Published in final edited form as:

Read Writ. 2011 July 1; 7: 581–586. doi:10.1007/s11145-010-9260-0.

How Prior Knowledge Affects Word Identification and Comprehension

Sarah J. Priebe, Janice M. Keenan, and Amanda C. Miller

University of Denver

Abstract

While prior knowledge of a passage topic is known to facilitate comprehension, little is known about how it affects word identification. We examined oral reading errors in good and poor readers when reading a passage where they either had prior knowledge of the passage topic or did not. Children who had prior knowledge of the topic were matched on decoding skill to children who did not know the topic so that the groups differed only on knowledge of the passage topic. Prior knowledge of the passage topic was found to significantly increase fluency and reduce reading errors, especially errors based on graphic information, in poor readers. Two possible mechanisms of how prior knowledge might operate to facilitate word identification were evaluated using the pattern of error types, as was the relationship of errors to comprehension. Implications of knowledge effects for assessment and educational policy are discussed.

Beginning with Bransford and Johnson's (1972) seminal study, many researchers have shown that having some prior knowledge about the topic of a passage enables both greater comprehension of the text and better memory for it (McNamara & Kintsch, 1996; Rawson & Van Overschelde, 2008; Recht & Leslie, 1988; Spilich, Vesonder, Chiesi, & Voss, 1979). The effects of prior knowledge on comprehension are so pronounced that researchers have advocated including assessments of prior knowledge in comprehension batteries (Johnston, 1984; Pearson & Hamm, 2005), and Hirsch (2006) designed a reading curriculum that aims to improve comprehension scores by establishing a core of common knowledge for all students.

In contrast, there has been very little research investigating whether prior domain knowledge might also improve word identification. Logically, it is possible that part of the benefit of prior knowledge on comprehension may be because when one has topic knowledge, then the words can be identified more readily. Accurate word identification is a strong predictor of comprehension (Fuchs, Fuchs, & Compton, 2004; Pinnell et al., 1995), and deficits in fluency are thought to be at least partly responsible for comprehension problems for many students (Duke, Pressley, & Hilden, 2004). Indeed, interventions targeted at improving fluent word identification in struggling readers can improve comprehension (Chard, Vaughn, & Tyler, 2002; National Institute of Child Health and Human Development, 2000; Nicholson & Tan, 1999; Stahl & Heubach, 2006; Tan & Nicholson, 1997). So, if it can be shown that differences in prior knowledge lead to differences in how words are identified within a passage, then it would provide further insights into the processes of word reading and comprehension and how one might improve both. We know that attentional resources are limited (LaBerge & Samuels, 1974) and therefore there can be tradeoffs between allocating resources to word decoding and comprehension (Perfetti, 1985). If prior knowledge facilitates word reading, then more resources could be allocated to comprehension processes rather than to word identification, thus facilitating comprehension.

But is there any evidence that having prior topic knowledge does in fact help word identification? There is plenty of evidence that context facilitates word identification, and undoubtedly context taps into prior knowledge to have its effects, but there can be knowledge differences even when there are no context differences and there is almost no research on how topic knowledge affects word identification.

Beginning with Tulving and Gold's (1963) classic study, it has often been shown that words can be read more accurately and quickly in context (e.g., Nation & Snowling, 1998; Perfetti, Goldman, & Hogaboam, 1979; Stanovich, 1984). This is especially true for poor readers, who seem to benefit more from the additional semantic activation provided by the context to compensate for their poor decoding skills (Gough, 1983; Nation & Snowling, 1998; West & Stanovich, 1978). Semantic effects on word identification have even been shown when words are presented in just a single word context or even in isolation (e.g. Nation & Cocksey, 2009; Rodd, 2004; see Keenan & Betjemann, 2007, for a review). However, even though context must exert its effects by evoking prior knowledge, having context is not the same as having prior topic knowledge. This is evident from the fact that it is possible for readers to differ in the amount of background knowledge they have when reading the same text and thus having the same amount of context available. The question we raise in the current study is whether such differences in the amount of prior knowledge matter for word identification. When reading words in the context of a text passage, is it possible that a reader who has prior knowledge of the passage topic will identify the words in the passage better than another reader who has the same level of word decoding skill but does not have prior knowledge about the topic?

The only study that directly assessed the effects of prior topic knowledge on reading errors in developing readers was done by Taft and Leslie (1985). They examined oral reading accuracy in third graders by comparing readings of a passage about the food chain from children who attended different schools, where only one of the schools had provided instruction in the food chain as part of their curriculum. They found that the children with prior knowledge of the subject made few graphically similar errors relative to their counterparts who had no knowledge of the topic, and their errors tended to preserve meaning. This was a promising finding about the potential of topic knowledge to facilitate word identification, but to our knowledge, it has not been followed up in subsequent research. However, a study by Malik (1990) of adults learning English as a foreign language provides some corroborating support in that errors were more likely to be semantically acceptable when reading familiar than unfamiliar passages.

Children with better word decoding skill are also often those who have a wide range of domain knowledge. For example, in our own sample, we find that it is quite easy to find poor readers who do not have prior knowledge of assorted passage topics; however, it is much harder to find good readers who do not have prior knowledge. Better word decoding skill appears to go hand in hand with an increased chance of having prior knowledge. This raises the question of whether the effects that Taft and Leslie (1985) observed were due to better word reading skills or more prior knowledge of the passage. Because they did not have an independent assessment of word decoding skill, we thought it important to determine whether knowledge differences would have an effect even when children are matched on word decoding skill across knowledge groups, defined here as performance on reading words in a list.

In the present study we unconfound word decoding skill and knowledge by matching the children in our knowledge groups on word decoding skills and by including both good and poor readers in both the knowledge group and the no knowledge group. The logic behind matching our knowledge groups on decoding skill is as follows. If we know that two

children perform equally well at reading words in a list, when we then assess their skill at reading a passage for which they have different levels of domain knowledge, we would expect those children to perform equally on reading that passage if domain knowledge has no effect. If, however, we do find differences in oral reading accuracy on the passage, we can be confident that those differences are not due to decoding skill differences and that it is likely domain knowledge differences that underlie the difference in performance on the passage.

This study examines oral reading accuracy in both typically developing readers and poor readers in order to determine whether the effects of prior topic knowledge depend on the reader's skill level in word decoding. Stanovich's (1984) interactive compensatory model of reading suggests that, just as context has been found to have a greater influence on word identification in poor readers than in those with typical reading skill, so too differences in knowledge about the passage topic may have a greater influence on decoding accuracy in poor readers. In fact, a recent study by Miller and Keenan (2009) showed that prior knowledge effects on comprehension are strongest for poor readers. In this paper we therefore examine knowledge effects both in typical readers and in children who are poor readers.

Using Substitution Errors to Identify Processing Mechanisms Underlying Knowledge Effects

This paper not only examines whether prior knowledge facilitates word recognition, but also how. We assume that word identification consists of a set of neural networks that map orthographic input onto sounds and meaning, as shown in Figure 1 (adapted from Plaut, McClelland, Seidenberg, & Patterson, 1996). We examine two possible mechanisms that might operate on this model to yield knowledge effects. One is whether prior knowledge facilitates reading by allowing the reader to skip mapping the orthographic onto the phonological units because the word is so predictable based on the child's knowledge, e.g., "hill" in "Jack and Jill climbed up the..." A child familiar with the nursery rhyme would not even need to look at the word, much less process the orthographic information, because their prior knowledge of the rhyme would allow them to generate the correct word. We refer to this process of bypassing orthographic information as the *bypass process*.

Support for the bypass process would be evident if the child makes a substitution that is semantically similar, but not graphically related; for example, if the text stated "Jack and Jill climbed up the mountain," and the child said "hill" instead of "mountain." Although it may seem unlikely that semantic information can overwhelm the graphemic input so as to produce such errors, even highly skilled readers have shown evidence for just this in a phenomenon known as the Moses Illusion. Studies of semantic illusions like the Moses Illusion have participants answer questions like "How many animals of each kind did Moses take on the ark?" Even after reading the sentence out loud, and being told to look out for distorted questions, participants tend to answer "two", as if "Noah" instead of "Moses" had been included in the question (Erickson & Mattson, 1981; Reder & Kusbit, 1991).

An alternative mechanism by which prior knowledge may facilitate word identification is that it gives the network that maps orthographic to phonological and semantic units a boost, rather than bypassing the mapping, so that it has more information to use in the constraint satisfaction process and can more readily and more accurately settle on a solution. We refer to this process as the *constraint satisfaction* use of semantic activation from prior knowledge. Although this semantic boost would generally keep one from making errors, evidence for this process would be reflected by a decrease in substitutions that are graphically similar and semantically dissimilar to the original word. Whereas a bypass

mechanism leads to substitutions that may not be at all related graphically, like “hill” is to “mountain”, the constraint satisfaction process would lead to fewer substitutions that are only graphically similar, and not semantically similar, for example something like “maintain” for “mountain.” Just as context constrains possible choices for a word, so might having prior knowledge, resulting in fewer miscues that are based solely on graphic information.

It is important to note that both models predict more semantically similar miscues as an outcome of having prior knowledge, due to the added contribution of semantics. The difference in the two models lies in the relative contributions of graphic and semantic information, where the constraint satisfaction model predicts fewer errors based on only graphic information, and the bypass model predicts more errors based on only semantic information.

In order to determine the relative occurrence of these two processing mechanisms, we examined oral reading substitution errors for their graphic and semantic similarity to the original word. This approach is labeled miscue analysis (Goodman, 1969), and has a long history of use, with demonstrated validity and reliability (Murphy, 1999; Sadoski, Carey, & Page, 1999). It was used by Taft and Leslie (1985) when they explored the effects of prior knowledge; however, because they averaged over semantic similarity when examining graphic similarity, and vice versa, it is not possible to determine from their results the mechanism by which prior knowledge is operating. To do that, we need to do more than just examine overall graphic similarity and overall semantic similarity; we need to classify a substitution on both dimensions simultaneously, as we do in the current study. Table 1 shows the types of substitutions and examples of each kind.

If prior knowledge allows readers to bypass word identification and “read” what they think the word should be based on their knowledge, then we expect to find that they will show this by having more G–S+ substitutions than readers with no knowledge. Alternatively, if prior knowledge gives a boost to the constraint satisfaction process of mapping, then we may see fewer substitutions that are G+S–. It is common to find that when children who are better readers do make errors, their oral reading errors are more semantically similar than those of poorer readers (e.g., Goodman & Goodman, 1977; Leslie, 1980). More semantically similar errors is also what Taft and Leslie (1985) reported when children had prior domain knowledge. Again, to dissociate any effects that might be due to better word decoding instead of prior knowledge, we matched our knowledge groups on decoding.

Because we expect that having prior knowledge leads to more automatic word recognition, we will also examine the fluency of good and poor readers with and without prior knowledge. We expect that having prior knowledge will lead to more accurate and faster word identification, particularly for poor readers. In addition, because previous research has demonstrated the importance of studying self-corrections (Beebe, 1980; Sadoski et al., 1999), and because previous research has failed to find a relationship between prior knowledge and self-corrections (Malik, 1990; Taft & Leslie, 1985), we will also examine the rate of self-corrections as a function of prior knowledge and reading ability.

A final aim of this study is to examine the relationship between oral reading accuracy and comprehension to explore whether certain types of errors predict reading comprehension more strongly than the overall number of errors. In particular, based on previous research that shows that poor readers make more miscues that change the meaning of the text (Goodman & Goodman 1977; Leslie, 1980), and given previous findings that semantically dissimilar miscues are more related to comprehension than other types of errors (Pinnell et

al., 1995), we expect that semantically dissimilar substitutions may be the most strongly related to comprehension.

In sum, the purpose of the present study is to investigate the mechanisms by which prior knowledge might facilitate word identification. By being able to match our participants on age and word decoding skill so that they only differ in their prior knowledge of the passage topic and by examining both fluency and the types of oral reading errors, we can determine whether and how knowledge facilitates word identification. By examining knowledge effects in both typical readers and in poor readers, we can also determine whether there is differential facilitation on oral reading for poor readers with prior knowledge. Finally, we examine the relationship between oral reading errors and comprehension to determine if the type of oral reading error, or the total number of errors, is more predictive of reading comprehension; if a specific type of error is more predictive, then it would suggest that the extra effort required to classify error types is warranted.

Method

Participants

The data analyzed for this study were previously collected as part of a language comprehension assessment battery in a larger ongoing research project conducted by the Colorado Learning Disabilities Research Center (c.f. Keenan et al, 2006; Olson, 2006). The sample for the present study was much smaller than the full sample, which consisted of over 130 children in the 4th grade, because we had to match the participants in the prior knowledge group with children in the no prior knowledge group on single word reading. Because we were doing this matching for both poor readers and for controls and because there is a lower rate of poor readers who had topic knowledge and controls who had no prior knowledge of the topic, we were fortunate to have a large number of children to select from so that we could closely match our knowledge groups on decoding skill. For this study the participants consisted of 60 4th graders (33 female), with a mean age of 9.7, originally recruited to the larger study either on referral for a reading disability (RD) or as controls.

Thirty participants comprised the prior knowledge group, with 15 being poor readers and 15 good readers. The no prior knowledge groups also had 30 participants who were matched to the individuals in the knowledge group on word decoding skill. The method for defining knowledge groups and reading abilities groups is given below.

Table 2 presents descriptive statistics on age and word decoding skill for the four groups, as well as vocabulary skill and listening comprehension. Decoding skill was measured with two tests, the Timed Oral Reading of Single Words (Olson, Forsberg, Wise, & Rack, 1994) and the Peabody Individual Achievement Test (PIAT) word recognition subtest (Dunn & Markwardt, 1970). Each raw decoding score was standardized across the larger sample, and these were combined into a composite in order to have a more reliable measure. The poor readers had an age-adjusted composite decoding z-score of -1 or below. The controls all had word decoding composite z-scores above zero, indicating above average word decoding skills. The matching of individuals in the prior and no prior knowledge groups on word decoding was evaluated by t-tests that revealed no significant differences in decoding skill between the knowledge groups either for poor readers ($t(28) = .12, p = .91$) or for good readers ($t(28) = .46, p = .65$).

Vocabulary was measured using the vocabulary subtest of the Wechsler Intelligence Scale for Children, 3rd. ed. (WISC-III; Wechsler, 1991), and listening comprehension was measured using the oral comprehension subtest of the Woodcock-Johnson Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001). Interestingly, although

readers were not explicitly matched on either listening comprehension or vocabulary, t-tests revealed no significant differences on these variables between those who had topic knowledge and those who did not. For listening comprehension there were no significant differences between the knowledge groups either among the poor readers ($t(28) = .08, p = .94$) or among the good readers ($t(28) = 1.12, p = .23$), and likewise for vocabulary, among poor readers ($t(28) = .71, p = .49$) or among good readers ($t(28) = 1.1, p = .30$).

Materials and Background Knowledge Assessment

This study used a 263-word expository passage, titled “Amelia Earhart” from the *Qualitative Reading Inventory* (QRI; Leslie & Caldwell, 2001). Although the children were tested on other QRI passages as well, they were not suitable to testing our hypotheses because either all children knew about the topic (e.g., octopus) or none knew anything (e.g., Andrew Carnegie). The Amelia Earhart passage provided a sufficient variety in children's background knowledge because some children knew about her and some children had never heard of her. Topic familiarity was assessed by a concept question that accompanies the passage in the QRI manual; it always preceded the child's reading. It was scored either as a 0 (no prior knowledge of the topic) or 1 (some prior knowledge of the topic). Reliability of assessing knowledge was assessed by having multiple examiners score a subset of concept questions, and there were no discrepancies among the ratings of knowledge. Previous research using this method of knowledge assessment on this same passage (Miller & Keenan, 2009) showed it to have sufficient sensitivity to detect knowledge differences that significantly impact comprehension. Given that it has established sensitivity, it seemed preferable to assess knowledge with it rather than more extensive questioning because it avoids the potential problem of telling the child, by way of multiple knowledge assessment questions, what the text is about before they even read it.

Procedure

Before the child read the passage, the tester asked the child a background knowledge question, “Who is Amelia Earhart?” to gauge how much the child knew about the topic of the passage. Answers such as “A pilot” or “The first woman to fly around the world” would be scored as demonstrating prior knowledge, whereas incorrect responses or “I don't know” would be scored as demonstrating no prior knowledge. The child then read the passage orally, while the examiner scored the passage online for rate and accuracy. The child then provided a free recall of the passage. After the testing session, the recalls were scored according to the number of idea units recalled from the idea checklist that accompanies the test. A subset of recalls were scored by multiple raters, and inter-rater reliability of this subset was very high (Cronbach's $\alpha = .97$). The readings were transcribed and coded according to the taxonomy described in the next section.

Substitution Error Scoring

Any substitution that was fully sounded out was counted as a substitution error. The taxonomy outlined here is based on the one used by Taft and Leslie (1985). Table 1 gives examples of each type of substitution.

Graphically Similar (G+)—Substitutions were classified to the degree to which they were graphically similar to the target word by assigning 1 point for each letter shared by the substitution and the target word, and in the same relative order in both words. After totaling the points, they were divided by the number of letters in the target word. Results greater than .5 were judged to be graphically similar, and results less than .5 were judged graphically dissimilar.

For example, if a child is reading the word *mechanical* and produces *michinical*, the substitution would be assigned a graphic similarity score of .80 (8 shared letters, divided by the total number of letters in the word). A substitution of *machine*, in contrast, would receive a score of .40 and would not be judged graphically similar. Using Visual Basic and Microsoft Excel, a macro was created which automatically calculates the graphic similarity of each substitution (Priebe & Keenan, 2009).

Semantically Similar (S+)—While we would have preferred to use a quantitative measure of semantic similarity, previous attempts to do so by using latent semantic analysis (LSA; Landauer, 2002) resulted in very low agreement with human ratings of semantic similarity (Priebe & Keenan, 2009). Therefore, we relied on human ratings of semantic similarity for this study. Substitutions were categorized according to the degree to which they changed the intended meaning of the author of the passage, by reading the sentence up to the substitution and determining if the resulting sentence retained the author's intended meaning for the passage. For example, if a child substituted “inventor” in the sentence “Amelia Earhart was an adventurer”, the substitution would be judged as semantically dissimilar since it changed the intended meaning. A substitution of “adventuresome”, in contrast, would be judged semantically similar.

Other Types of Errors—The number of substitutions, omissions, insertions, repetitions, and skipped lines were summed to create an overall number of errors for each participant. Self-corrections were defined as any error that was subsequently successfully corrected by the child, and these were also summed to create an overall number of self-corrections for each participant.

Reliability of Error Scoring—Intra-rater reliability was assessed through use of BBEDIT, a text-editing software that allows multi-file searches. Each coded word in a transcript was checked against all other occurrences of that word in the remaining transcripts, ensuring that all miscues were categorized in the same way. Inter-rater reliability was obtained by having a trained second observer code 10% of the transcripts for the above categories. Reliability between observers was quite high for all categories, with Cronbach's alphas of .9 for all categories except for substitutions that were semantically dissimilar (Cronbach's $\alpha = .80$).

Results

Comprehension

A 2×2 analysis of variance (ANOVA) was conducted with prior knowledge (0,1) and reading ability (Poor Readers, Controls) as the between-participants independent variables and the total amount of idea units recalled as the dependent variable. Controls recalled significantly more of the text ($M=19.47$) than poor readers ($M=16.10$; $F(1, 56) = 5.19, p < .05$; partial $\eta^2 = .09$), but there was no main effect of prior knowledge ($F(1, 56) < 1$). Interestingly, reading ability interacted with prior knowledge ($F(1, 56) = 4.99, p < .05$; partial $\eta^2 = .08$). When poor readers had prior knowledge, they recalled significantly more idea units ($M = 18.23$) than poor readers without prior knowledge ($M = 13.93$); $t(28) = 1.9, p < .05$; however, for controls, knowledge did not have a significant effect. Thus, in this case prior knowledge appears to play a compensatory role for poor decoding skills.

¹Exceptions to the above rules concern substitutions for 2-letter words. Because these words are so short, applying the same rules resulted in many substitutions being classified as graphically similar when they only shared one letter and did not really seem to be retaining much graphic similarity (e.g. the substitution of *that* for *an*). In the case of 2-letter words, only substitutions that share the initial letter of the target word were counted as graphically similar.

Fluency

A fluency score was calculated for each participant by dividing the total words read correctly over the time taken to read the passage. Table 3 shows average fluency scores for each group. As expected, typical readers had higher fluency scores than poor readers ($F(1, 56) = 55.15, p < .05$). There was a trend toward an effect of prior knowledge improving fluency ($F(1, 56) = 3.10, p = .08$) and a trend toward an interaction between reading ability and prior knowledge ($F(1, 56) = 3.05, p = .09$). Planned t-tests revealed that poor readers with prior knowledge read significantly more words correctly per minute compared to poor readers without prior knowledge ($t(24.67) = 2.25, p < .05$), whereas there was no significant difference in fluency among good readers with and without prior knowledge.

Total Number of Oral Reading Errors

Table 3 shows the average total of all oral reading errors, substitution errors, and self-corrections for each group as a function of prior knowledge and reading ability. As would be expected, poor readers made more total errors ($F(1, 56) = 17.00, p < .05$; partial $\eta^2 = .23$) and more substitutions ($F(1, 56) = 28.10, p < .05$; partial $\eta^2 = .33$) than controls. Poor readers also made more self-corrections compared to good readers ($F(1, 56) = 7.91, p < .05$; partial $\eta^2 = .12$), but there was no effect of prior knowledge, and no interaction between prior knowledge and reading ability on the number of self-corrections made. There was a trend towards an effect of prior knowledge on the total number of errors ($F(1, 56) = 3.07, p = .09$). The interaction between prior knowledge and reading ability on total errors failed to reach significance ($F(1, 56) = 2.20, p = .14$), but planned t-tests revealed a significant difference in the total number of errors between poor readers with and without prior knowledge ($t(17.7) = 1.83, p < .05$). Prior knowledge did not have a significant effect on the total number of substitutions ($F(1, 56) = 2.01, p = .16$), and although there was no significant difference in the number of substitutions between poor readers with and without prior knowledge ($t(28) = 1.3, p = .10$), knowledge effects were evident in the types of substitution errors made, as discussed next.

Types of Substitution Errors

More substitutions were graphically similar than dissimilar ($F(1, 56) = 78.93, p < .05$; partial $\eta^2 = .59$), but the opposite pattern was seen for semantic similarity: more substitutions were semantically dissimilar than semantically similar ($F(1, 56) = 13.04, p < .05$; partial $\eta^2 = .19$). As a result, the interaction of graphic and semantic similarity was significant ($F(1, 56) = 18.24, p < .05$; partial $\eta^2 = .25$) and suggested that overall, when a word could not be identified, readers used graphic cues more than substituting words that made sense within the passage. What is interesting, however, is that the specific type of substitution differed by reading ability and prior knowledge; the predicted four-way interaction between reading ability, prior knowledge, graphic similarity, and semantic similarity was significant ($F(1, 56) = 4.09, p < .05$; partial $\eta^2 = .07$). Figure 2 shows the group differences for each of the four types of specific errors (G+S+, G-S+, G+S-, G-S-), and Table 4 shows the means and standard deviations for each of these points.

We can examine this interaction to determine exactly how prior knowledge of the topic is affecting the word identification process in poor readers. We postulated that one way knowledge might be used by poor readers is to allow them to bypass orthographic-phonological mappings when the orthographic pattern isn't immediately identifiable and produce a word that fits the context; that is, produce a word based only on the semantics of the passage. If this were the case, poor readers who had prior knowledge would produce more substitutions that were semantically related but graphically dissimilar than poor readers who had no prior knowledge. Although the number of substitutions that were semantically similar but graphically dissimilar (G-S+) was greater in poor readers who had

prior knowledge ($M=2.20$) than those without ($M=1.40$), the difference was not significant ($t(28) = .88, p = .39$), and thus did not provide much support for the bypass hypothesis as the model underlying knowledge effects.

The other way we proposed that knowledge could facilitate performance in poor readers is that instead of allowing bypass of normal word identification processes, knowledge could simply add to the process and increase the contribution of semantics to the network's settling. If this is the effect knowledge is having, then substitutions should have some semantic similarity to the target; substitutions that are semantically unrelated and just look like the target should decline. Indeed, as Figure 2 illustrates, this appears to be the effect that prior knowledge had: it dramatically reduced the number of semantically unrelated substitutions of the G+S- type. Poor readers who had no prior knowledge of the topic made almost twice as many substitutions that were graphically similar but semantically dissimilar compared to poor readers with prior knowledge ($t(28) = 2.19, p < .05$). Interestingly, control readers showed similar numbers of G+S- errors whether or not they had prior knowledge.

Error Types and Comprehension

Because we found that prior knowledge and reading ability affected both comprehension and the types of errors made in word identification, we examined whether some error types predicted comprehension more than others. Table 5 shows descriptive statistics for the error types and comprehension measures, as well as the correlations between the types of errors and performance on our measure of comprehension. The total number of errors correlated strongly with performance on comprehension, and all but one of the error types were also strongly correlated with comprehension performance. Interestingly, G+S- errors, or those that were solely based on graphic information without using contextual or semantic cues, were diagnostic of the lowest levels of comprehension. Using the Fisher's-z test for the significance of the difference between two correlations, we found that G+S- errors predicted recall performance more strongly than the overall number of errors ($Z = 3.71, p < .05$).

Discussion

The goal of this study was to determine whether differences due to prior knowledge of the passage topic might extend beyond their well-established impact on comprehension processes and be found to influence word identification processes as well. We examined the types of errors produced in oral reading by readers who were matched on word reading accuracy on lists but who differed in their knowledge of the passage topic to see if differences in prior knowledge affect the types of errors they produced. We found that knowledge did indeed affect word identification, but the effect was confined to the poor readers. Poor readers with prior knowledge read the passage more fluently and made fewer errors than poor readers without prior knowledge. More interesting, they showed a different pattern of errors. Having prior topic knowledge greatly reduced the tendency poor readers had to produce substitutions that were only graphically similar to the target. Thus, poor readers with prior knowledge showed a pattern of errors more closely resembling that of controls. In contrast, poor readers without prior knowledge showed more reliance on graphic information as opposed to semantic information from the text.

As in previous research (Malik, 1990; Taft & Leslie, 1985), we found no effect of prior knowledge on self-corrections. Perhaps the effects of prior knowledge are constrained to initial word identification only. Participants in our study, however, made relatively few self-corrections, which may explain the null results. Future studies of the effects of prior knowledge on self-correction behavior might use longer and harder passages to try to

increase the number of errors and self-corrections in order to be able to more accurately assess its effects.

Our results show not only that topic knowledge affects word identification in a passage but also how it does so. We examined the types of oral reading errors to assess two possible mechanisms by which prior knowledge might be operating: a general speeding up of the constraint satisfaction process, in which we would see fewer G+S- (e.g. “maintain” for “mountain” in the Jack and Jill example) substitutions among those with prior knowledge, and a bypass of the orthographic-phonological information, which would result in more G-S+ substitutions for those with prior knowledge (e.g., “hill” for “mountain” in the Jack and Jill example). Although it is possible that both processes might underlie how knowledge facilitates oral reading, we did not find much support for the notion that knowledge leads readers to bypass orthographic processing because there was not a significant difference in G-S+ errors as a function of knowledge. It is possible, however, that the relatively low rate of G-S+ errors overall might have led to our lack of support for the bypass model. We did, however, find support for the idea that the locus of the knowledge effect in poor readers is in facilitating the constraint satisfaction process. Strong activation of semantic information through prior knowledge of the passage topic seemed to constrain the available choices for a word's identification, which led to a reduction in errors that were solely based on graphic information (G+S-). Thus, prior knowledge appears to serve a compensatory function for poor readers, allowing them to utilize the additional semantic information to help determine the word's identity.

The model of knowledge effects that our data supported for the poor readers suggests one reason why we may not have found similar knowledge effects in our typical readers. Perhaps because typical readers had adequate orthographic-phonological mappings, at least when they are reading grade-level passages, their networks may have settled too quickly on a solution, leaving little room for semantics to exert its effects. Future research could explore this interpretation by determining whether typical readers would show knowledge effects when reading texts above their decoding level or with less familiar vocabulary, in other words when some word identification would be expected to be slowed.

A possible mechanism underlying prior knowledge facilitation in constraint satisfaction is that it may allow poor readers to gain new decoding skills. For instance, having knowledge that allows one to anticipate the word *climbed* in “*Jack and Jill climbed up...*” permits the phonological representation of *climbed* to get activated more quickly than it would from the orthography alone since the mapping involves the relatively low frequency silent *b*. Having the phonological representation activated would allow connections to be established between the phonology and the orthography which could then feed back to facilitate connections between orthography and phonology and allow the child to develop knowledge about the silent *b*. If so, then prior topic knowledge would function like other language effects, such as grammatical sensitivity and vocabulary, which, as Tunmer and Chapman (2004) pointed out, have more effect on word identification when the miscues are graphically similar. It may be that prior knowledge is only effective for poor readers when their graphonological decoding skills are developed enough to make miscues that are graphically similar to the text word. Our participants were in the fourth grade, so presumably their decoding skills were developed enough for this to happen, but future studies could examine younger readers to see if the same pattern holds for readers with less developed decoding skills.

Our results reinforce the earlier findings of Taft and Leslie (1985) showing that topic knowledge influences word identification. They also extend those results by elucidating the mechanisms underlying knowledge effects and also by separating knowledge effects from

reading skill. Children who have broad prior topic knowledge often tend to be better readers, making it difficult to know whether differences between knowledge groups are reflecting just knowledge differences or also other differences. Indeed, when we selected children to form our knowledge groups, there were far fewer typical readers than poor readers who had no prior knowledge of the topic. However, because we started with such a large sample, we were able to define knowledge groups so that children were matched to have equivalent word decoding accuracy when reading word lists and only differed in prior knowledge. In addition, there were no vocabulary or comprehension skill differences between children who had knowledge of the passage topic and those who did not. Thus, our findings provide a straightforward, unconfounded demonstration of the effects of knowledge on oral reading.

We also replicated previous research showing that prior knowledge facilitates comprehension (Miller & Keenan, 2009; Recht & Leslie, 1988; Spilich, Vesonder, Chiesi, & Voss, 1979), although the effect was apparent only in the poor readers. Miller and Keenan (2009) posited that one way prior knowledge might help poor readers is by allowing less effortful construction of the text representation when ideas presented in the passage are already known to the reader. Our results support this idea, as poor readers with prior knowledge had significantly higher comprehension scores compared to poor readers without prior knowledge. Because we found that topic knowledge also improves fluency for poor readers, it is possible that the facilitation of knowledge on comprehension for poor readers also stems from their fluency – fewer cognitive resources are required for word identification and thus more are available for comprehension (Fuchs et al., 2004; Pinnell et al., 1995).

One type of error, G+S–, predicted reading comprehension more strongly than the overall number of errors did. Readers who relied solely on graphic information and did not integrate semantic information from the text had the lowest reading comprehension. One might think that our findings on error type predicting comprehension could be explained by restriction of range because, compared to the other types of errors, G+S– errors had a larger range in our sample. However, it is important to note that the total number of errors actually had a larger range, but it predicted comprehension less strongly. Thus, it appears that examining the types of errors is useful in predicting reading comprehension and may be worth the extra effort because it can provide a more precise method of predicting reading comprehension compared to only looking at the total number of errors.

We have shown here that a child's skill at accurately reading a passage depends not just on their decoding skill but also on their knowledge of that passage topic; knowledge that also facilitates comprehension. While it will be important to replicate this study with other materials, our findings could have important implications for both remediation and assessment. Expanding children's background knowledge could serve as an additional avenue to pursue not only for facilitating comprehension processes, but also word identification. In addition, these studies point to the importance of taking knowledge into account when assessing a child's reading skill level. As we have shown, a poor reader can look more like a typical reader in their errors if the child has prior knowledge of the topic. Ironically, domain knowledge can be a double-edged sword for poor readers: they may appear to be stronger readers than they actually are when reading passages on familiar topics and may not be recognized as needing special services. More likely, struggling readers will be seen as having poor word decoding when in fact part of their struggles may also reflect low familiarity with the topic. Either way, the results from this study indicate that taking a reader's prior knowledge of a topic into account is important.

Our participants were closely matched – decoding skill, vocabulary, and listening comprehension – and yet we found a difference in oral reading errors, fluency, and

comprehension between poor readers who had topic knowledge and those that did not. This suggests that building knowledge might be an additional avenue to pursue for improving literacy. When one hears data such as those from the National Institute of Child Health and Human Development Early Child Care Research Network's (2005) study showing that 57% of the time in 3rd grade classrooms is spent in literacy and language activities, whereas only 6% is spent in science and social studies activities, it is tempting to view them as in direct conflict. However, with results such as ours, and those of Taft and Leslie (1985) and Malik (1990), it appears that gaining subject knowledge can also improve literacy. Perhaps then we need not view instruction in content areas as taking time away from literacy but as a potential addition to it (Hirsch, 2006; Kamhi, 2007).

References

- Beebe MJ. The effect of different types of substitution miscues on reading. *Reading Research Quarterly*. 1980; 3:324–336.
- Bransford J, Johnson M. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*. 1972; 11(6):717–726.
- Chard D, Vaughn S, Tyler B. A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*. 2002; 35(5):386–406. [PubMed: 15490537]
- Duke, NK.; Pressley, M.; Hilden, K. Difficulties with reading comprehension. In: Stone, CA.; Silliman, ER.; Ehren, BJ.; Apel, K., editors. *Handbook of language and literacy: Development and Disorders*. Guilford Press; New York: 2004. p. 501–520.
- Dunn, LM.; Markwardt, FC. *Examiner's Manual: Peabody Individual Achievement Test*. American Guidance Service; Circle Pines, MN: 1970.
- Erickson T, Mattson M. From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*. 1981; 20(5):540–551.
- Fuchs L, Fuchs D, Compton D. Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*. 2004; 71(1):7–21.
- Goodman KS. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly*. 1969; 5:9–30.
- Goodman K, Goodman Y. Learning about psycholinguistic processes by analyzing oral reading. *Harvard Educational Review*. 1977; 47(3):317–333.
- Gough, P. Context, form, and interaction. In: Rayner, K., editor. *Eye Movements in Reading*. Academic Press; New York: 1983. p. 203–211.
- Hirsch, ED. Building Knowledge: The case for bringing content into the language arts block and for a knowledge-rich curriculum core for all children. *American Educator*. 2006. Retrieved from http://www.aft.org/pubs-reports/american_educator.htm
- Johnston P. Prior knowledge and reading comprehension test bias. *Reading Research Quarterly*. 1984; 19:219–239.
- Kamhi A. Knowledge deficits: The true crisis in education. *The ASHA Leader*. 2007; 12:28–29.
- Keenan JM, Betjemann RS, Wadsworth SJ, DeFries JC, Olson RK. Genetic and environmental influences on reading and listening comprehension. *Journal of Research in Reading*. 2006; 29:79–91.
- Keenan, JM.; Betjemann, RS. Comprehension of single words: The role of semantics in reading and reading disability. In: Grigorenko, E.; Naples, A., editors. *Single Word Reading*. Taylor & Francis; London: 2007. p. 191–210.
- LaBerge D, Samuels S. Toward a theory of automatic information processing in reading. *Cognitive Psychology*. 1974; 6(2):293–323.
- Landauer, TK. Applications of latent semantic analysis. In: Gray, WD.; Schunn, CD., editors. *Proceedings of the Twenty-fourth Annual Conference of the Cognitive Science Society*. Psychology Press; London: 2002.

- Leslie L. The use of graphic and contextual information by average and below-average readers. *Journal of Reading Behavior*. 1980; 12:139–149.
- Leslie, L.; Caldwell, J. *Qualitative Reading Inventory – 3*. Addison Wesley Longman, Inc.; New York: 2001.
- Malik AA. A psycholinguistic analysis of the reading behavior of EFL-proficient readers using culturally familiar and culturally unfamiliar expository tests. *American Educational Research Journal*. 1990; 27:205–223.
- McNamara DS, Kintsch W. Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*. 1996; 22:247–282.
- Miller AC, Keenan JM. How word decoding skill impacts text memory: The centrality deficit and how domain knowledge can compensate. *Annals of Dyslexia*. 2009; 59:99–113. [PubMed: 19475514]
- Murphy, S. The validity and reliability of miscue analysis. In: Marek, AM.; Edelsky, C., editors. *Reflections and connections: Essays in honor of Kenneth S. Goodman's influence on language education*. Hampton Press; Cresskill, NJ: 1999. p. 95-122.
- Nation K, Cocksey J. Beginning readers activate semantics from sub-word orthography. *Cognition*. 2009; 110:273–278. [PubMed: 19108820]
- Nation K, Snowling M. Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*. 1998; 69(4):996–1011. [PubMed: 9768483]
- National Institute of Child Health and Human Development. Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. U.S. Government Printing Office; Washington, DC: 2000. NIH Publication No. 00-4769
- National institute of Child Health and Human Development. A day in third grade: A large-scale study of classroom quality and teacher and student behavior. *The Elementary School Journal*. 2005; 105(3):305–323.
- Nicholson, T.; Tan, A. Word identification processes. In: Thompson, GB.; Nicholson, T., editors. *Learning to read: Beyond phonics and whole language*. Teachers College Press; New York: 1999. p. 150-173.
- Olson RK. Genes, environment, and dyslexia: The 2005 Norman Geschwind memorial lecture. *Annals of Dyslexia*. 2006; 56:205–238. [PubMed: 17849199]
- Olson, RK.; Forsberg, H.; Wise, B.; Rack, J. Measurement of word recognition, orthographic, and phonological skills. In: Lyon, GR., editor. *Frames of Reference for the Assessment of Learning Disabilities: New Views on Measurement Issues*. Paul H. Brookes Publishing Co.; Baltimore: 1994. p. 243-277.
- Pearson, P.; Hamm, D. The assessment of reading comprehension: A review of practices-past, present, and future. In: Paris, SG.; Stahl, SA., editors. *Children's reading comprehension and assessment*. Lawrence Erlbaum Associates Publishers; Mahwah, NJ US: 2005. p. 13-69.
- Perfetti, CA. *Reading ability*. Oxford University Press; Oxford: 1985.
- Perfetti C, Goldman S, Hogaboam T. Reading skill and the identification of words in discourse context. *Memory & Cognition*. 1979; 7(4):273–282.
- Pinnell, GS.; Pikulski, JJ.; Wixson, KK.; Campbell, JR.; Gough, PB.; Beatty, AS. US Department of Education, Office of Educational Research and Improvement; Washington, DC: 1995. *Listening to children read aloud*.
- Plaut D, McClelland J, Seidenberg M, Patterson K. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*. 1996; 103(1):56–115. [PubMed: 8650300]
- Priebe, SJ.; Keenan, J. Automated Methods of Assessing Oral Reading Accuracy; Poster presented at the annual meeting of the Society for the Scientific Study of Reading; Boston, MA: June. 2009
- Rawson KA, Van Overschelde JP. How does knowledge promote memory? The distinctiveness theory of skilled memory. *Journal of Memory and Language*. 2008; 58:646–668.
- Recht DR, Leslie L. Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*. 1988; 80:16–20.
- Reder L, Kusbit G. Locus of the Moses Illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language*. 1991; 30(4):385–406.

- Rodd J. When do leotards get their spots? Semantic activation of lexical neighbors in visual word recognition. *Psychonomic Bulletin & Review*. 2004; 11(3):434–439. [PubMed: 15376791]
- Sadoski, M.; Carey, RF.; Page, WD. Empirical evidence for the validity and reliability of miscue analysis as a measure of reading comprehension. In: Marek, AM.; Edelsky, C., editors. *Reflections and connections: Essays in honor of Kenneth S. Goodman's influence on language education*. Hampton Press; Cresskill, NJ: 1999. p. 123-144.
- Spilich G, Vesonder G, Chiesi H, Voss J. Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior*. 1979; 18(3):275–290.
- Stahl, S.; Heubach, K. Fluency-oriented reading instruction. In: Stahl, KAD.; McKenna, MC., editors. *Reading research at work: Foundations of effective practice*. Guilford Press; New York, NY US: 2006. p. 177-204.
- Stanovich K. The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *RASE: Remedial & Special Education*. 1984; 5(3):11–19.
- Taft M, Leslie L. The effects of prior knowledge and oral reading accuracy on miscues and comprehension. *Journal of Reading Behavior*. 1985; 17(2):163–179.
- Tan A, Nicholson T. Flashcards revisited: Training poor readers to read words faster improves their comprehension of text. *Journal of Educational Psychology*. 1997; 89:276–288.
- Tunmer, WE.; Chapman, JW. The use of context in learning to read. In: Bryant, P.; Nuñez, T., editors. *Handbook of Children's Literacy*. Dordrecht, the Netherlands; Kluwer Academic.: 2004. p. 199-212.
- Tulving E, Gold C. Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*. 1963; 66(4):319–327. [PubMed: 14053250]
- Wechsler, D. *Manual for the Wechsler Intelligence Scale for Children- Third Edition*. The Psychological Corporation; San Antonio, TX: 1991.
- West R, Stanovich K. Automatic contextual facilitation in readers of three ages. *Child Development*. 1978; 49(31):717–727.
- Woodcock, RW.; McGrew, KS.; Mather, N. *Woodcock-Johnson III Tests of Achievement*. Riverside Publishing; Itasca, IL: 2001. Examiner's manual.

Figure 1. Illustration of two hypothesized mechanisms underlying knowledge effects: a) bypass hypothesis, and b) constraint satisfaction hypothesis. Figure adapted from Plaut et al., (1996).

Figure 2.

Mean number of each type of oral reading substitution as a function of child's prior knowledge (PK) of passage topic and reading ability.

*Note: G= Graphic, S= Semantic, NPK= No prior knowledge, PK = Prior knowledge

Table 1

Examples of the Types of Graphic and Semantic Errors, Using the Example Sentence: “Jack and Jill climbed up the mountain.”

Type of error	Example (target word: <i>mountain</i>)
Graphically and semantically similar (G+S+)	mound
Semantically similar but graphically dissimilar (G-S+)	hill
Graphically similar but semantically dissimilar (G+S-)	maintain
Neither graphically nor semantically similar (G-S-)	apple

Table 2

Average Age, Word Decoding, Listening Comprehension, and Vocabulary Skill for Participants

	Poor Readers No Prior Knowledge (n = 15)	Poor Readers Prior Knowledge (n = 15)	Controls No Prior Knowledge (n = 15)	Controls Prior Knowledge (n = 15)
Age (years)	9.62 (.41)	9.72 (.53)	9.80 (.53)	9.60 (.44)
Word Decoding (composite z-score)	-1.54 (.56)	-1.52 (.32)	.49 (.36)	.55 (.34)
Piat Word Recognition Raw Score	36.60 (6.14)	36.87 (3.52)	50.27 (3.97)	51.60 (4.55)
Timed Oral Reading of Single Words	67.87 (22.71)	67.80 (16.40)	119.47 (17.36)	118.67 (17.97)
Woodcock-Johnson Oral Comprehension Raw Score	19.27 (4.23)	19.13 (4.67)	21.40 (3.02)	22.47 (2.13)
WISC Vocabulary Raw Score	26.27 (7.25)	27.87 (4.98)	32.20 (3.97)	33.80 (4.35)

Average Fluency and Accuracy Measures for Each Group (standard deviations in parentheses)

Table 3

	Words read correctly per minute	Total number of errors	Total number of substitutions	Total Number of self-corrections
Poor Readers				
No Prior Knowledge	66.02 (16.75)	27.47 (15.56)	19.73 (12.37)	3.47 (2.48)
Prior Knowledge	83.30 (24.63)	19.07 (5.86)	14.73 (8.12)	4.27 (2.60)
Controls				
No Prior Knowledge	111.23 (13.76)	13.27 (6.56)	6.87 (2.90)	2.33 (1.84)
Prior Knowledge	111.30 (19.52)	12.6 (.39)	6.13 (4.31)	2.27 (1.53)

Table 4

Mean Number of Oral Reading Substitution Errors (Standard Deviations in Parentheses) as a Function of Error Type, Topic Knowledge and Reading Ability

	Error Type	No Prior Knowledge	Prior Knowledge
Poor Readers	G-S-	3.20 (4.33)	2.27 (1.98)
	G-S+	1.40 (1.40)	2.20 (3.23)
	G+S-	10.93 (8.08)	6.00 (3.30)
	G+S+	4.20 (2.27)	4.27 (2.25)
Good Readers	G-S-	.67 (.82)	.40 (.63)
	G-S+	1.07 (1.39)	1.33 (1.35)
	G+S-	2.87 (1.77)	3.00 (3.32)
	G+S+	2.27 (1.16)	1.40 (.91)

Table 5

Means (Standard Deviations in Parentheses) for Oral Reading Errors and for Comprehension and the Correlation of Errors with Comprehension

	Mean	<i>Correlation with Comprehension</i>
Comprehension	17.78 (6.08)	
Total Errors	18.03 (11.04)	-.43 **
G-S-	1.63 (2.64)	-.31 *
G-S+	1.50 (2.00)	.15
G+S-	5.70 (.69)	-.59 **
G+S+	3.03 (2.12)	-.29 *

*
 $p < .05$.

**
 $p < .01$