading disadvantage disappeared after controlling for them. This is consistent with the general finding that the gap between L1 and L2 students in their reading performance diminishes if differences in their socio-cultural background are taken into account. However, while other studies still found significant group differences (Müller & Stanat, 2006), in the present study, the effects disappeared completely. This finding underlines the importance of assessing students' skills in different language domains, which might be a more powerful way to predict students' reading performance. In contrast to research that focused primarily on the effects of executive functions, our results show that if the impact of linguistic skills is additionally taken into account, executive skills only contribute to performance differences on the most basic level of reading. This, in turn, suggests that even if bilinguals had an advantage on executive functioning relative to their monolingual peers, this would not explain the gap in reading literacy as measured by PISA.

Turning to the educational implications of our study, it seems reasonable to conclude that if groups only differ in their performance on the text level, which, in turn, is explained by differences in linguistic skills, the gap in reading literacy could be closed by fostering these skills. Although our findings are limited to the specific set of measures we have used, our data suggests that interventions to improve L2 speakers' reading performance should target at their vocabulary knowledge and listening comprehension skills. This is in line with Babayiğit and Stainthorp (2011), who found listening skills to be the most powerful predictor for reading comprehension. To improve their skills, for instance, L2 students might benefit from a distinction between top-down as well as bottom-up approaches to teaching listening (for an overview, refer to Vandergrift, 2007). As our research has shown, however, these measures should be directed at all students with weaker linguistic skills, regardless of whether they are L1 or L2 speakers. Although linguistic deficits naturally emerge more often in children who have started to learn the respective language later in life, such problems are not confined to L2 speakers. Thus, educational policy should expand its focus from improving L2 reading acquisition towards monitoring beginning readers more closely and promoting poor readers' development of their linguistic skills.

References

Anger, H., Bargmann, R. & Voigt, M. (1972). *Verständiges Lesen 7–9 [Reading comprehension, Grade 7–9]*. Weinheim: Beltz.

Babayiğit, S. & Stainthorp, R. (2011). Modeling the relationships between cognitive–linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology*, 103(1), 169–189. DOI:10.1037/a0021671.

Baron, J. & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology*, 4(2), 207–228. DOI:10.1016/0010-0285(73)90012-1.

Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Eds.), *PISA 2000. Basiskompetenten von Schülerinnen und Schülern im internationalen Vergleich*. (pp. 323–407). Opladen: Leske & Budrich. doi:10.1007/978-3-322-83412-6_10

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press. doi:10.1017/cbo9780511605963

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204. DOI:10.1037/0033-295x.108.1.204.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19(4), 450–466. DOI:10.1016/s0022-5371(80)90312-6.

Droop, M. & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38(1), 78–103. DOI:10.1598/rrq.38.1.4.

Duzy, D., Ehm, J.H., Souvignier, E., Schneider, W. & Gold, A. (2013). Prädiktoren der Lesekompetenz bei Kindern mit Deutsch als Zweitsprache. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 45(4), 173–190. DOI:10.1026/0049-8637/a000093.

Fedorenko, E., Gibson, E. & Rohde, D. (2004). Verbal working memory in sentence comprehension. In *Proceedings of the 26th annual meeting of the Cognitive Science Society.*

Ganzeboom, H.B., De Graaf, P.M. & Treiman, D.J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56. DOI:10.1016/0049-089x(92)90017-b.

Goldenberg, C., Rueda, R.S. & August, D. (2006). Sociocultural influences on the literacy development. *Developing literacy in second-language learners*: *report of the national literacy panel on language-minority children and youth*. doi:10.1037/0012-1649.42.1.117

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, 23(1), 1–35. DOI:10.1080/01690960701578013.

Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, 1(02), 67–81. DOI:10.1017/s1366728998000133.

Heathcote, A., Popiel, S.J. & Mewhort, D.J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, 109(2), 340. DOI:10.1037/0033-2909.109.2.340.

Heller, K.A. & Perleth, C. (2000). *KFT 4-12+ R kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz Test.

Hoover, W.A. & Gough, P.B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127–160. DOI:10.1007/bf00401799.

Johnson, W., Bouchard, T.J., Jr., Segal, N.L., Keyes, M. & Samuels, J. (2003). The Stroop Color-Word Test: Genetic and environmental influences; reading, mental ability, and personality correlates. *Journal of Educational Psychology*, 95(1), 58. DOI:10.1037/0022-0663.95.1.58.

Just, M.A. & Carpenter, P.A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99(1), 122. DOI:10.1037/0033-295x.99.1.122.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, 22(6), 529–540. DOI:10.1111/j.1944-9720.1989.tb02780.x.

Köller, O., Knigge, M. & Tesch, B. (Eds.) (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster Waxmann Verlag.

Lenhard, W. & Schneider, W. (2006). *Ein Leseverständnistest für Erst-und Sechstklässler: ELFE 1-6*. Göttingen: Hogrefe.

Limbird, C.K. (2007). *Phonological processing*, *verbal abilities*, *and second language literacy development among bilingual Turkish children in Germany* (Doctoral dissertation, Freie Universität Berlin, Germany).

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P. & Robertson, I.H. (2001). The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample

and ADHD performance. *Journal of Child Psychology and Psychiatry*, 42(08), 1065–1081. DOI:10.1111/1469-7610.00806.

Mol, S.E. & Bus, A.G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267–296. DOI:10.1037/a0021890.

McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99(3), 440. DOI:10.1037/0033-295x.99.3.440.

Mitchell, D.C. (1994). Sentence parsing. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. (pp. 375–409). San Diego: Academic Press.

Müller, A.G. & Stanat, P. (2006). Schulischer Erfolg von Schülerinnen und Schülern mit Migrationshintergrund: Analysen zur Situation von Zuwanderern aus der ehemaligen Sowjetunion und aus der Türkei. In *Herkunftsbedingte Disparitäten im Bildungswesen*. (pp. 221–255). Wiesbaden: VS Verlag für Sozialwissenschaften.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A. & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. DOI:10.1006/cogp.1999.0734.

Oakhill, J.V., Cain, K. & Bryant, P.E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, 18(4), 443–468. DOI:10.1080/01690960344000008.

OECD (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris: OECD. doi:10.1787/9789264195905-en

OECD (2010). *PISA 2009. Assessment Framework: Key competencies in reading, mathematics and science*. Paris: OECD Pub. doi:10.1787/9789264062658-en

Oller, D.K. (2005). The distributed characteristic in bilingual learning. In *ISB4*: *Proceedings of the 4th International Symposium on Bilingualism* (pp. 1744–1749). Somerville, MA: Cascadilla Press.

Paap, K.R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology*, 26(3), 242–255. DOI:10.1080/20445911.2014.891597.

Perfetti, C.A. (1985). *Reading ability.* Oxford: Oxford University Press. doi:10.1017/s0033291700005304

Perfetti, C.A., Marron, A.M. & Foltz, P.W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties – processes and intervention*. (pp. 33–68). Hillsdale, N.J: Lawrence Erlbaum Associates.

Richter, T., Isberner, M.-B., Naumann, J. & Kutzner, Y. (2012). Prozessbezogene Diagnostik von Lesefähigkeiten bei Grundschulkindern. *Zeitschrift für Pädagogische Psychologie*, 26, 313–331. DOI:10.1024/1010-0652/a000079.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643. DOI:10.1037/h0054651.

Thapar, A., Ratcliff, R. & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, 18(3), 415. DOI:10.1037/0882-7974.18.3.415.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, 4 (4), 313–330. DOI:10.1207/s1532799xssr0404_4.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(03), 191–210. DOI:10.1017/s0261444807004338.

Weiß, R.H. (1998). *Grundintelligenztest Skala 2-CFT 20 mit Wortschatztest (WS) und Zahlenfolgentest (ZF): mit Handanweisung und Testheft*. Göttingen: Hogrefe.

Zwaan, R.A. & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162. DOI:10.1037/0033-2909.123.2.162.

Pauline Schröter is a predoctoral research fellow in the research group Reading Education and Development (REaD) at the Max Planck Institute for Human Development Berlin. She studied Psychology, TESOL and Teaching German as a Foreign Language and currently works on completing her PhD thesis on the development of visual word recognition in German bilinguals. Her research focuses on orthographic processing in beginning readers of L2 German as well as in German-English balanced bilingual children.

Sascha Schroeder received his PhD in psychology from the University of Cologne in 2008 and habilitated in 2011 at the Freie Universität Berlin. After working at the Max Planck Institute for Human Development Berlin and the University of Kassel, as well as holding guest professorships at the

University of Potsdam and the Freie Universität Berlin, he established the Max Planck Research Group REaD in 2012. He investigates the cognitive processes underlying written language acquisition and the cognitive determinants of reading development in childhood and adolescence. His research combines longitudinal and experimental approaches and uses linguistic as well as cognitive methods to analyse and simulate developmental data. He is a member of the International Max Planck Research School LIFE and supervises students at both the undergraduate and graduate levels.

Received 24 August 2015; revised version received 3 February 2016. Accepted 1 June 2016.

Address for correspondence: Pauline Schröter, Reading Education and Development, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: *schroeter@mpib-berlin.mpg.de*