ProVoc : an app to train vocabulary depth in order to foster children's reading comprehension

Anna Potocki, Mathilde Chailleux, Manuel Gimenes, Jean Pylouster

To cite this version:

Anna Potocki, Mathilde Chailleux, Manuel Gimenes, Jean Pylouster. ProVoc : an app to train vocabulary depth in order to foster children's reading comprehension. Journal of Computer Assisted Learning, 2021, 37 (5), pp.1324-1335. 10.1111/jcal.12572 . hal-03357789

ProVoc: An app to train vocabulary depth in order to foster children's reading comprehension

Anna Potocki[1,2], Mathilde Chailleux[1], Manuel Gimenes[1], & Jean Pylouster[1]

[1]Centre de Recherches sur la Cognition et l'Apprentissage, UMR 7295, CNRS, Université de Poitiers, Université François Rabelais of Tours, MSHS - Bâtiment A5, 5, rue T. Lefebvre, TSA 21103, 86073 POITIERS Cedex 9, France

[2]Laboratoire de Recherches sur les Apprentissages en Contexte, EA 602, Université Grenoble Alpes, bâtiment BSHM, 1251 Avenue Centrale, 38400 Saint-Martin-d'Hères, France

Author Note

Correspondence concerning this article should be addressed to Anna Potocki, LaRAC, Université Grenoble Alpes, Bâtiment BSHM, Aile C, 1251 Avenue Centrale, 38400 Saint-Martin-d'Hères, France; e-mail: anna.potocki@univ-grenoble-alpes.fr.

Acknowledgments

Support for this research was provided by a grant from La Région Nouvelle Aquitaine (CPER-FEDER P-2017-BAFE-68), in partnership with the European Union (FEDER/ERDF, European Regional Development Fund). The authors would like to thank Justine Robin and Magali Lamboley for their help in implementing the training sessions and data collection within the schools; as well as Yann Novelli for his help in data formatting.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

ProVoc: An app to train vocabulary depth in order to foster children's reading comprehension

Abstract

Background: Many studies have demonstrated the crucial role of vocabulary in predicting reading performance in general. More recent work has indicated that one particular facet of vocabulary (its depth) is more closely related to language comprehension, especially inferential comprehension. On this basis, we developed a training application to specifically improve vocabulary depth.

Objectives: The objective of this study was to test the effectiveness of a mobile application designed to improve vocabulary depth.

Methods: The effectiveness of this training was examined on 3rd and 4th grade children's vocabulary (breadth and depth), decoding and comprehension performances. A *randomized waiting-list control* paradigm was used in which an experimental group first received the intervention during the first 4 weeks (between pretest and post-test1), thereafter, a waiting control group received the training for the next 4 weeks (between postest1 and posttest2).

Results: Results showed that the developed application led to significant improvements in terms of vocabulary depth performance, as well as a significant transfer effect to reading comprehension. However, we did not observe such a beneficial effect on either vocabulary breadth or written word identification.

Implications: These results are discussed in terms of the links between vocabulary depth and comprehension, and the opportunities the app presents for remedying language comprehension deficits in children.

Keywords: Comprehension, Vocabulary, Mobile learning

Introduction

Links between vocabulary and reading comprehension

From the very outset, research on reading literacy has highlighted a strong relationship between reading skills and vocabulary (e.g., National Institute of Child Health and Human Development, 2000). According to Gough and Tunmer (1986)'s framework, reading is made up of three components: word identification, language comprehension, and reading comprehension, which is the result of the interaction between the first two processes. Vocabulary has been found to be a powerful predictor of these different components, playing an important role in both written word identification and comprehension. First, as regards word identification, numerous developmental studies have reported correlations between vocabulary and word-reading performances (see Scarborough, 2001; Sénéchal, Ouellette, & Rodney, 2006). From the perspective of *dual-route* models of word identification (Coltheart, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Perry, Ziegler, & Zorzi, 2010), a larger lexicon may facilitate the immediate identification of written words via the *lexical route.* However, it has also been suggested that vocabulary growth is also associated with word decoding (nonlexical procedure), through its role in the development of phonological awareness (Goswami, 2001; Metsala, 1999; Walley, Metsala, & Garlock, 2003). According to this view, as more word forms are added to the lexicon, children become increasingly sensitive to sublexical details, leading to greater phonemic awareness. Second, lexical knowledge and language (either oral or written) comprehension have been shown to be strongly correlated (e.g., Nagy, Berninger & Abbot, 2006; Sénéchal et al., 2006) in adults (e.g., Carroll, 1993), adolescents (e.g., Cromley & Azevedo, 2007), school-aged children (e.g., Cain, Oakhill, & Lemmon, 2004; Lonigan, Burgess, & Anthony, 2000, Oakhill & Cain, 2012; Oakhill, Cain, & Bryant, 2003; Ricketts, Nation, & Bishop, 2007), and preschool readers (Florit, Roch, Altoè, & Levorato, 2009, Roth, Speece, & Cooper, 2002). For example, Vellutino, Tunmer, Jaccard, and Chen (2007) found a .64 correlation between listening comprehension and vocabulary in school-aged children. In preschoolers, Florit et al. (2009) showed that vocabulary explains 22% of the variance in performances on listening text comprehension. For obvious reasons, readers cannot understand a text without knowing what most of its constituent words mean (e.g., Fayol, Gombert, Lecocq, Sprenger-Charolles, & Zagar, 1992; Oakhill et al., 2003).

Research conducted on struggling readers also provides interesting insights into the role of vocabulary in decoding and comprehension. In general, three types of poor readers can be identified, based on Gough and Tunmer's (1986) *simple view of reading* framework: those who have difficulties with identifying written words but not with language comprehension (so-called *poor decoders* or *dyslexics*); those with persistent difficulties with comprehension but average word-reading performances (*poor comprehenders*); and those who exhibit difficulties with both types of processes (*garden-variety poor readers*; see, for example, Catts, Hogan, & Fey, 2003; Potocki, Magnan, & Ecalle, 2015). Dyslexic individuals are supposed to have relatively minor lexical representation deficits (Bishop & Snowling, 2004; Snowling, 2000), and their potentially greater semantic knowledge may compensate for their word-reading difficulties (Cavalli et al., 2016; Miller-Shaul, 2005; van Viersen, de Bree, & de Jong, 2019). Poor comprehenders, however, tend to experience difficulties in tasks requiring access to word meanings. For instance, compared with age-matched control children, they have been described as being slower and less precise in naming infrequent word pictures (Nation, Marshall, & Snowling, 2001), exhibiting reduced semantic priming effects for categorically related but poorly associated words (e.g., *bed-office*; Nation & Snowling, 1999), and having difficulty making synonym judgments (e.g., *fast-quick*; Nation & Snowling, 1998). In addition, the *lexical quality hypothesis* (Perfetti, 2007; Perfetti & Hart, 2002, see also Perfetti & Stafura, 2014) assumes that word knowledge (both form and meaning) is central to reading skills, and suggests that poor comprehenders have poorer lexical representations that may, in turn, hinder their higher-level comprehension processes (Perfetti & Stafura, 2014; Perfetti, Yang, & Schmalhofer, 2008; Yang, Perfetti, & Schmalhofer, 2005). It should be noted that poor access to word meanings is one of several possible causes of poor comprehension, and there is no clear consensus as to the importance of these word-level difficulties in explaining comprehension difficulties (Nation, 2005).

The role of vocabulary as a predictor of reading skills has therefore been well documented. However, a number of studies have also suggested the existence of a two-way link between vocabulary and comprehension, meaning that comprehension level may be predictive of vocabulary growth (e.g., Eldredge, Quinn, & Butterfield, 1990). For instance, Cain and Oakhill (2011) highlighted a *Matthew effect* (Stanovich, 1986) in the development of vocabulary in good versus poor comprehenders. In a longitudinal follow-up of children between the ages of 8 and 16 years, these authors showed that poor comprehenders identified at 8 years had 1) a lower starting level of vocabulary, but also 2) less vocabulary growth, compared with good comprehenders, during this 8-year period. The gap between good and poor comprehenders in terms of vocabulary therefore widened with age. One possible explanation for this group difference is that poor comprehenders are less motivated to read during their free time because they have greater difficulty understanding what they read. However, there is no clear evidence to date that children with poor reading comprehension do indeed have less print exposure (Cain, Oakhill, & Bryant, 2000; Ricketts et al., 2007). Another explanation could be that children with a lower level of comprehension find it more difficult to infer the meaning of the new words they encounter during reading, thus preventing them from increasing their lexicon while reading (Cain et al., 2004). Poor comprehenders may indeed lack the strategic knowledge needed to make use of context to infer word meanings, and seem to construct less stable memory traces for newly encountered words (Nation &

Snowling, 1999; Perfetti, Landi, & Oakhill, 2005). In sum, reading is generally described as the best way of acquiring new vocabulary (Hargrave & Sénéchal, 2000; Stanovich, 1993), but this statement may only be true for children with good comprehension skills, who can use context to deduce the meaning of the unknown words they read (Cain et al., 2004).

The above-mentioned studies examining the role of vocabulary in reading (and vice versa) generally assessed vocabulary using *quantitative* measures, simply counting the number of words known to readers. However, recent research has highlighted the need to distinguish between two facets of vocabulary when examining this relationship between vocabulary and reading skills.

Distinction between two aspects of vocabulary

Less than 20 years ago, researchers started to make a distinction between two facets of vocabulary, namely *breadth* and *depth* (e.g., Cain & Oakhill, 2014; Ouellette, 2006; Tannenbaum, Torgesen, & Wagner, 2006). Before that, knowing a word was more or less always judged according to whether individuals knew the form-meaning connection. The proposed distinction corresponds to the difference between quantitative and qualitative aspects of vocabulary, that is, the difference between how many words are known (i.e., breadth) and how well the meanings of these words are known (i.e., depth). Whereas vocabulary breadth has a clear definition, no real consensus has been reached for vocabulary depth, and there is a plethora of definitions (Yanagisawa & Webb, 2019). This concept relating to the quality of word knowledge can therefore refer to semantic, syntactic or morphological knowledge, as well as links to and/or differences from other words, and *contextualized* knowledge about words, that is, knowledge about variations in word meanings according to context (Binder, Cote, Lee, Bessette, & Vu, 2017; Proctor, Silverman, Harring, & Montecillo, 2012; Read, 2004).

A number of different tasks have been used to assess these two facets of vocabulary. There are already several standardized tests for assessing vocabulary breadth, including the famous Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), in which participants have to match images to words. Although no tests are explicitly referenced as assessing vocabulary depth, one task that is commonly used for this purpose is a word definition task, which assesses whether individuals know the word, but can also probe more qualitative aspects related to the richness of their responses.

Breadth and depth are thus distinct facets of the vocabulary construct, and recognizing this distinction has led to a better understanding of vocabulary's role in reading acquisition. Research has demonstrated that each one can predict specific aspects of reading skills, as defined in the simple view of reading framework.

Using a 3-year-longitudinal paradigm with typically developing children, a Roth et al. (2002) examined the links between different linguistic abilities (including vocabulary and phonological awareness) measured in kindergarten, and early reading skills in first and second grade. Their main finding was that out of all the measures they performed in kindergarten, only vocabulary (assessed by both oral definitions and word retrieval) was predictive of early reading comprehension, accounting for 23% of the variance after controlling for print awareness skills. By contrast, single-word reading was mainly predicted by phonological awareness and by word retrieval (more related to vocabulary breadth), and only in Grade 2. With a sample of older children in Grade 4, Ouellette (2006) confirmed and extended this first result by assessing different aspects of reading (decoding, visual word recognition, and reading comprehension), as well as different aspects of vocabulary: receptive and expressive single-word vocabulary to account for vocabulary breadth; and a semantic information task to account for vocabulary depth. The author found that reading comprehension was related to both vocabulary breadth and depth, whereas only receptive vocabulary (i.e., breadth) predicted significant variance in word reading. He therefore concluded that written text comprehension places more demands on vocabulary knowledge than word reading, and children who have more comprehensive word knowledge (i.e., vocabulary depth) should have better language comprehension. Binder et al. (2017) reached the same conclusion with college-aged students.

In sum, it seems that comprehension is more closely related to vocabulary depth than to vocabulary breadth, although Tannenbaum et al. (2006) drew the opposite conclusion in their study, conducted in a large sample of third graders. Nevertheless, it seems important to specify here that the same task (oral definition of words) was interpreted by some (e.g., Ouellette, 2006) as a measure of vocabulary depth, and by others (Tannenbaum et al. 2006) as a measure of vocabulary breadth. Like Ouellette (2006; but see also Binder et al., 2017; Cain & Oakhill, 2014), we believe that definition tasks reflect vocabulary depth rather than breadth, for beyond whether the person knows the word or not, the scoring usually also takes account of the accuracy and detail with which the word is defined.

With a view to further examining the links between vocabulary and comprehension, Cain and Oakhill (2014) recently proposed distinguishing between breadth and depth when examining the impact of vocabulary on two aspects of comprehension: literal and inferential. They even went further, by distinguishing between two types of inferences: local and global. First, they observed that vocabulary was more predictive of inferential than literal comprehension skills. More specifically, they reported that vocabulary was more explanatory of global inferences than of local ones. Finally, they observed that vocabulary depth rather than breadth was more critical in explaining comprehension skills, especially global inferences, even after literal comprehension and word-reading skills were taken into account in the analyses. These two authors therefore concluded that some "aspects of vocabulary knowledge may be more important for higher level comprehension skills than others" (p. 2). This conclusion could have a major impact on vocabulary training.

Vocabulary training: effects on reading comprehension

The above literature review highlighted strong relationships between vocabulary and reading comprehension skills. In addition, most of the cited studies used statistical techniques to examine the nature of these relationships, and concluded that there is a causal link between vocabulary and comprehension. On this basis, we can hypothesize that increasing vocabulary has a beneficial effect on the level of comprehension. As mentioned earlier, one way of increasing vocabulary may simply be to read texts (e.g., Nagy, 2010; Sénéchal, 2000). However, this is only true for good readers/good comprehenders (e.g., Cain et al., 2004; Oakhill & Cain, 2000) who are able to infer unknown word meanings when reading them in context. Therefore, a better strategy for fostering reading comprehension in poorer readers/comprehenders would be to specifically train vocabulary. However, reviews of such training studies have actually highlighted relatively mixed results (e.g., Elleman, Lindo, Morphy, & Compton, 2009; Wright & Cervetti, 2017). For example, Mezynski (1983) analyzed eight studies that had investigated the effect of vocabulary training on comprehension. While these studies generally showed improvements in students' word knowledge, few-if any-reported a transfer to comprehension. Two other meta-analyses (Elleman et al., 2009; Stahl & Fairbanks, 1986) concluded that vocabulary interventions generally had positive effects on comprehension outcomes, but only for the comprehension of passages containing the words taught during the interventions (Stahl & Fairbanks, 1986), and minimal effects were observed in standardized measures of comprehension (Elleman et al., 2009).

In their recent review of the literature, Wright and Cervetti (2017) systematically examined vocabulary interventions that could positively impact comprehension. They identified 36 studies conducted in children from pre-kindergarten to secondary school, and published between 1965 and 2015. They distinguished between them according to two main criteria: 1) type of intervention (i.e., direct teaching of word meanings or teaching of strategies for making sense of unknown words during reading); and 2) type of comprehension measures used (i.e., comprehension of passages containing the taught words, or more generalized/standardized comprehension measures). In line with Elleman et al. (2009), their findings showed that teaching word meanings supports the comprehension of texts containing the taught words, but such interventions – even long-term ones – barely improve comprehension assessed with general texts. The fact that there was scarcely any transfer to the comprehension of texts that did not contain the words used for the vocabulary training led the authors to question the relevance of using classroom/teaching time for vocabulary training, at least if the aim is to improve comprehension.

One potential reason why previous vocabulary interventions may have been unproductive is that most of them may not have succeeded in fostering vocabulary depth. First, numerous intervention studies still focus on helping students gain knowledge about form-meaning connections and enlarge their lexicon (Webb, 2019). And yet, as seen earlier, lexicon size (i.e., vocabulary breadth) is virtually unrelated to text comprehension, or at least is less related to higher-order comprehension processes such as inferencing than depth of vocabulary is. In addition, as pointed out by Wright and Cervetti (2017) in the conclusion of their meta-analysis, training based on active processing of word meanings, for example through work on semantic features, seems particularly promising. In our study, we therefore developed a new training program in which students had to actively work on words, by linking words and their semantic features or linking words together from a semantic point of view. Our aim was also to examine the effects of this training on the comprehension of texts that did not contain the trained words and in reading comprehension tasks tapping more into inferential processes. We reasoned that because research has demonstrated that vocabulary depth is more predictive than vocabulary breadth of comprehension (Ouellette, 2006), and more precisely of inferential comprehension (Cain & Oakhill, 2014), our study would highlight the appropriate conditions for enhancing the reading comprehension skills of children in general.

To explore this hypothesis, we developed a new French application called ProVoc that was specifically aimed at training vocabulary depth, and tested its effectiveness on children's vocabulary and comprehension performances. No such app was currently available for French-speaking students, who only had access to a few French vocabulary tools intended to promote second-language learning or extend lexicon breadth (quantity of words known) rather than depth. Finally, to our knowledge, no research had been conducted to validate these tools for enhancing students' reading comprehension performance within a controlled paradigm. The present study was carried out to fill in this gap, as well as to make it possible to explore theoretical hypotheses concerning the links between vocabulary and comprehension.

Method

Participants

The initial sample was composed of 173 French-speaking children in Grades 3 (n = 79) and 4 (n = 92) from two schools in a small town in western France. These schools were located in so-called *priority education zones*, that is, areas known to contain families with lower socio-economic status. Participants' mean age was 9.1 years. They were randomly assigned to either an experimental or a control group. The initial experimental group was composed of 82 children (mean age = 9.2 years), and the initial control group of 91 children (mean age = 9.1 years). The groups' roles were then switched, so that each child was exposed to the training, either in the first period, or in the second period (see description of the procedure below). The statistical analyses that we then carried out on the data only concerned a sample of 123 students: 73 in the initial experimental group (19 in Grade 3; 54 in Grade 4) and 50 in the initial control group (17 in Grade 3; 33 in Grade 4). Although all the children could take part in the study (with their assent and the consent of their parents), the data of students who had a proven developmental disorder (who were present in the classes in inclusion time), had only recently arrived in France (i.e., nonFrench speaking children), had missing data for the pretest or one of the posttests, or had attended too few training sessions (fewer than five training sessions out of eight) were removed from the analyses.

Material

Pre- and posttest measures

Vocabulary breadth. Children's breadth of vocabulary was assessed with a French-language adaptation of the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn, Theriault-Whalen, & Dunn, 1993). In each trial, the children had to correctly identify a named picture from four possible black-and-white drawings. The test featured 20 words, half taken from the original Peabody test (which did not appear in the app), and half taken from the app. This allowed us to distinguish between trained and nontrained words in the subsequent analyses. For the nontrained words from the PPVT-R, we chose 10 items for the relevant age group (8-10 years), selecting those that were least likely to be known by the students in the pretest. These were the words *uniforme*, *edifice*, *demeure*, *cubique*, *taquin*, *nutritive, breuvage*, *cultivateur*, *rive*, *oratoire* [uniform, edifice, dwelling, cubic, teaser, nutritious, beverage, cultivator, shore, oratory]. For the trained words from the application, we again selected 10 words from among all the target words presented in the exercises, choosing rare words to avoid a ceiling effect from the pretest. These were the words *macaron*, *chicorée*, *cassoulet*, *pédiatre*, *jaguar*, *théière*, *braille*, *cachalot*, *cordonnier*, *châtaigne* [macaroon, chicory, cassoulet, pediatrician, jaguar, teapot, braille, sperm whale, shoemaker, chestnut].

The total number of correct answers constituted the total raw score (max = 20), and we also calculated two subscores for the trained versus nontrained words (max = 10 for each).

Vocabulary depth. Vocabulary depth was assessed with a French-language version of the Vocabulary subtest of the Wechsler Intelligence Scale for Children, Four Edition (WISC-IV; Wechsler, 2003). We adapted the procedure so that the children had to write down their definitions of the words, instead of orally defining them. The test featured 20 words: 10 taken from the WISC-IV (which were not presented in the app), and 10 taken from the app. This allowed us to distinguish between trained and nontrained words in the subsequent analyses. The nontrained words from the WISC-IV were *parapluie*, *horloge*, *gant*, *silence*, *bicyclette*, *cambrioleur*, *hiberner*, *débrouillard*, *rivalité*, *fasciné* [umbrella, clock, glove, silence, bicycle, burglar, hibernate, resourceful, rivalry, fascinated]. The trained words from the application were the same as those used for the vocabulary breadth task: *macaron*, *chicorée*, *cassoulet*, *pédiatre*, *jaguar*, *théière*, *braille*, *cachalot*, *cordonnier*, *châtaigne* [macaroon, chicory, cassoulet, pediatrician, jaguar, teapot, braille, sperm whale, shoemaker, chestnut].

We applied the WISC-IV scoring criterion to the words from this test, meaning that each word definition could be scored 0, 1 or 2 points, depending on its accuracy. An equivalent scoring system was developed for the words taken from the app. Two of the experimenters (first two co-authors of this paper) first applied these scoring guidelines to the responses of the first 30 students, then modified the guidelines according to any disagreements or ambiguous responses. Next, these new scoring guidelines were applied to a new set of 30 responses. The scoring guidelines we developed allowed us to achieve interrater agreement > 90%. The dependent variables were the total number of correct definitions (max = 40), and two subscores for the trained versus nontrained words (max = 20 for each).

Reading comprehension. To assess children's reading comprehension skills, we used a test taken from Potocki, Bouchafa, Magnan, and Ecalle (2014). This test featured a 232-word narrative followed by 12 open-ended questions. After reading the text to themselves, the children answered the questions. They could not refer back to the text. The questions elicited both literal and inferential processes. Two types of inferences (Cain & Oakhill, 1999; Potocki, Ecalle, & Magnan, 2013) were targeted: text-connected inferences (i.e., causal or referential relationships derived from successive textual statements); and knowledge-based inferences (i.e., relationships that require general knowledge about the world to construct a coherent representation of the text as a whole). Children's answers were scored 0, 0.5 or 1 point, depending on their accuracy. As for vocabulary depth, two of the experimenters (first two co-authors of this paper) first applied these scoring guidelines to the responses of the first 30 students, and then modified the guidelines according to any disagreements or ambiguous responses. Next, these new scoring guidelines were applied to a new set of 30 responses. The scoring guidelines we developed allowed us to achieve interrater agreement > 90%. The maximum score was 12.

Word identification. A standardized test of word reading (Timé3; Ecalle, 2006) was used to assess the children's written word identification skills. This test comprised two forced-choice tasks. In the first task, children had to correctly identify written target words corresponding to pictures. In the second task, they had to correctly identify written target words corresponding to semantically associated words. Each task featured 20 trials, in each of which the children had to circle the correctly written word among four distractors. The dependent variable was the total number of correctly identified words (max = 40).

App to train vocabulary depth.

ProVoc, the app designed to train children's depth of vocabulary, featured three types of exercises. Chosen to correspond to the definitions of vocabulary depth given in the literature (e.g., Binder et al., 2017; Proctor et al., 2012; Read, 2004), two exercises ("find the correct attributes" and "find the correct word") targeted the words' semantic features, and one exercise ("find the synonyms") targeted the semantic links between the words. Based on active processing of the words' semantic features, the proposed exercises also met Wright and Cervetti (2017)'s recommendations based on their meta-analysis.

Each exercise was composed of 20 series of 20 words each (i.e., 400 words per exercise). The series were made increasingly difficult by including higher numbers of rare words. Lexical frequency was controlled using the Manulex database (Lété, Sprenger-Charolles, & Colé, 2004). The first series therefore contained a majority of frequent words, and this proportion then decreased across the series in favor of rare words.

Find the correct attributes. In each trial, children saw a target word at the top of the screen and had to select semantically related attributes from a list of 12 words (see Figure 1). In each trial, this list contained a mixture of distractors and 2-10 correct attributes. For each word, we had previously defined 2-5 *essential* attributes. A pilot validation of these attributes was carried out among 15 university students, who were asked to give the first five semantic features they thought were associated with each target word. For their answer to be deemed correct, children had to select either all the *essential attributes* or at least 70% of all the semantically related words.

After giving an initial response, the children received corrective feedback with information on the correct/incorrect nature of their selection (i.e., “Great!” / “Try again!”). They then had the opportunity to modify their response, either by clicking on new attributes or by removing previously validated ones. After this, they received further corrective feedback on the correct/incorrect nature of their answer (i.e., “Great!” / “Pity!”), followed by the expected correct answer where all the correct attributes were highlighted by a colored frame.

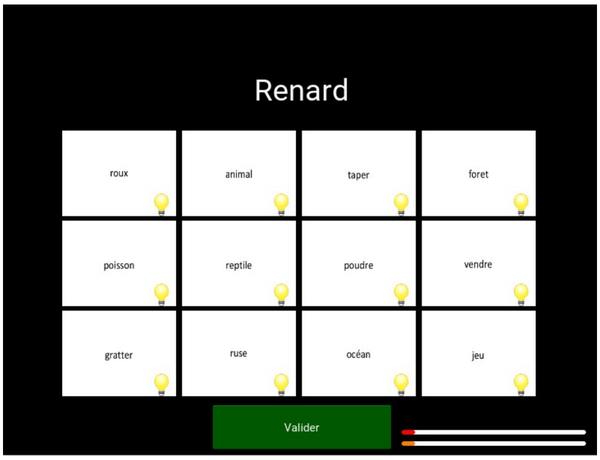

Figure 1. Screenshot from the "find the correct attributes" exercise.

Note. The target word is "fox". The proposed attributes are "ginger, animal, to hit, forest, fish, reptile, powder, sell, to scratch, smart, ocean, game".

Find the correct word. In each trial, the children were given a series of 3-6 semantic attributes and asked to identify the word that these attributes correctly defined among six response options (see Figure 2). For example, when the attributes were *garden*, *raw*, *to eat*, *green*, *plant* and *leaves*, the children had to choose *lettuce* from among the distractors *carrot*, *strawberry*, *meat*, *radish* and *chestnut*. After giving their answer, the children received corrective feedback informing them of the accuracy (or otherwise) of their answer (i.e. "Great!" / "Pity!"), along with the correct answer highlighted by a colored frame.

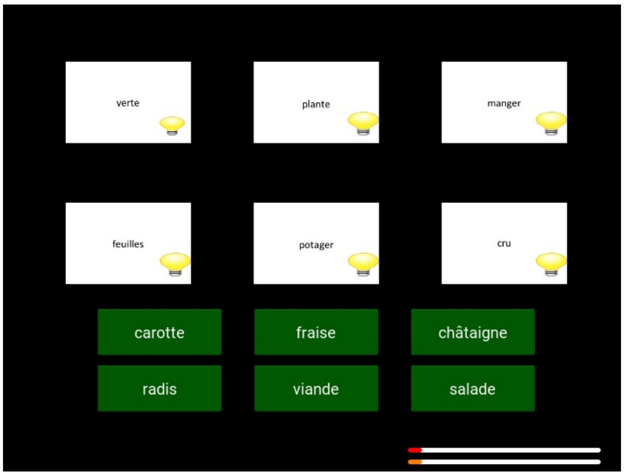

Figure 2. Screenshot from the "find the correct word" exercise.

Note. The given attributes are "green, plant, eat, leaves, vegetable garden, raw". The proposed answers are "carrot, strawberry, chestnut, radish, meat, salad"

Find the synonyms. In each trial, children saw a target word at the top of the screen and had to find its synonym (i.e., the most semantically related word) among six response options (correct answer and five distractors). For example, in Figure 3 below, the target word is *mist*, the synonym is *fog*, and the five distractors are *night*, *wind*, *hail*, *bright* and *thunderstorm*. After making their initial choice, the children received corrective feedback informing them of the (in)correct nature of their answer (i.e., "Great!" / "Try again!"). They could then modify their response by clicking on another word. They again received corrective feedback on this response (i.e., "Great!"/ "Pity!"), this time with the correct answer highlighted by a colored frame.

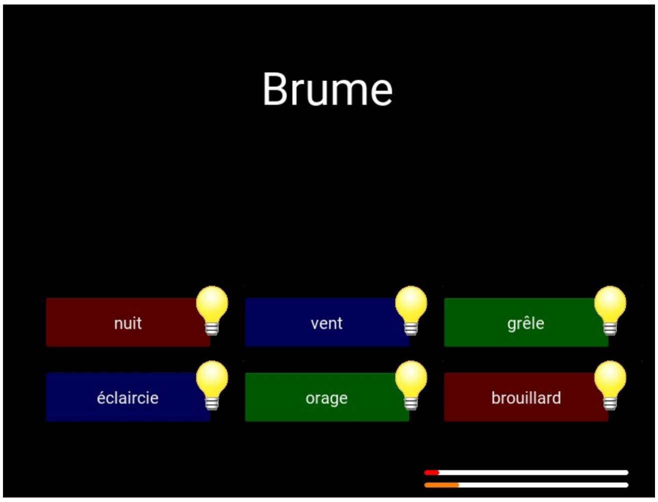

Figure 3. Screenshot from the "find the synonym" exercise.

Note. The target word is "mist". The proposed synonyms are "night, wind, hail, clearing, thunderstorm, fog".

Procedure

We implemented a pretest/training/posttest paradigm with an experimental group and a control group. The training sessions were spread over 4 weeks, with two 30-minute sessions per week. During the training sessions, children were seen in small groups (about 10 students) in a quiet room of the school. They performed the training on Android tablets provided by the research team. At each session, the children worked on one particular exercise and did as many series of this exercise as they could within the 30 minutes of training. The control group followed their usual classes. A *randomized waiting-list control* paradigm was used so that ultimately all the children underwent the training (see Figure 4). This paradigm also allowed us to test the longer-term effects of the training.

The researchers administered the pre- and posttest measures over two sessions in the children's classroom, in presence of their usual teacher. In one session, the students performed the reading comprehension and vocabulary breadth tasks. In the other session, they performed the vocabulary depth and written word identification tasks. The order of these two sessions was counterbalanced across classes. The sessions each lasted approximately 40-45 minutes, and took place on different days.

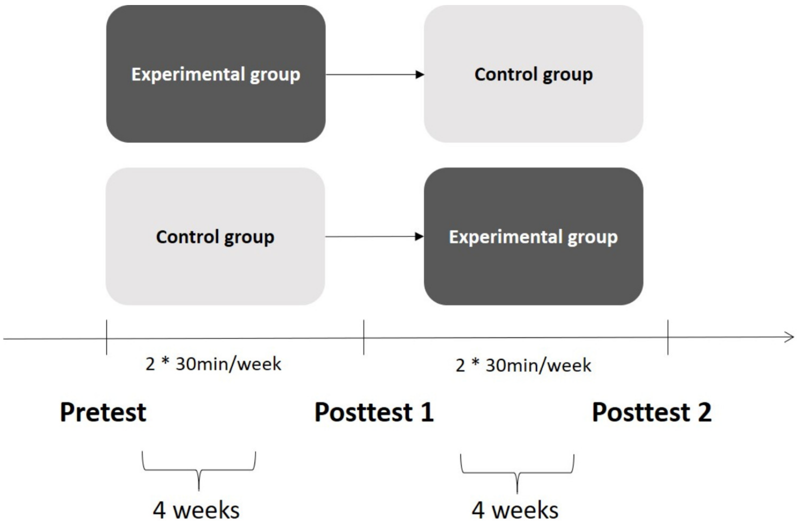

Figure 4. Presentation of the "randomized waiting-list control" procedure used for the implementation of the training program

Results

Descriptive statistics are set out in Table 1.

To determine the effectiveness of the training for each variable of interest (vocabulary depth, reading comprehension, vocabulary breadth, and written word identification), the progress made by each child for each item on each task has been calculated between the different evaluation time points (pretest/posttest1 and posttest1/posttest2). We then ran mixed model analyses with group (experimental-control vs. control-experimental) as a fixed effect[1], and participant and item as random effects. In these analyses, the initial level of children at pretest was also controlled for by entering it as a covariate. The analyses were performed on jamovi (version 1.1.9, retrieved from www.jamovi.org), using the GAMLj module (general, mixed and generalized models module for jamovi; Gallucci, 2019).

Table 1. Mean (*SD*) scores on each task for each group at the pretest and first and second posttests.

Task	Group	Pretest	Posttest 1	Posttest 2
Word identification	1st experimental group	24.1 (7.5)	26.5 (6.9)	27.5 (6.5)
	1st control group	26.6 (5.6)	27.9 (5.4)	27.9 (5.7)
Reading comprehension	1st experimental group	5.67 (3.1)	7.65 (2.8)	7.61 (2.8)
	1st control group	5.44 (2.9)	6.41 (3.2)	7.96 (3)
Vocabulary depth	1st experimental group	14.9 (4.9)	17.7 (6.1)	18.5 (6.4)
	1st control group	15 (5.5)	15.4 (5.8)	17.8 (6.5)
Vocabulary breadth	1st experimental group	11.2 (3.5)	12.7 (3.1)	13.1 (3.5)
	1st control group	10.5 (4)	12.0 (3.4)	12.5 (3.8)

Impact on vocabulary depth.

For the vocabulary depth task, we distinguish between the words that were trained (from the app) and the words that were not trained (from WISC-IV) in the analyses. For the trained words, we observed a significant Group effect ($F(1,118) = 7.9$, $p < .01$) for the first period (pretest/postest1) with the children trained during this period (experimental-control group) showing higher progress than the children in the control-experimental group. For the second

[1] Analyses with grade level as a fixed effect were carried out but as grade did not significantly interact with group and because we did not have any hypothesis as regards different progression as a function of children's grade level, it was removed from the final analyses.

period (posttest1/posttest2), the Group effect was also significant (F(1,121 = 4.7, $p < .05$) with the children in the control-experimental group (trained during this period) exhibiting higher progress in vocabulary breath for the trained words (see Figure 5).

For the nontrained words (i.e., taken from WISC-IV and not presented in the app), the Group effect was not significant at any period (pretest/postest1: $F(1, 120) = 2.3$, $p = .13$; posttest1/posttest2: $F(1, 120) = 3.8$, $p = .053$).

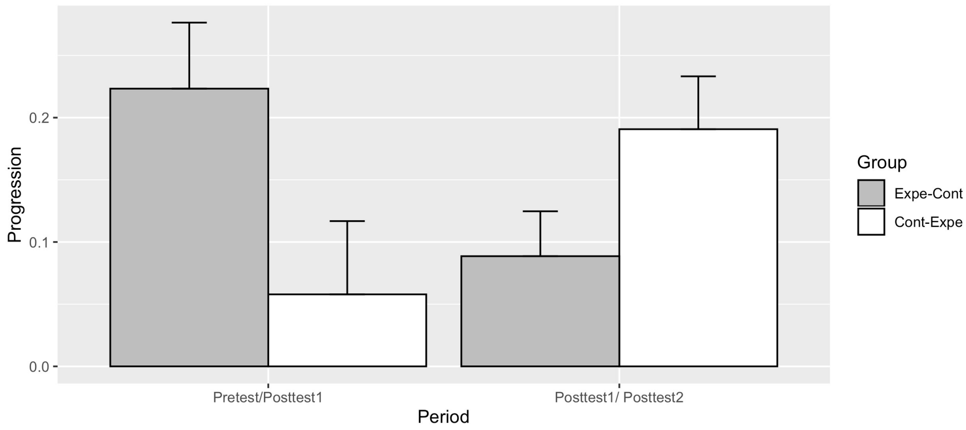

Figure 5. Progress made by each group on the trained words of the vocabulary depth task between the different evaluation sessions.

Note. Progress are expressed at the item level on the figure. Expe-cont= experimental-control group (trained between pretest and postests1). Cont-expe = control-experimental group (trained between posttest1 and posttest2).

Impact on reading comprehension.

For the first period (pretest/posttest1), we observed a significant Group effect (F(1,118 = 6.8, $p = .01$) with the children trained during this period (experimental-control group) having higher progress in reading comprehension than the control-experimental children. For the second period (posttest1/posttest2), the Group effect was also significant (F(1,121 = 9, $p < .01$) with the children in the control-experimental group trained during this period exhibiting higher progress in reading comprehension (see Figure 6).

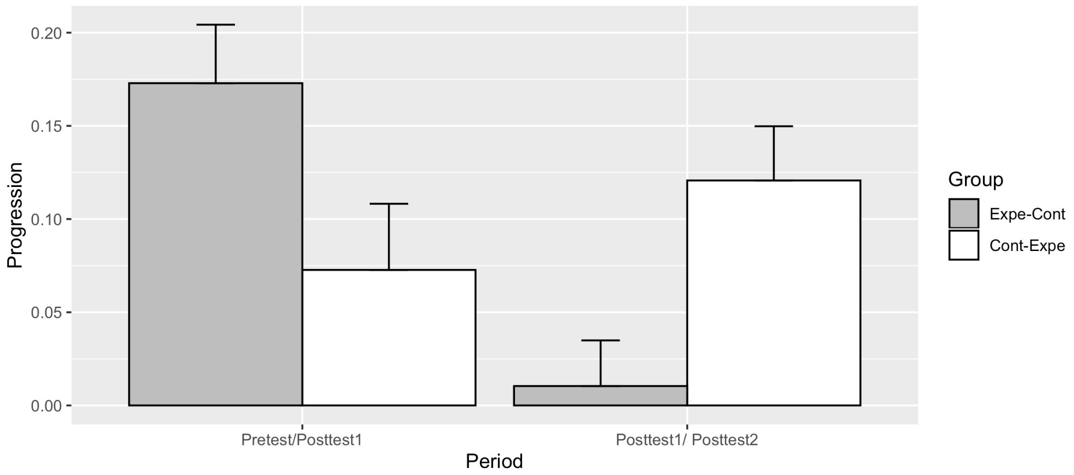

Figure 6. Progress made by each group on the reading comprehension task between the different evaluation sessions.

Note. Progress are expressed at the item level on the figure. Expe-cont= experimental-control group (trained between pretest and postests1). Cont-expe = control-experimental group (trained between posttest1 and posttest2).

Impact on vocabulary breadth.

For the vocabulary breadth task, the Group effect was not significant neither for the trained (pretest/postest1: $F(1, 119) = 2.8$, $p = .09$; posttest1/posttest2: $F(1, 122) = .11$, $p = .74$) nor the untrained words (pretest/postest1: $F(1, 118) = .79$, $p = .38$; posttest1/posttest2: $F(1, 120) = .11$, $p = .75$) at any period.

Impact on written word identification.

For the written word identification task, the Group effect is not significant, at any period (pretest/postest1: $F(1, 135) = .18$, $p = .67$; posttest1/posttest2: $F(1, 130) = 1.1$, $p = .29$).

Discussion

Many studies have shown that a higher level of vocabulary results in better comprehension performance (e.g., Ricketts et al., 2007; Sénéchal et al., 2006), and conversely, that poor comprehenders generally exhibit less developed-or less accessible-semantic knowledge

(Nation et al., 2001; Perfetti & Hart, 2001). In addition, some recent studies have underlined the need to distinguish between two facets of vocabulary, namely breadth and depth, the latter appearing as the one that is mainly recruited for text comprehension (Cain & Oakhill, 2014; Ouellette, 2006). Based on this, we have hypothesized that a vocabulary training specifically targeting depth has a greater transfer effect on comprehension than training focused on breadth. To investigate this hypothesis, we developed a new application specifically aimed at improving vocabulary depth, and tested its effectiveness on vocabulary and comprehension performances for texts that did not contain the trained words. After 8 hours of training spread over 4 weeks, results were very encouraging.

First, we observed a beneficial effect of the application on vocabulary depth. It should, however, be noted that this effect was limited to the trained words, and did not extend to the nontrained ones. This finding is not surprising in itself as there was no reason to observe an impact of training on the knowledge of words that children did not encounter during the training sessions. The training exercises we administered, inspired by the methods used to assess vocabulary depth, therefore seemed to be an effective means of improving children's lexical knowledge of words. These exercises focused on the semantic traits associated with each word, as well as on the semantic links between words. Other types of exercise could also be considered. Several studies have provided evidence for the effectiveness of instruction about words' morphological structure with regard to literacy skills (see Bowers, Kirby, & Deacon, 2010, for a review). For example, Lyster, Lervag, and Hulme (2016) evaluated the long-term effectiveness of morphological awareness training delivered in preschool in improving children's reading ability (both word reading and reading comprehension) at the end of Grade 1 and, 5 years later, in Grade 6. They observed both short- and long-term positive effects of the morphological training. Children in the morphological awareness training group had significantly higher scores than children in the control group on both word reading and reading comprehension measures in Grade 1, and still outperformed the control children on reading comprehension measures in Grade 6. Adding morphological training therefore represents a promising possible extension of the ProVoc application.

The main finding of the present study was the transfer we observed from vocabulary training to reading comprehension. This is an original finding, insofar as it shows a clearly significant effect, where previous vocabulary intervention studies found at most a weak transfer to comprehension (Elleman et al., 2009; Wright & Cervetti, 2017). In addition, this transfer effect was observed in a comprehension test using a narrative text that did not include the trained words. At this level, the present study yielded very promising results. The majority of vocabulary intervention studies so far have been able to show a beneficial transfer effect on comprehension tasks featuring the trained words, leading Wright and Cervetti (2017) to conclude that "teaching students the meanings of the words in a passage supported students' comprehension of that passage" (p. 13).

Several hypotheses can be put forward to explain this beneficial effect. First, we believe that the vocabulary training we provided, which encouraged active processing of semantic features of words and semantic links between words, can be an efficient way of training vocabulary depth. As a result, given the causal nature of the relationships between vocabulary depth and reading comprehension highlighted in previous research (e.g., Cain & Oakhill, 2014; Ouellette, 2006), this vocabulary depth training led children to improve their text comprehension skills. Second, it is worth noting that in the reading comprehension test (Potocki et al., 2014) used in the present study, two thirds of the questions were inferential ones (local and global). In line with what Cain and Oakhill (2014) showed, as regards the particular involvement of vocabulary depth in inferential abilities, the use of predominantly inferential comprehension questions may have maximized the effects we observed. This point, however, needs to be confirmed by specific analysis of the effects of our training on different types of comprehension processes. This analysis was not possible here, given the small number (4) of questions in each category (literal, text-connecting inferences, and knowledge-based inferences). Third, it is important to interpret the transfer effect observed in this study with regard to the particular population in which the training was conducted. The children in our study came from schools in priority education zones, where families have a lower socio-economic status. Within this population, comprehension difficulties may tend to be more related to vocabulary difficulties. Many studies (see, for example, Hart & Risley, 1995; Hoff, 2003) have indeed shown that children from families with a low socio-economic status generally have less vocabulary knowledge, owing to differences in their language-learning experiences at home. In line with this consideration, it is interesting to note that the few studies to have shown a beneficial effect of vocabulary training on the level of comprehension also concerned students from disadvantaged socio-economic backgrounds (Beck, Perfetti, & McKeown, 1982), or students with low reading levels (see meta-analysis of Elleman et al., 2009) or poor initial vocabulary (Nelson & Stage, 2007). Thus, different effects might have been observed with a more varied population of students whose comprehension difficulties were linked to other, higher-level deficits. This hypothesis should be explored in future studies. Finally, some authors (Stahl & Fairbanks, 1986; see also Wright & Cervetti, 2017) have suggested that transfer effects on comprehension tasks are also related to incidental effects of vocabulary training (e.g., greater interest in or increased attention to words in texts). As a result, the active semantic processing of words in the ProVoc app may have led students to develop word comprehension strategies such as the one that makes it possible to infer the meaning of new words encountered in texts. This inferencing process is precisely lacking in children with comprehension difficulties (Cain et al. 2004). This last hypothesis particularly holds our attention. Indeed, as seen previously, we did not observe any effect of the intervention on the vocabulary depth task for the untrained words. But in this task, the children had to give - out of any context - definitions of words on the basis of their names only. On the contrary, one can hypothesize that the significant transfer effect on text comprehension (that did not contain the trained words) was observed because the children have benefitted from the intervention to foster the process of inferring the meaning of words presented in context. This hypothesis should be specifically tested in a future study, for example by administering, before and after the training with the ProVoc app, a task in which children have to infer the meaning of new words (e.g, pseudo-words) embedded in texts (see for example the material proposed by Cain et al., 2004 or recently by Beauvais, Pfeiffer, Habib, Beauvais, 2020).

Perfetti and Stafura (2014) put forward another potential explanation for this transfer effect. For these authors, the effects of vocabulary on reading comprehension may be mediated by the influence of vocabulary knowledge on word reading (Tunmer & Chapman, 2012). The retrieval of word meanings through orthographic representations (and their integration with text meaning) would therefore be the critical factor for reading comprehension. Although this explanation sounds plausible, our own study did not support it, as we found no effect of vocabulary training on the word reading measure. This result questions the possibility of mediation by word reading, and instead suggests the existence of a direct link between vocabulary depth and reading comprehension. However, it should be noted that in our app, words were presented in a written format, which may enhance the storage of orthographic form of words in relation to their meanings, even if we did not observe any transfer effect in the word reading task.

Despite these positive results, the present study had several limitations that deserve to be addressed, so that they can be remedied in future research. First, we recruited whole classes of children, meaning that we included both good and poor readers. Future research could focus on more specific profiles of children, especially poor comprehenders, as analysis of previous vocabulary training studies showed that these were particularly beneficial (up to three times more) for less skilled readers (Elleman et al., 2009). However, readers with these profiles are very much in the minority (3-10% of a school-aged population; e.g., Cain & Oakhill, 2006; Catts, Hogan, & Fey, 2003), and the purpose of the present study was not to conduct this more fine-grained analysis, but to provide an initial validation of the app we had developed. In addition, it should be noted that our sample was composed of children from schools in priority education zones, that is, areas where families have a lower socio-economic status and more students have special needs. As discussed earlier, this may partly explain the results observed in this study (including the transfer effect on the comprehension task), and future studies should include a more diverse population. Second, only 8 hours were devoted to the training, and we can postulate that the impact of the intervention would have been greater had the training period been longer. Some research has demonstrated that multiple exposure to target words is more beneficial than single exposure (Beck et al., 1982; McKeown, Beck, Omanson, & Perfetti, 1983). Third, we did not observe any training effect on either vocabulary breadth or word reading performances. For the former, this was particularly surprising, as one might have assumed that, at least for the trained words, fostering their semantic traits would necessarily enhance basic knowledge of the words within the lexicon. We believe that this absence of effect may have been due to the type of task we used to assess vocabulary breadth. As its name makes clear, the Peabody Picture Vocabulary Test uses pictorial representations of words, the aim being to find the correct picture among four options, whereas in our training, no pictures were used at any juncture. As a result, children may have had difficulty matching the word they had seen during the training session with its correct pictorial representation. This explanation should be tested in future studies by using other types of task (with no pictures) to assess vocabulary breadth. Another possible explanation for this lack of effect on vocabulary breadth is that we used a task made of a binary scale (correct *vs* incorrect answers) which is perhaps less sensitive than another type of scale (for example the 3-points one used for the vocabulary depth task). This point should also be considered in future studies.

We also observed an absence of training effect in a standardized test of word reading (Ecalle, 2006). This result might also seem surprising, as vocabulary has been shown to be related to both components of reading, as described in the simple view of reading model (i.e., also related to word reading; Scarborough, 2001; Sénéchal et al., 2006). This absence of effect can therefore be interpreted as the confirmation of a weaker influence of vocabulary depth on word reading (Ouellette, 2006), but it can also be explained by other parameters. For instance, it would have been interesting to look at differences in reading performance (fluency and adequacy) between trained and nontrained words, as has been done for the two vocabulary tasks. In the word reading task we used, all the words were nontrained. We can postulate that more positive effects would have been observed had we included trained words, because of the facilitating effect on reading of already having the word stored in the lexicon (Coltheart, 2005; Harm & Seidenberg, 2004). Finally, even if the use of a randomized waiting-list control group protocol allowed us to test the medium-term maintenance of the training effects (i.e., results on the second posttest for students in the first experimental group), we have yet to study the longer-term effects of this training on children's vocabulary and comprehension performances.

In summary, although several improvements to the current training program or study protocol could be implemented in future studies, the present research nevertheless yielded promising results in terms of vocabulary interventions for improving reading comprehension. This training could have direct implications in the fields of education, remediation or speech therapy targeting children with poor vocabulary knowledge (Nash & Snowling, 2006) and/or language disabilities, especially comprehension difficulties.

References

Beauvais, C., Pfeiffer, N., Habib, M., & Beauvais, L. (2020). Effet de la valence émotionnelle de textes sur leur compréhension et l'apprentissage de nouveaux mots chez des élèves de 5éme année de primaire. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 74*(2), 144.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*(4), 506.

Binder, K. S., Cote, N. G., Lee, C., Bessette, E., & Vu, H. (2017). Beyond breadth: The contributions of vocabulary depth to reading comprehension among skilled readers. *Journal of Research in Reading*, *40*(3), 333-343.

Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, *130*(6), 858.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144-179.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, *11*(5-6), 489-503.

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, *76*(4), 683-696.

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.

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, *114*(4), 647-662.

Cain, K., Oakhill, J., & Bryant, P. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing*, *12*(1-2), 31-40.

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.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151-164.

Cavalli, E., Casalis, S., El Ahmadi, A., Zira, M., Poracchia-George, F., & Cole, P. (2016). Vocabulary skills are well developed in university students with dyslexia: Evidence from multiple case studies. *Research in Developmental Disabilities*, *51*, 89-102.

Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 6–23).Oxford: Blackwell.

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*, 204–256.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311.

Dunn, L., & Dunn, L. (1981). *The Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Dunn, L. M., Dunn, L. M., & Theriault-Whalen, C. (1993). *EVIP: Echelle de vocabulaire en images Peabody*. Toronto, Ontario, Canada: Psycan.

Écalle, J. (2006). Timé-3: Test d'identification de mots écrits. Paris : Mot à Mot Editions.

Eldredge, J., Quinn, B., & Butterfield, D. D. (1990). Causal relationships between phonics, reading comprehension, and vocabulary achievement in the second grade. *The Journal of Educational Research*, *83*(4), 201-214.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.

Fayol, M., Gombert, J. E., Lecocq, P., Sprenger-Charolles, L., & Zagar, D. (1992). *Psychologie cognitive de la lecture*. Paris: PUF.

Florit, E., Roch, M., Altoè, G., & Levorato, M. C. (2009). Listening comprehension in preschoolers: The role of memory. *British Journal of Developmental Psychology*, *27*(4), 935-951.

Gallucci, M. (2019). *GMLj: General analysis for linear models. GAMLj: General analyses for linear models*. [jamovi module].

Goswami, U. (2001). Early phonological development and the acquisition of literacy. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 111–125). New York: Guilford Press.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6-10.

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75-90.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.

Hart, B. & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 156-166.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, *36*(5), 596.

Lyster, S. A. H., Lervåg, A. O., & Hulme, C. (2016). Preschool morphological training produces long-term improvements in reading comprehension. Reading and Writing, *29*(6), 1269-1288.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior*, *15*(1), 3-18.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*, 3–19.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, *53*(2), 253-279.

Miller-Shaul, S. (2005). The characteristics of young and adult dyslexics readers on reading and reading-related cognitive tasks as compared to normal readers. *Dyslexia, 11*, 132–151.

Nagy, W. (2010). The word games. In M. G. McKeown & L. Kucan (Eds.), *Bringing reading research to life* (pp. 72-91). New York: The Guilford Press.

Nagy, W., Berninger, V., & Abbot, R. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle school students. *Journal of Educational Psychology*, *98*(1), 134-147.

Nash, H., & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary knowledge: A controlled evaluation of the definition and context methods. *International Journal of Language & Communication Disorders*, *41*(3), 335-354.

Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (pp. 248–265). Blackwell Publishing.

Nation, K., Marshall, C. M., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. *Language and Cognitive Processes*, *16*(2-3), 241-259.

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.

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.

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel: Teaching children to read*. Washington, DC.

Nelson, J. R., & Stage, S. A. (2007). Fostering the development of vocabulary knowledge and reading comprehension though contextually-based multiple meaning vocabulary instruction. *Education and Treatment of Children*, *30*, 1-22.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*(2), 91-121.

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

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.

Perfetti, C. A. (1985). *Reading ability*. Oxford: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Gorfien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67-86). Washington, DC: American Psychological Association.

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. 227-255). Oxford: Blackwell.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22-37.

Perfetti, C., Yang, C. L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*(3), 303-318.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology, 61*(2), 106-151

Potocki, A., Bouchafa, H., Magnan, A., & Ecalle, J. (2014). Évaluation de la compréhension écrite de récits chez l'enfant de 7 à 10 ans : Vers des profils de compreneurs. *European Review of Applied Psychology*, *64*(5), 229-239.

Potocki, A., Ecalle, J., & Magnan, A. (2013). Effects of computer-assisted comprehension training in less skilled comprehenders in second grade: A one-year follow-up study. *Computers & Education*, *63*, 131-140.

Potocki, A., Magnan, A., & Ecalle, J. (2015). Computerized trainings in four groups of struggling readers: Specific effects on word reading and comprehension. *Research in Developmental Disabilities, 45*, 83-92.

Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing*, *25*(7), 1635-1664.

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: Selection, acquisition, and testing* (pp. 209-227). Amsterdam: John Benjamins.

Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235-257.

Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research*, *95*(5), 259-272.

Scarborough, H.S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp.97–110). New York: Guilford Press.

Sénéchal, M. (2000). Examen du lien entre la lecture de livres et le développement du vocabulaire chez l'enfant préscolaire. *Enfance*, *53*(2), 169-186.

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

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia, 7*(1), 37-46.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72-110.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360-407.

Stanovich, K. E. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. *Advances in Child Development and Behavior*, 24, 133-180).

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.

Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, *16*(2), 122-140.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

van Viersen, S., de Bree, E. H., & de Jong, P. F. (2019). Protective factors and compensation in resolving dyslexia. Scientific Studies of Reading, 23(6), 461-477.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, *11*(1), 3-32.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phonological awareness and early reading ability. *Reading and Writing*, *16*, 5–20.

Webb, S. (Ed.) (2019). *The Routledge handbook of vocabulary studies*. New York: Routledge.

Wechsler, D. (2003b). Wechsler Intelligence Scale for Children–Fourth Edition: Technical and interpretive manual. San Antonio, TX: Psychological Corporation.

Wright, T. S., & Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, *52*(2), 203-226.

Yanagisawa, A., & Webb, S. (2019). Measuring depth of vocabulary knowledge. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 371-386). New York: Routledge.

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. *Written Language & Literacy*, *8*(2), 157-181.