International Journal of Educational Research Open 5 (2023) 100278

Contents lists available at ScienceDirect

International Journal of Educational Research Open

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

Matthew effect in vocabulary and reading: A comparison of good and average readers in Grade 1 to Grade 3

Martina Röthlisberger [a,*], Christoph Zangger [b], Britta Juska-Bacher [a]

[a] Bern University of Teacher Education, Institute of Research, Development and Evaluation, Fabrikstrasse 8, 3012 Berne, Switzerland
[b] University of Bern, Institute of Sociology, Fabrikstrasse 8, 3012 Bern, Switzerland

ARTICLE INFO

Keywords:
Reading comprehension
Average readers
Good readers
Vocabulary breadth
Vocabulary depth

ABSTRACT

There is quite a large body of research about the influence that vocabulary plays in poor comprehenders reading. Yet, little is known about the role of vocabulary in good comprehenders' reading skills. In the present study, third graders with a one-year advantage in reading comprehension ($N = 105$) were compared to average comprehenders ($N = 177$) in terms of vocabulary knowledge at three timepoints (Grade 1 to 3). Vocabulary was operationalized as vocabulary breadth (receptive and expressive) and vocabulary depth in form of semantic word knowledge. Latent growth models showed a superiority of good over average comprehenders both cross-sectionally and longitudinally in all assessed vocabulary skills. Moreover, the impact of vocabulary skills on sentence and text reading increased significantly from Grade 1 to Grade 3 in both groups. These results substantiate the increasing importance of vocabulary skills on reading comprehension yielding a Matthew effect.

1. Introduction

The ability to read texts is based on decoding sequences of letters to which the associated speech sounds must be assigned (Castles, Coltheart, Wilson, Valpied & Wedgwood, 2009). On the other hand, the sequence of speech sounds must be assigned a meaning (van Zeeland & Schmitt, 2013). Thus, there are two clearly different processes that enable comprehending reading. However, while decoding is a rather "technical" process, which depends on the successful neuronal recycling of certain cortical areas destined for face and object recognition (Dehaene & Cohen, 2007), reading comprehension is a complex skill based on different cognitive and linguistic abilities (Castles, Rastle & Nation, 2018). It is obvious and uncontested that access to word meanings forms the basis for reading comprehension – Perfetti, Landi and Oakhill (2005) called vocabulary "a bottleneck in comprehension" – reading comprehension also depends on other factors such as working memory or inference formation (Perfetti et al., 2005, p. 240). Moreover, vocabulary is composed of two different dimensions: breadth and depth. Vocabulary breadth refers to the number of words of which some aspect of meaning is known by a person (Anderson & Freebody, 1981: 92–93), vocabulary depth describes the knowledge of semantic subtleties and of the correct use of a word. While preschoolers' vocabulary breadth is predictive for later reading comprehension (e.g., Sénéchal, Ouellette & Rodney, 2006), vocabulary depth directly predicts reading comprehension in Grade 4 (Ouellette, 2006). However, the division into vocabulary breadth and depth is not undisputed. Vermeer (2001) argues that vocabulary breadth and depth are so highly correlated that they should not be operationalized as different constructs.

Despite a growing recent interest in the role of the different vocabulary components in children with poor reading comprehension (see Spencer & Wagner, 2018 for a meta-analysis), the influence of vocabulary in children with high reading comprehension skills has been neglected so far or considered only in the context of general cognitive abilities (Borella, Carretti & Pelegrina, 2010; Cain, 2006).

The focus of the present article is thus a differentiated analysis of the vocabulary's contribution on skilled readers reading comprehension in Grade 1 to 3. The central question is whether and to what extent the different vocabulary components such as vocabulary breadth (receptive and expressive) and semantic word knowledge of these children differ from those of average comprehenders. The resulting findings should help reinforce the strengths of poor comprehenders, in addition to eliminating their shortcomings.

* Corresponding author.
E-mail address: martina.roethlisberger@phbern.ch (M. Röthlisberger).

https://doi.org/10.1016/j.ijedro.2023.100278
Received 6 January 2023; Received in revised form 28 August 2023; Accepted 28 August 2023

1.1. Development of reading comprehension

1.1.1. Decoding phase

Reading involves the ability to arbitrarily match written symbols to speech sounds and synthesize them into larger units. Once the complete set of grapheme-phoneme correspondences is acquired, children are able to decode unknown words as well and retrieve the words meaning even if not all grapheme-phoneme matches may be correct (Castles et al., 2018). Vocabulary thus already plays a role at this early stage of reading, albeit a less important one than in the following phase (e.g., Quinn, Wagner, Petscher & Lopez, 2015).

In addition to vocabulary, the prerequisites for decoding are of course the knowledge of letters (Foulin, 2005), the ability to segment, synthesize and manipulate spoken language (phonological awareness; Anthony & Francis, 2005) and the ability to realize linguistic units quickly and accurately (rapid naming, RAN; Georgiou, Parrila, Cui & Papadopoulos, 2013).

1.1.2. Fluent reading phase

With increasing experience in decoding, automatisms generate a solid orthographic knowledge. According to Ehri (2005), this process contains several phases to culminate in the consolidated orthographic phase, where words are recognized by sight. The meaning of familiar words is retrieved directly from print, at least for familiar words (dual route cascaded model DRC; Coltheart, Rastle, Perry, Langdon & Ziegler, 2001).

The foundation for comprehensive reading is thus the knowledge of the meaning and the form of the individual word, a knowledge which depends on the lexical quality of the child's mental word representations (Perfetti, 2007) and is linked to vocabulary (Perfetti et al., 2005). If there is not at least a fragmentary knowledge of the word's meaning, it must be inferred from the context, which places high demands on the executive functions such as working memory and is hardly manageable for a child who has just made the transition to direct word reading (Cain, Oakhill & Lemmon, 2004). However, decoding skills still influence reading fluency and thus reading comprehension. In a recent study, Gentaz, Sprenger-Charolles and Theurel (2015) assessed first graders reading comprehension in relation to their decoding skills, which were measured by pseudoword reading. The sample was divided into three groups according to decoding skills. A higher impact of decoding and phonological awareness was found in the poor and average decoders compared to the good decoders. Interestingly, receptive vocabulary was a significant predictor of reading comprehension exclusively in poor and average decoders. The authors attribute this difference to an increased reliance on vocabulary in less skilled decoders.

With improving lexical quality, cognitive resources are freed and available for the comprehension of larger text chunks. Instead of reading word by word, multiple word strings are processed. Yet, understanding a larger text implies the knowledge of how words fit into a sentence, for instance, which grammatical morphemes mark a direct object. Obviously, there is a reciprocal link between morphological knowledge, word reading, and vocabulary, the latter potentially also affecting the direct link between morphological knowledge and reading comprehension through relational word knowledge such as finding or identifying synonyms (Nation & Snowling, 1998), antonyms and words of the same category (Nation & Snowling, 1999), which is stronger in good comprehenders compared to poor comprehenders.

1.2. The connection between vocabulary and reading comprehension

Assuming that reading comprehension depends on the understanding of word meaning, children's skills in vocabulary are to be predictive for the development of text comprehension (e.g., de Jong & van der Leij, 2002). However, the conception of vocabulary is far from being trivial. It is not only the sheer number of entries in a mental lexicon – operationalized as vocabulary breadth – but also the quality of these entries, the vocabulary depth (Read, 2004; Wesche & Paribakht, 1996). However, vocabulary depth has been defined in different ways. While Anderson and Freebody (1981) described vocabulary depth as lexical quality (how well a mental lexicon entry is known), Read (2004) described a total of three dimensions of vocabulary depth, i.e., in addition to the knowledge of the precise meaning of a word, the extent of the knowledge about a word including semantic and formal aspects, and the knowledge about the network in which a word is framed.

The central aspect to these definitions is that vocabulary depth has been shown to explain reading comprehension beyond vocabulary size. Quinn et al. (2015) for instance found that vocabulary predicted reading comprehension (but not vice versa) with vocabulary depth operationalized as semantic word knowledge. Assessing basic receptive vocabulary in Grade 1 and 2, using a more sophisticated assessment method in Grade 3 (finding synonyms based on sentences), Verhoeven, van Leeuwe and Vermeer (2011) showed that vocabulary was strongly associated to reading comprehension. Ouellette (2006) established a high influence of vocabulary depth – measured as word definitions – and a somewhat lower influence of receptive vocabulary breadth on reading comprehension in fourth graders, while expressive vocabulary breadth (picture naming) did not explain any further variance. Finally, in a study operationalizing vocabulary breadth as receptive vocabulary and word definitions, and vocabulary depth/fluency as knowledge of multiple meanings and attributes, Tannenbaum, Torgesen and Wagner (2006) found significant contributions of both vocabulary dimensions in the prediction of variance of reading comprehension in third graders.

In all these studies, reading comprehension was measured exclusively in terms of text comprehension, usually with cloze texts or questions to be answered about the text. However, for young children these paradigms might be too difficult or measure decoding instead of comprehension (Keenan, Betjemann & Olson, 2008). Following Perfetti's lexical quality hypothesis (Perfetti, 2007), it makes sense to step back and to assess lower hierarchy processes like word and sentence reading comprehension first (Lenhard, Lenhard & Schneider, 2017), and to control explicitly for decoding. De Jong and van der Leij (2002) for instance used a sentence-based paradigm for the first assessment (Grade 1, children's age about 7 years), which yielded high correlations with receptive and expressive vocabulary breadth.

1.3. Good and poor comprehenders

While many children develop average to good reading comprehension skills once their decoding abilities are consolidated, some of them exhibit reading comprehension issues despite adequate decoding. Almost one-third of the third graders in the United States reading comprehension scores below the basic level (Spencer & Wagner, 2018), in Germany about one-fifth of the fourth graders achieve below the third benchmark[1] of the Progress in International Reading Literacy Study PIRLS (Hußmann et al., 2017: 14).

At primary school age, children with good text comprehension typically differ from those with low text comprehension in verbal working memory (Cain et al., 2004), listening comprehension (Nation, Clarke, Marshall & Durand, 2004), inference making (Cain, Oakhill, Barnes & Bryant, 2001), monitoring (Cain et al., 2004), cognitive abilities, and vocabulary (Cain & Oakhill, 2006). Good and poor comprehenders do usually not diverge in phonological awareness (e.g., Cain, Oakhill & Bryant, 2000; Nation et al., 2004) nor text reading fluency and accuracy (Cain & Oakhill, 2007; Nation, 2005). In the context of the Dual Route Cascading Model (Coltheart et al., 2001), the absence of decoding variation between good and poor comprehenders may be

[1] Four benchmarks divide the performance scale into five sections, with the middle of the third section corresponding to the 50th percentile. Children who do not reach the third benchmark are therefore below the 40th percentile (Hußmann et al., 2017).

explained by an intact phonological pathway but a deficient direct path, pointing to semantic weaknesses. On the other side, several studies have shown lower performances in poor comprehenders reading irregular and pseudowords (Ricketts, Bishop & Nation, 2008), a finding that is consistent with Perfetti's Lexical Quality Hypothesis (Perfetti, 2007; Perfetti & Hart, 2002). According to his hypothesis, an entry in the mental lexicon includes semantic, orthographic, and phonological information, which depend on each other, implying that poor comprehenders are likely to have issues in the latter areas as well. For German, this theory is supported by a study investigating the link between the phonological, orthographic, and semantic representations to reading comprehension, where the impact of semantic aspects on reading comprehension was moderated by the phonological and orthographic quality of the lexical representations (Richter, Isberner, Naumann & Neeb, 2013).

Since oral vocabulary assessment mainly reveals aspects of the semantic knowledge, its inclusion may shed light on the structure of mental lexicon entries. However, many of the cited studies controlled for vocabulary but did not analyze it in more detail. Studies looking more closely at vocabulary in the context of low reading comprehension are somewhat sparser.

Nation, Snowling and Clarke (2007) tested 24 8 to 9-years-olds, 12 good and 12 poor comprehenders on their facility to learn new words. The children were taught nonsense words defined by four attributes and illustrations depicting the non-words. While in terms of attributing the correct phonological label to the pictures there was no difference between the groups, the poor comprehenders scored significantly lower in defining the non-words and in delayed recall. Moreover, the poor comprehenders scored at the lowest limit of the average range on a standardized expressive vocabulary breadth test. The authors conclude that poor comprehenders have difficulties to map phonological forms on their meanings and to consolidate knowledge about these mappings, which explains their low vocabulary scores when presented with more or less familiar words. In a well-designed small-scale study, Colenbrander, Kohnen, Smith-Lock and Nickels (2016) reported semantic difficulties in poor comprehenders of age 9 to 11 as a group in active and multiple-choice definition tasks and in listening comprehension, while reaching comparable results in reading accuracy and fluency. On an individual level, ten out of thirteen poor comprehenders scored below the control group in vocabulary breadth, definition tasks, conceptual semantics, and picture naming. The authors conclude that at least in their sample no direct relationship between semantic, orthographic, and phonological quality of lexical entries could be found.

All studies cited above examined the difference in vocabulary between good and poor comprehenders, defining poor comprehenders as having reading comprehension skills that lag behind their chronological age at least one year. However, some children have above-average reading skills including higher reading comprehension compared to their peers. Usually, this constellation is discussed in studies about reading skills and print exposure (e.g., Van Bergen et al., 2018) or in terms of cognitive development (Borella et al., 2010; Cain, 2006). Nevertheless, insights from the relations between reading comprehension, vocabulary and decoding in these children may be not only interesting per se but may also provide information that are beneficial for poor comprehenders. The question of whether a so-called Matthew effect also occurs between children with average reading comprehension and those who are above average is particularly intriguing. Stanovich (1986) defined the Matthew effect in reading as a "reciprocal relationship" between reading ability and vocabulary, which remains stable into adulthood (p. 379). In the present context, a Matthew effect can be described as following: children who have good reading comprehension skills experience higher gains in vocabulary, which in turn fosters their reading comprehension, while poor readers often do not understand what they read because their word reading skills are already limited (Cain & Oakhill, 2011). The gap between good and poor comprehenders in terms of vocabulary gets therefore larger over time because the poor comprehenders cannot incorporate the new words into their vocabulary. Cain and Oakhill (2011) have shown such an effect in children from 8 to 16 years but not in reading comprehension. Interestingly, in the reciprocal relationship between reading comprehension and vocabulary, there was only an interaction between receptive vocabulary and reading comprehension, but not for sight vocabulary. The authors explain this with the frequency with which the words presented in receptive vocabulary tests occur in print media. Expressive vocabulary was not included in this study.

To date, however, no study has ever been done to examine the impact of vocabulary breadth and depth on the reading comprehension of good and average comprehenders.

1.4. The present study

The purpose of the present study is to investigate the development of the interplay between reading comprehension and vocabulary in German speaking L1 children with average and above average reading comprehension skills between Grade 1 and Grade 3. The aim was to disentangle the influence of the different aspects of vocabulary on reading comprehension to identify those, which are particularly well developed in above average comprehenders. The sample was thus divided into two groups. Children with a one-year head start on their chronological age in terms of text reading comprehension were assigned to the "good comprehenders" group (GC), all other children were classified as "normal comprehenders" (NC). It should be noted that the total sample did not contain any children with a diagnosis of reading acquisition disorders, and that decoding faculties were controlled. Reading comprehension was operationalized as word, sentence, and text reading, vocabulary as vocabulary breadth (receptive and expressive) and semantic word knowledge. The developmental trajectories were modeled in a latent growth curve model based on the following hypotheses:

(1) The group assignment has an impact on reading comprehension in the context of all vocabulary variables

(1.a) at each single time point.
(1.b) on the development between Grade 1 and 3.

(2) The single vocabulary measures, i.e., receptive, and expressive vocabulary breadth, and semantic word knowledge have a differential impact on the reading variables. Due to the fundamental role of vocabulary depth on reading comprehension, we expected a larger effect of semantic word knowledge compared to receptive and expressive vocabulary breadth.
(3) The higher the reading comprehension requirements are (word, sentence, and text reading), the higher the influence of the vocabulary measures are.

2. Method

2.1. Participants

An initial sample consisted of 342 German speaking L1 children. With 33 children having dropped out during the first and 23 children during the second year of the assessment period, the attrition rate was 9.6% and 7.5% respectively. From the remaining 292 children, 10 could not be assessed because school access was denied due to COVID, resulting in a final sample of 282 children (163 girls, 119 boys). The 282 children were first assessed in Grade 1 (t1; *M* age = 7 years, 5 months, *SD* = 4.5), in Grade 2 (t2; *M* age = 8 years, 5 months, *SD* = 4.3) and in Grade 3 (t3; *M* age = 9 years, 7 months, *SD* = 4.9). The higher standard deviation at t3 is due to the extension of survey period because of COVID.

Classes were recruited mainly through teachers at partner schools of the university, who were contacted by mail. The parents of the children

were informed by e-mails or letters. Children without parental consent were not included. The study was approved in advance by the university's internal review board.

The origin of the participating 40 classes was well distributed among urban areas ($N = 11$), agglomerations ($N = 14$), and rural areas ($N = 15$) to achieve a representative socioeconomic background. For the coding of the socioeconomic status the ISEI-08 score of occupation was used according to the International Standard Classification of Occupations ISCO (Ganzeboom, Graaf & Treiman, 1992), which ranges between 10 and 89. The average socioeconomic status was rather high, $M = 56.94$, $SD = 14.25$, with a minimum of 23, a maximum of 89 and a range of 66.

The children were divided into two reading comprehension groups based on the following criteria: children whose third-grade standard score in text comprehension were above the 50th percentile according to the fourth-grade standard norms were classified as "good comprehenders (GC)" ($N = 105$, 64 girls, 41 boys), all other children as "normal comprehenders (NC)" ($N = 177$, 99 girls, 78 boys). The proportion of girls and boys was balanced, $\chi^2(1, N = 282) = 0.68$, $p = .41$. The two groups did not differ in age, all $Fs \langle 1.9$, all $ps \rangle 0.16$. In contrast, the socioeconomic status of the good comprehenders was higher, $M = 60$, $SD = 15.22$, than of normal comprehenders, $M = 55.12$, $SD = 13.36$, $F(1, 281) = 7.99$, $p = .005$, $\eta^2 = 0.03$. The normal comprehender group included a small number of below average comprehenders, i.e., children whose third-grade standard text reading scores were at least one year below their chronological age ($N = 40$, 17 girls, 23 boys). A repeated measures ANOVA showed that the between-subject effects of the 40 below average comprehenders did not differ from the average comprehenders ($N = 147$, 82 girls, 55 boys) in terms of vocabulary variables, all $Fs \langle 2.5$, all $Ps \rangle 0.118$, all $\eta^2 < 0.15$. Therefore, they were included in the group of average comprehenders.

2.2. Materials

At all three timepoints in Grade 1, 2 and 3, the following data were collected: word, sentence, and text reading, decoding, vocabulary breadth (receptive and expressive), and semantic word knowledge.

2.2.1. Reading comprehension

Reading comprehension was assessed with the standardized German reading test "ELFE II, Ein Leseverständnistest für Erst- bis Siebtklässler (ELFE II)" (Lenhard et al., 2017). The ELFE-II consists of three parts, a word, sentence, and text reading part.

Word reading. The 75 items of the word reading part are composed of a picture with 5 words each, the task is to underline the matching word to the picture. The distractor words are either graphemically or phonologically related to the target word and contain the same number of syllables. The order of presentation is according to the difficulty of the items. The duration is three minutes, the test is preceded by three practice trials.

Sentence reading. After two practice trials, 36 sentences with a cloze are presented, accompanied with five words each. The task is to select the correct word, which fits into the cloze. Different types of words are represented, nouns, adjectives, verbs, prepositions, and conjunctions. The distractor words are either visually, phonologically, or semantically related and belong to the same word class as the target. The time limit is three minutes.

Text reading. The text reading part consists of short texts followed by one or several questions relating to the text, resulting in 26 items. Each item is presented with a target and three distractor responses. The texts are either narrative or non-fiction texts that require simple information retrieval (easy), processing of cross-sentence references (intermediate), or inference formation (difficult). The text reading task is preceded by two practice trials and limited to seven minutes.

2.2.2. Decoding

The pseudoword reading subtest from the "Salzburger Lese- und Rechtschreibtest SLRT-II" (Moll & Landerl, 2014) was used to measure decoding ability. The children were presented with a list of 156 pseudowords, of which they are asked to read as many as possible correctly aloud for 1 min.

2.2.3. Vocabulary breadth (receptive)

The German version of the "Peabody Picture Vocabulary Test PPVT-4″ (Lenhard, Segerer, Lenhard & Suggate, 2015) was used to assess vocabulary breadth. Due to time constraints, only the odd item blocks were presented.

2.2.4. Vocabulary breadth (expressive)

Expressive vocabulary was measured with the expressive subtest (abbreviated form, items 16 to 55, for 7 to 9 years old children) of the "Wortschatz- und Wortfindungstest WWT 6–10″ (Glück, 2011). In the WWT, children have to produce the precise name of objects, actions, antonyms, and hyopernyms with the help of pictures.

2.2.5. Semantic word knowledge

Since there is no specialized test instrument for measuring semantic word knowledge in German, the 15 odd items of the definition task for measuring verbal intelligence from the German "Wechsler Intelligence Test HAWIK-IV" (Petermann & Petermann, 2007) were administered. The children were presented with a spoken word and asked to describe its meaning. The standard scoring system of HAWIK was extended by one level to better capture semantic knowledge, i.e., instead of three, four levels with a maximum of three points were applied. All responses were recorded, transcribed, and scored by independent raters. In case of differences in the two scores, a third rater scored again. Inter-rater reliability was decent, Cohen's kappa $\kappa = 0.69$, $SD = 0.15$ for Grade 1, $\kappa = 0.77$, $SD = 0.08$ for Grade 2 and $\kappa = 0.74$, $SD = 0.12$ for Grade 3.

2.3. Procedure

Data collection was assumed by project collaborators and trained students visiting the children at their classrooms. While at t1 the reading skills were tested individually, the reading assessments at t2 and t3 were groupwise. All other measures such as vocabulary were applied in an individual setting. Divided in two blocks, each individual test session had a duration of about 20 min. Data were collected at t1 and t2 between March and May 2018 and 2019 respectively. At t3, the regular data collection was interrupted in 2020 by COVID-19 after the assessment of 37 children. Once the schools were open again, there was only one person allowed in the classrooms. A staff member who brought along the required technical equipment (laptops, earphones, and a mobile hotspot) thus did the group testing on site, while the individual sessions were conducted online via Microsoft Teams®. The effects of online assessments were tested with a univariate analysis of variance with raw test scores (receptive and expressive vocabulary breadth, and semantic word knowledge) as dependent variables, and the factor assessment (onsite vs. online) as between-subjects variable.[2] The results showed that the children assessed online showed a slightly lower performance in expressive vocabulary, $M = 26.96$, $SD = 5.75$, relative to those tested onsite, $M = 29$, $SD = 4.06$, $F(1, 281) = 4.83$, $p = .029$, $\eta^2 = 0.06$ (expressive vocabulary). All other scores did not differ significantly, all $Fs < 3.7$, all $ps > 0.05$.

2.4. Data analysis

Latent growth curve models (LGCMs) were used to map the development of the two groups' reading skills in relation to the vocabulary

[2] Additionally, rapid automatized naming RAN, and phonological awareness were defined as dependent variables, as these were particularly influenced by the survey form.

and context variables across the three time points (Hertzog, Oertzen, Ghisletta & Lindenberger, 2008). Data analysis was conducted using the R package lavaan (Rosseel, 2012) for the latent growth curve models. Except for the descriptive statistics, for all analyses, the z-standardized values based on t1 were used, according to the formula $(X_{tn} - \mu_{t1}) / \sigma_{t1}$.

Table 3 illustrates the basic setup of the models estimated in this article. η_0 and η_1 depict the latent intercept and slope, respectively, that capture the development of individual reading scores over the three time points. They are characterized by an estimated mean, α, and a corresponding variance, ψ. Together, they predict the average change in the dependent variable, here the different reading scores (word, sentence, and text) across the three time points. The two latent constructs are usually allowed to be correlated, captured by ψ_{01}. The (residual) variances of the dependent variable is denoted by the different ε. Group is a time-invariant dichotomous variable, that allows us to investigate differences at the baseline and growth of students' reading scores for good comprehnders at t3 vs. average comprehenders. Finally, the different vocabulary measures (PPVT for vocabulary breadth (receptive), WWT for vocabulary breadth (expressive), and definitions for semantic word knowledge) are used as time-variant predictors of students' reading scores.

3. Results

3.1. Descriptive statistics

The mean values and the differences between the good (GC) and the normal comprehenders (NC) are documented in Table 1. In addition to the strongly increasing differences in all reading variables, growing disparity between the groups was mainly found in vocabulary breadth (receptive) and semantic word knowledge. Group differences in vocabulary breadth (expressive) showed a steeply rising trajectory between t1 and t2, which leveled off between t2 and t3 due to a slightly smaller gain in the good comprehenders group.

3.2. Correlations

Data distribution was controlled with Kolmogorv-Smirnov tests. As some of the data – mainly text reading – were not distributed normally, non-parametric Spearman correlations were calculated. The correlations for both groups and the three timepoints are shown in Table 2. Mainly the correlations between text reading and vocabulary variables got stronger over time in both groups. The correlations between word reading and all vocabulary variables remained low at all three timepoints. Sentence reading's correlations with all vocabulary variables on the other hand increased steadily in the good comprehenders group but not in the normal comprehenders. The correlations of the text reading task exhibited a more differential pattern: increasing values over time for both groups when correlated with receptive vocabulary breadth (although considerably higher for the good comprehenders) but decreasing in the good comprehenders when correlated with expressive vocabulary breadth. The correlations between text reading and semantic word knowledge were clearly increasing in good comprehenders, while they remained stable for normal comprehenders.

3.3. Latent growth curve model

Starting at the bottom of Table 3, we first note that the models fit the data well. All robust CFI and TLI values are above 0.95 and 0.9, respectively. With the exception of the LGCMs predicting word, sentence, and text reading using students' definition scores as a time-variant predictor, the RMSEA values do not exceed 0.08 and the SRMR is below or equal to 0.05 in all models (Hu & Bentler, 1999). To reduce the complexity of the model, decoding was related to word reading exclusively, since word reading best reflects decoding.

4. Discussion

In the present study, the differential influence of receptive and expressive vocabulary breadth and depth – operationalized as semantic word knowledge – on first to third graders' word, sentence, and text reading comprehension and its development between Grade 1 and 3 have been investigated by a latent growth curve model. The children were divided into two groups according to their text reading comprehension skills in good and normal comprehenders. Generally, the results showed – with some variation – an increasing vocabulary impact and a rather decreasing impact of decoding on all reading comprehension measures between Grade 1 and 3 beyond the groups.

In terms of between-subject differences, the good comprehenders persistently exhibited higher word, sentence, and text reading scores at the baseline (intercept), a result which could be expected. But moreover – indicated by the highly significant and positive effect on the latent slope – they were also found to make bigger gains in reading throughout the first three years of primary school. This is also applicable in relation with the vocabulary coefficients, which are always significant on the intercept and slope and positive, i.e., the good comprehenders have

Table 1
Descriptive statistics and t-tests of differences between normal (NC) and good comprehenders (GC) for Grade 1 to 3.

	t1			t2			t3		
Group Measure	NC *M (SD)*	GC *M (SD)*	Group Differences *F* (η^2)	NC *M (SD)*	GC *M (SD)*	Group Differences *F* (η^2)	NC *M (SD)*	GC *M (SD)*	Group Differences *F* (η^2)
Word Reading	17.53	23.45	52.03***	29.36	38.46	78.95***	42.66	54.31	118.52***
	(5.98)	(7.69)	(0.16)	(7.78)	(9.14)	(0.22)	(8.91)	(8.30)	(0.3)
Sentence Reading	3.70	7.04	68.54***	9.76	16.96	159.14***	17.02	24.96	194.30***
	(2.50)	(4.27)	(0.2)	(4.41)	(5.00)	(0.36)	(5.12)	(3.64)	(0.41)
Text Reading	1.78	4.12	73.28***	5.92	11.87	190.78***	11.01	18.77	402.70***
	(1.59)	(3.00)	(0.21)	(3.12)	(4.06)	(0.41)	(3.52)	(2.38)	(0.59)
T-Value Reading	41.66	50.71	85.69***	43.65	56.39	174.87***	45.90	59.96	264.28***
	(7.26)	(8.97)	(0.23)	(7.63)	(8.14)	(0.38)	(7.60)	(5.91)	(0.49)
Vocabulary Breadth (receptive)	140.01	145.97	6.40*	150.99	160.50	16.96***	164.46	176.08	21.16***
	(19.79)	(17.96)	(0.02)	(18.51)	(19.11)	(0.06)	(19.06)	(22.71)	(0.07)
Vocabulary Breadth (expressive)	16.98	19.83	13.74***	22.24	26.89	40.70***	25.60	29.24	30.43***
	(6.35)	(6.06)	(0.05)	(6.23)	(5.23)	(0.13)	(5.63)	(4.86)	(0.10)
Semantic Word Knowledge	14.20	14.90	1.43	19.40	21.07	5.64*	22.07	24.40	13.53***
	(4.87)	(4.68)	(< 0.01)	(5.85)	(5.46)	(0.02)	(5.40)	(4.65)	(0.05)
Decoding	17.46	23.62	53.17***	26.27	35.79	90.32***	31.22	42.18	94.02***
	(6.52)	(7.40)	(0.16)	(7.86)	(8.58)	(0.24)	(8.70)	(9.94)	(0.25)

Note: NC: normal comprehenders, $N = 177$; GC: good comprehenders, $N = 105$; Mean number of correct items with standard deviations in parentheses; F-values with η2 in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 2
Correlations between reading and vocabulary at all timepoints for both groups.

		t1						t2						t3					
		1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	Word Reading	–	.707***	.594***	−0.082	.058	−0.112	–	.806***	.547***	.085	.078	.149	–	.711***	.447***	.093	.179	.162
2	Sentence Reading	.578***	–	.797***	.052	.109	.064	.738***	–	.712***	.110	.121	.155	.732***	–	.693***	.151	.201*	.247*
3	Text Reading	.572***	.635***	–	.063	.220*	.168	.497***	.673***	–	.102	.154	.238*	.611***	.717***	–	.220*	.163	.247*
4	Vocabulary Breadth (receptive)	−0.010	.138	.029	–	.219*	.427***	−0.016	.114	.051	–	.285**	.489***	.019	.125	.137	–	.372***	.552***
5	Vocabulary Breadth (expressive)	.065	.128	.068	.423***	–	.413***	.026	.186*	.155*	.424***	–	.440***	.048	.145	.217*	.500***	–	.445***
6	Semantic Word Knowledge	.061	.152*	.175*	.350***	.360***	–	.144	.206**	.175*	.391***	.316**	–	.045	.112	.127	.380***	.520***	–

Note: NC: normal comprehenders (below the diagonal), $N = 177$; GC: good comprehenders (above the diagonal), $N = 105$; all correlation coefficients are Spearman's rank coefficients; * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3
Estimated coefficients and model fits for latent growth curve models.

	Word reading			Sentence reading			Text reading		
	Vocabulary Breadth (receptive)	Vocabulary Breadth (expressive)	Semantic Word Knowledge	Vocabulary Breadth (receptive)	Vocabulary Breadth (expressive)	Semantic Word Knowledge	Vocabulary Breadth (receptive)	Vocabulary Breadth (expressive)	Semantic Word Knowledge
Group differences									
on	0.31***	0.31***	0.30***	0.42***	0.42***	0.44***	0.37***	0.37***	0.37***
η_0 (intercept)	−4.45	−4.4	−4.32	−6.65	−6.74	−7.06	−7.03	−7.08	−7.19
on η_1	0.23**	0.23**	0.25***	0.43***	0.43***	0.44***	0.65***	0.64***	0.65***
(intercept)	−3	−2.88	−3.32	−8.36	−8.39	−8.73	−14.88	−14.45	−15.77
Time-variant: vocabulary									
at *t1*	−0.08	−0.03	−0.02	0.08	0.10*	0.07	0.03	0.04	0.12**
	(−1.64)	(−0.7)	(−0.38)	−1.71	−2.15	−1.61	−0.71	−0.94	−2.88
at *t2*	0.02	0.02	0	0.12***	0.13***	0.11***	0.07	0.11***	0.08*
	−0.52	−0.47	(−0.03)	−3.92	−3.47	−3.36	−1.86	−3.34	−2.25
at *t3*	0.06	0.09*	0.05	0.08**	0.11***	0.07*	0.07*	0.13***	0.06
	−1.87	−2.32	−1.58	−2.87	−3.28	−2.11	−2.32	−4.03	−1.95
Time-variant: Decoding									
at *t1*	0.53***	0.53***	0.53***						
	−9.61	−9.85	−9.69						
at *t2*	0.41***	0.42***	0.41***						
	−8.62	−8.83	−8.45						
at *t3*	0.48***	0.48***	0.48***						
	−10.29	−10.38	−10.19						
CFI (robust)	0.96	0.96	0.95	0.99	0.98	0.97	0.99	0.98	0.97
TLI (robust)	0.93	0.93	0.91	0.97	0.97	0.95	0.97	0.97	0.94
RMSEA (robust)	0.08	0.08	0.09	0.07	0.07	0.09	0.07	0.07	0.1
SRMR	0.05	0.05	0.05	0.04	0.04	0.05	0.04	0.05	0.05

Note: $N = 282$; the coefficients are standardized, empirical Z-values in parentheses. * $p < .05$, ** $p < .01$; *** $p < .001$.

overall higher values than the normal comprehenders. Hypothesis (1) is therefore correct with respect to both sub hypotheses, claiming that group assignment modulates reading comprehension with regard to all vocabulary variables. This contrasts with the results in Cain et al. (2011), where between-subject effects between good and poor comprehenders were shown only with respect to vocabulary.

The picture is somewhat more complex about the time-variant factors, namely the influence of different measures of vocabulary. On the word reading level, the impact of all vocabulary measures remained marginal up to Grade 3 once group differences and overall trajectories were accounted for. This suggests that the word comprehension part of the ELFE-II reading assessment primarily measures decoding abilities, which remain steady between Grade 1 and 3 (see Table 3). The task of retrieving the meaning of single words is apparently not a challenge for children with average and above average reading comprehension.

Differentiated patterns – but similar for both groups – can be found for sentence reading. For sentence reading, all three vocabulary measures at t1 had almost no influence on the children's reading scores. While the missing effects at t1 most likely reflect a floor effect in terms of sentence and text reading (see Table 1), the influence of the three vocabulary variables at t2 was comparable. Interestingly, in sentence reading there was a decreasing impact of vocabulary breadth (receptive) and semantic word knowledge at t3, whereas the impact of vocabulary breadth (expressive) remained stable. However, both receptive vocabulary and expressive vocabulary breadth significantly predicted the children's sentence reading at the following time points.

A similar picture can be found for text reading. While receptive and expressive vocabulary breadth did not show any association with text reading in first grade, their influence grew stronger over time, especially for expressive vocabulary. Receptive vocabulary breadth at t2 showed a similar influence as at t3. Meanwhile, an opposite development was observed for semantic word knowledge. It was clearly associated with text reading at t1, but its influence decreased over time to become insignificant in third grade. The second hypothesis, that the influence of the distinct vocabulary variables on the reading variables is differential, is thus supported by these data.

Semantic word knowledge in form of definitions did not exhibit the same pattern as receptive and expressive vocabulary breadth. A closer look at the correlations reveals that while the normal comprehenders group indeed showed a decreasing influence of the semantic word knowledge on sentence and text reading at t3, the good comprehenders presented an increase of the correlations between t2 and t3. Moreover, the group differences in semantic word knowledge considerably increased between t1 and t3 (see Table 1). Taken together, these results show that the demanding task of defining words is more sensitive for good readers than for readers with an average or slightly below average comprehension. This is in line with the insight that definitional abilities in children are part of metalinguistic skills (Benelli, Belacchi, Gini & Lucangeli, 2006) and is supported by the findings from studies comparing good and poor comprehenders, which have demonstrated that poor comprehenders score in the lowest range in word definition tasks (Colenbrander et al., 2016; Nation et al., 2004, 2007). Nevertheless, hypothesis (3), stating that the impact of the vocabulary measures is higher in more complex reading tasks such as sentence and text reading, can thus be confirmed with a small restriction. It should however be noted that working memory was not taken into account in the present study.

5. Conclusions

In conclusion, it has to be acknowledged that the development of reading comprehension in relation to vocabulary from Grade 1 to Grade 3 is subject to a Matthew effect not only in poor – such as the findings of Cain and Oakhill (2011) show – but also in average comprehenders compared to above average comprehenders. Thus, compared with the study by Cain et al. (2011), the Matthew effect was evident in the present study not only in vocabulary development but also in reading comprehension. The founders for this lie on the one hand in the different composition of the sample and in particular the sample size, which was almost ten times larger in the present study. This allows mapping also more subtle changes in reading comprehension. The largely superior gains in reading comprehension of the above average comprehenders leave no doubt that they have a growing advantage over average readers. It is therefore not only important for poor comprehenders to foster their reading skills and to increase the amount of reading, but also for average readers who may sometimes be neglected. Vocabulary, on the other hand, will not only be boosted by a higher reading intensity but should also have a place in the instruction of average readers. In addition, reading comprehension and vocabulary skills should be promoted in a coordinated manner among poor comprehenders.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the Swiss National Science Foundation for funding the present study (100019_173245). Further, we thank Dr. Ladina Brugger for the valuable input and the joint work in the project, as well as the anonymous reviewers. Finally, we want to thank the participating teachers, the school administrators and especially the children.

References

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. *Comprehension and Teaching: Research reviews*, 77–117.

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14*(5), 255–259. https://doi.org/10.1111/j.0963-7214.2005.00376

Benelli, B., Belacchi, C., Gini, G., & Lucangeli, D. (2006). To define means to say what you know about things": The development of definitional skills as metalinguistic acquisition. *Journal of Child Language, 33*(1), 71–97. https://doi.org/10.1017/S0305000905007312

Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities, 43*(6), 541–552. https://doi.org/10.1177/0022219410371676

Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory, 14*(5), 553–569. https://doi.org/10.1080/09658210600624481

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology, 76*(4), 683–696. https://doi.org/10.1348/000709905X67610

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities, 44*(5), 431–443. https://doi.org/10.1177/0022219411410042

Cain, K., Oakhill, J., & Bryant, P. (2000). Phonological skills and comprehension failure: A test of the phonological processing deficit hypothesis. *Reading and Writing, 13*(1), 31–56. https://doi.org/10.1023/A:1008051414854

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*(4), 671–681. https://doi.org/10.1037/0022-0663.96.4.671

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29*(6), 850–859. https://doi.org/10.3758/BF03196414

Cain, K., & Oakhill, J. (2007). *Children's comprehension problems in oral and written language: A cognitive perspective*. Guilford Press.

Castles, A., Coltheart, M., Wilson, K., Valpied, J., & Wedgwood, J. (2009). The genesis of reading ability: What helps children learn letter–sound correspondences? *Journal of Experimental Child Psychology, 104*(1), 68–88. https://doi.org/10.1016/j.jecp.2008.12.003

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Colenbrander, D., Kohnen, S., Smith-Lock, K., & Nickels, L. (2016). Individual differences in the vocabulary skills of children with poor reading comprehension. *Learning and Individual Differences, 50*, 210–220. https://doi.org/10.1016/j.lindif.2016.07.021

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384–398. https://doi.org/10.1016/j.neuron.2007.10.004

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M.& J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135–154). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch8.

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

Gentaz, E., Sprenger-Charolles, L., & Theurel, A. (2015). Differences in the predictors of reading comprehension in first graders from low socio-economic status families with either good or poor decoding skills. *PloS one, 10*(3), Article e0119581. https://doi.org/10.1371/journal.pone.0119581

Georgiou, G. K., Parrila, R., Cui, Y., & Papadopoulos, T. C. (2013). Why is rapid automatized naming related to reading? *Journal of Experimental Child Psychology, 115*(1), 218–225. https://doi.org/10.1016/j.jecp.2012.10.015

Glück, C. (2011). *Wortschatz- und Wortfindungstest für6- bis 10-Jährige (WWT 6-10)*. Elsevier.

Hertzog, C., Oertzen, T.von, Ghisletta, P., & Lindenberger, U. (2008). Evaluating the power of latent growth curve models to detect individual differences in change. *Structural Equation Modeling: A Multidisciplinary Journal, 15*(4), 541–563. https://doi.org/10.1080/10705510802338983

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., et al. (2017). *IGLU 2016: Lesekompetenzen von grundschulkindern in deutschland im internationalen vergleich*. Waxmann.

Jong, P. F.de, & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6*(1), 51–77. https://doi.org/10.1207/S1532799XSSR0601_03

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*(3), 281–300. https://doi.org/10.1080/10888430802132279

Lenhard, A., Segerer, R., Lenhard, W., & Suggate, S. P. (2015). *Peabody picture vocabulary test–revision 4 (PPVT-4)*. Pearson. German version.

Lenhard, W., Lenhard, A., & Schneider, W. (2017). *ELFE II: A reading comprehension test for Grades 1 to 7*. Hogrefe.

Moll, K., & Landerl, K. (2014). *Lese-und rechtschreibtest (SLRT-II). weiterentwicklung des salzburger lese-und rechtschreibtests (SLRT), 2., korrigierte auflage mit erweiterten Normen*. Verlag Hans Huber.

Nation, K. (2005). Children's reading comprehension difficulties. In M. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248–265). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch14.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children. *Journal of Speech, Language, and Hearing Research, 47*(1), 199–211. https://doi.org/10.1044/1092-4388(2004/017

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*(1), 85–101. https://doi.org/10.1006/jmla.1998.2564

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*(1), B1–B13. https://doi.org/10.1016/S0010-0277(99)00004-9

Nation, K., Snowling, M. J., & Clarke, P. (2007). Dissecting the relationship between language skills and learning to read: Semantic and phonological contributions to new vocabulary learning in children with poor reading comprehension. *Advances in Speech Language Pathology, 9*(2), 131–139. https://doi.org/10.1080/14417040601145166

Foulin, N. J. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing, 18*(2), 129–155. https://doi.org/10.1007/s11145-004-5892-2

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554–566. https://doi.org/10.1037/0022-0663.98.3.554

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 67–86). John Benjamins Publishing, 11 http://digital.casalini.it/9789027296108-Casalini-id:5015797.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M.& J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch13.

Petermann, F., & Petermann, U. (2007). *Hamburg-Wechsler-Intelligenztest für kinder iv (HAWIK-IV)*. Huber.

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86*(1), 159–175. https://doi.org/10.1111/cdev.12292

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 209–227). John Benjamins.

Richter, T., Isberner, M.-B., Naumann, J., & Neeb, Y. (2013). Lexical quality and reading comprehension in primary school children. *Scientific studies of reading, 17*(6), 415–434. https://doi.org/10.1080/10888438.2013.764879

Ricketts, J., Bishop, D. V. M., & Nation, K. (2008). Investigating orthographic and semantic aspects of word learning in poor comprehenders. *Journal of Research in Reading, 31*(1), 117–135. https://doi.org/10.1111/j.1467-9817.2007.00365.x

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson, & S. B. Neuman (Eds.), *Handbook of early literacy research* (pp. 173–182). Guilford, 2.

Spencer, M., & Wagner, R. K. (2018). The comprehension problems of children with poor reading comprehension despite adequate decoding: A meta-analysis. *Review of Educational Research, 88*(3), 366–400. https://doi.org/10.3102/0034654317749187

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407. Retrieved from http://www.jstor.org/stable/747612.

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381–398. https://doi.org/10.1207/s1532799xssr1004_3

Van Bergen, E., Snowling, M. J., Zeeuw, E. L.de, van Beijsterveldt, C. E. M., Dolan, C. V., & Boomsma, D. I. (2018). Why do children read more? The influence of reading ability on voluntary reading practices. *Journal of Child Psychology and Psychiatry, 59*(11), 1205–1214. https://doi.org/10.1111/jcpp.12910

Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System, 41*(3), 609–624. https://doi.org/10.1016/j.system.2013.07.012

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading, 15*(1), 8–25. https://doi.org/10.1080/10888438.2011.536125

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to acquisition and frequency of input. *Applied Psycholinguistics, 22*, 217–234. https://doi.org/10.1017/S0142716401002041

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review, 53*(1), 13–40. https://doi.org/10.3138/cmlr.53.1.13