CHAPTER 17

Do Readers Use Phonological Codes to Activate Word Meanings? Evidence from Eye Movements

Meredyth Daneman and Eyal M. Reingold

University of Toronto

Abstract

Do readers use phonological codes to activate word meanings during natural silent reading? Many researchers have been arguing for the early involvement of phonology in activating word meanings. However, the evidence has depended largely on tasks that required readers to make some response in addition to reading (e.g., lexical decisions, proofreading). For example, when asked to proofread text for errors, readers have been less likely to detect homophonic errors (e.g., *He wore blew jeans*) than nonhomophonic errors (e.g., *He wore blow jeans*), a finding that has been taken as evidence that phonology is used to access word meanings. Although such findings are consistent with phonologically mediated lexical access, they do not provide conclusive evidence. There is always the concern that the secondary task may have changed the nature of the reading process itself, making the results less generalizable to natural reading. Furthermore, because the phonological involvement has been inferred from a slow secondary response rather than from reading itself, the findings are equivocal with respect to the time course of the phonological activation. Our research avoided these pitfalls. Readers were not asked to proofread for errors. They simply read for comprehension, and their eye movements were monitored for evidence of spontaneous disruptions when encountering homophonic errors (e.g., *blew*) versus orthographic control errors (e.g., *blow*). The evidence strongly suggests that readers use orthographic codes to access word meanings, with phonological codes playing a more restricted role in the accessing of meanings for low-frequency words. In this chapter we describe our research and we respond to a recent article by Rayner, Pollatsek and Binder (1998) which challenges our conclusions.
In 1993, Daneman and Reingold claimed to have provided evidence from eye movement data that skilled adult readers use orthographic codes rather than phonological codes to activate word meanings (see also Daneman, Reingold and Davidson, 1995). Evidence for the early involvement of phonology in activating word meanings had depended largely on paradigms that required readers to make some response in addition to reading (e.g., making semantic judgments about words and sentences, proofreading text for inconsistent words) and so there was always the concern that the secondary task requirement had changed the nature of the reading process itself. The significance of Daneman and Reingold's claims is that they were based on a task that demanded nothing other than normal reading for comprehension. However, a recent study by Rayner et al. (1998) has challenged the Daneman and Reingold (1993) and Daneman et al. (1995) claims. Using a variant of the Daneman and Reingold eye-movement monitoring task, Rayner et al. (1998) interpreted their data as evidence for the early and dominant involvement of phonology in accessing word meanings. In this chapter we critically evaluate the position of Rayner et al. (1998) in the context of new data that replicate and extend our earlier findings.

The background

Many proponents of phonologically mediated lexical access have drawn on evidence from tasks showing homophone confusion effects. One such task is lexical decision in which subjects judge whether a given letter string is a word. A typical finding is that subjects take more time to reject pseudohomophone foils, such as brane, than control foils, such as brene (Coltheart, Davelaar, Jonasson and Besner, 1977; Rubenstein, Lewis and Rubenstein, 1971). A common explanation for the effect is that the pseudohomophone brane activates the phonological representation /brein/, which in turn activates the lexical entry for the word brain. The activation of a lexical entry makes brane more difficult to classify as a nonword. Another task showing homophone confusions to single words is the semantic categorization task of Van Orden (1987). In this task, subjects are presented a category name (e.g., type of food) followed by a target word (e.g., meet, melt), and their task is to decide whether or not the target word is a member of the category. The typical finding is that subjects make more false categorization responses to meet which sounds like the genuine category member, meat, than they do to melt which is orthographically similar to meat but does not sound like it (Van Orden, 1987; Van Orden, Johnston

1 Although see also Pollatsek, Lesch, Morris and Rayner (1992) and Rayner, Sereno, Lesch and Pollatsek (1995) who use homophone preview/priming paradigms to argue for the early involvement of phonological codes in word identification and lexical access. The homophone preview paradigm will be discussed later.
and Hale, 1988). Homophone confusions have also been demonstrated in tasks that require readers to make semantic decisions about an entire sentence rather than an isolated word. The typical finding is that readers make more false-positive semantic acceptability decisions to incorrect sentences that sound correct (e.g., She has blond hare) than to incorrect sentences that do not sound correct (e.g., She has blond harm) (Coltheart, Avons and Trolley, 1990; Coltheart, Laxon, Rickard and Elton, 1988; Treiman, Freyd and Baron, 1983). And finally, homophonic confusions are not restricted to tasks that involve decisions to lists of unrelated words or simple sentences; they have been demonstrated for realistic, everyday prose as well (Daneman and Stainton, 1991; see also Van Orden, 1991). Daneman and Stainton (1991) had subjects proofread a lengthy and complex prose passage containing inconsistent words that were or were not homophones of consistent ones. Phonology was implicated by the finding that readers were less likely to detect homophonic errors (e.g., One night a week they would meat . . . ) than nonhomophonic orthographic control errors (e.g., One night a week they would moat . . . ). The finding that homophonic words are commonly misinterpreted as their sound-alike mates has been taken as evidence for phonology playing an early and dominant role in accessing word meanings.

Although demonstrations of homophone confusions are consistent with phonologically mediated lexical access, they do not provide conclusive evidence for early phonological influences. Because the demonstrations of homophone confusions have all relied on paradigms that required readers to make some response in addition to reading (e.g., lexical decisions, semantic decisions, proofreading), there is always the concern that the secondary task may have changed the nature of the reading process itself, making the results less generalizable to natural reading situations. Furthermore, because the phonological involvement has been inferred from the often slow secondary response rather than from reading itself, the findings are equivocal with respect to the time course of the phonological activation. Phonological interference could have resulted from the delayed effect of phonological codes that were not involved in the initial activation of word meanings. Consequently, it would be preferable to look for evidence of phonological involvement from an on-line reading measure (see also Van Orden, 1991).

The Daneman and Reingold studies

Daneman and Reingold (1993)

In Daneman and Reingold (1993), we looked for evidence of phonologically mediated lexical access in the most natural or typical of reading situations possible, reading for comprehension and enjoyment. Participants read the same 1100-word
text (called Russell Wood) used in the Daneman and Stainton (1991) proofreading study. However, they were not told that inconsistent words had been introduced into the text, nor were they given explicit instructions to proofread for the inconsistent words as they read; they were simply asked to read for comprehension and their eye fixations were recorded to examine whether phonological involvement was spontaneously revealed in the moment-to-moment computational processes of regular reading. Previous research has shown that readers gaze longer at words that are inconsistent with previously read information (Carpenter and Daneman, 1981; Frazier and Rayner, 1982; Just and Carpenter, 1980) and frequently make regressive fixations as they attempt to resolve the inconsistencies (Carpenter and Daneman, 1981). Thus, any additional time spent fixating an inconsistent word (e.g., … One night a week they would meat … or … One night a week they would moat …) relative to the consistent one (e.g., … One night a week they would meet …) could be attributed to the processes involved in inconsistency detection and recovery. If phonological codes are used to activate word meanings as the Daneman and Stainton (1991) proofreading data would have us believe, then readers should frequently fail to detect homophone errors, showing no additional processing time when first encountering meat relative to processing time spent when first encountering the contextually appropriate meet. On the other hand, if orthographic codes are used to activate word meanings, then readers should have no difficulty detecting homophone errors, showing as much disruption when initially encountering the homophone error meat as the orthographic control error moat.

The Daneman and Reingold (1993) eye fixation data suggested that readers use orthographic codes rather than phonological codes to activate the meanings of words during natural silent reading. Contrary to the results of Daneman and Stainton (1991) from their secondary proofreading task, Daneman and Reingold (1993) found no evidence that homophony interfered with the initial detection of homophone errors. In fact, first pass gaze durations showed that not only did the homophone errors cause more disruptions than the contextually correct homophones when first encountered (Experiments 1 and 2), but these homophone errors were as disruptive as the orthographic control errors (Experiment 2), suggesting that they were detected as easily. This lack of phonological interference in the early detection of homophonic errors was taken as evidence against those models that assume that phonological sources of activation invariably precede lexical access (Daneman and Stainton, 1991; Van Orden, 1987). Instead, the results suggested that readers bypass phonology, using the orthographic representations for meat and moat as a direct route to their contextually inconsistent meanings, ‘edible flesh’ and ‘trench’, respectively. The Daneman and Reingold (1993) eye-fixation data also revealed that the orthographic similarity between the homophone imposter and its mate did not affect the detectability of homophone errors; readers were as likely
to detect orthographically similar same-length homophone impostors (e.g. meat posing as meet) as they were to detect less orthographically similar different-length homophone impostors (e.g., wade posing as weighed).

Whereas the initial detection data of Daneman and Reingold (1993) did not provide evidence for the early engagement of phonological processes in activating a word’s meaning, the post-detection data provided some evidence for the delayed involvement of phonology in the error recovery processes. We found that homophony facilitated the error recovery processes that are initiated after an inconsistency is detected. The regressive eye fixations showed that readers spent less time rereading phrases containing homophonic errors (e.g., One night a week they would meat . . . ) relative to phrases containing the orthographically matched nonhomophonic errors (e.g., One night a week they would moat . . . ), presumably because for those cases in which readers had successfully detected the inconsistent impostor word (meat), they could exploit the shared phonology (e.g., /mit/) as a route to recovering the correct alternative (meet). These data suggested that phonology has its influence after lexical access.

Daneman, Reingold and Davidson (1995)

In Daneman et al. (1995), we provided evidence for the reliability of the Daneman and Reingold (1993) findings. In Experiment 1, participants read one of two new texts, and their eye movements were monitored for evidence of spontaneous disruptions when encountering homophonic and nonhomophonic error words. In one text (called Black Queen), all the homophone errors were low-frequency words relative to their contextually correct mates (e.g., meat substituted for meet; rein substituted for rain). In the other text (called Desjardins) all the homophone errors were high-frequency words relative to their contextually correct mates (e.g., meet substituted for meat; rain substituted for rein). As in the Daneman and Reingold (1993) study, we found that homophony did not interfere with the initial detection of homophone errors. First pass gaze durations showed that not only did homophone errors cause more disruption than correct homophones (Experiments 1A and 1B), but these homophone errors were as disruptive as the orthographic control errors (Experiment 1B), suggesting that orthography rather than phonology is used to activate word meanings. The only possible hint of an early influence of phonology came in the Black Queen text which contained the lower-frequency homophone errors; although the data showed no statistical support for any effect of relative word frequency on the initial detection of homophone errors, there was a hint (albeit a nonsignificant one) that the lower-frequency homophone errors were less disruptive than their orthographic controls (Experiment 1B), an issue that will be addressed in the present study. There was, of course, a delayed involvement of phonology in the error recovery process: Experiment 1 replicated the Daneman and Reingold (1993)
finding that homophony facilitated error recovery by showing that readers spent less time in regressive fixations to homophone errors than to orthographic control errors.

The Daneman et al. (1995) study also exposed the dangers of using a secondary proofreading response to make inferences about the time course of phonological activation during reading. In Experiment 2, participants read the original Daneman and Reingold (1993) Russell Wood text, but this time they were given explicit instructions to proofread for inconsistent words as they read. By collecting eye movement data in conjunction with an explicit proofreading task, we showed that overt proofreading responses are unreliable indices of error detection because even when readers failed to make an overt error detection response, their eye fixations revealed that they were disrupted by an error. Taken alone, the proofreading data would have led to the conclusion that phonology is used to activate word meanings because they showed that readers were less likely to make an overt detection response in the presence of homophonic errors than in the presence of orthographic control errors. However, the eye fixations revealed the same degree of disruption when readers first encountered homophonic errors as when they encountered orthographic control errors, leading to the conclusion that orthography rather than phonology is used to activate word meanings.

Our model

We believe that the on-line eye fixation data from our inconsistency detection paradigm provide compelling evidence concerning the time course of phonological activation during natural silent reading. With the possible exception of low-frequency words (Daneman et al., 1995, Experiment 1B), we take our data to be inconsistent with a theory of lexical access in which phonological codes play an early and/or dominant role (e.g. Daneman and Stainton, 1991; Inhoff and Topolski, 1994; Perfetti, Bell and Delaney, 1988; Pollasek et al., 1992; Rayner et al., 1995; Van Orden, 1987; Van Orden, Pennington and Stone, 1990). Rather, the data are consistent with a theory in which orthographic codes play the dominant role in activating word meanings (see also Coltheart, 1978; McCusker, Hillinger and Bias, 1981; McCutchen and Perfetti, 1982). Our model is depicted in Fig. 1. The figure highlights the main pathways of activation involved in accessing words meanings and detecting semantic inconsistencies in the kinds of low-constraint texts used in Daneman and Reingold (1993) and Daneman et al. (1995). Because most of our target words were not predictable from the preceding context, we believe that bottom-up sources of activation played the major role; later we will describe how the model can be extended to account for top-down sources of activation in the highly predictable contexts used by Rayner et al. (1998). As seen in Fig. 1, the predominant route from print to meaning is via the orthographic representation;
the printed stimulus *meat* activates the orthographic representation *[meat]* which in turn activates the semantic interpretation ‘edible flesh.’ The error word *meat* will be detected if the reader fails to integrate ‘edible flesh’ with the preceding context ‘*One night a week they would . . .*’ Although orthography plays the dominant role in lexical access, our model leaves open the possibility of phonologically mediated lexical access. As seen in Fig. 1, the orthographic representation *[meat]* also activates the phonological representation */mit/* which can activate the semantic representations ‘flesh’ and ‘gather’; if the latter meaning is selected, semantic integration will succeed and the error word *meat* will go undetected. However, word meaning is less likely to be activated by the phonological representation than by the orthographic representation because activation from phonology is considerably weaker and/or delayed.

**The Rayner, Pollatsek, and Binder critique and our response**

Using a variant of our eye-movement inconsistency detection paradigm, Rayner et al. (1998) have reported a pattern of data that is inconsistent with our pattern in several ways, and they have interpreted their data as evidence for the early involvement of phonology in accessing word meanings. We will first describe their
model of lexical access and inconsistency detection and the predictions derived from the model. Then we will describe their main findings and critically evaluate their interpretation of these findings. Finally we will present new data from a norming study and a replication study to support our position.

Rayner, Pollatsek, and Binder’s model

Rayner et al. (1998) appeal to the verification model of lexical access (see also Van Orden, 1987) to make their predictions and to explain their results. The verification model is depicted in Fig. 2. According to the verification model, the phonological representation is activated immediately and used exclusively to gain access to a word’s semantic representation, with the orthographic representation playing a post-activation verification role. As seen in Fig. 2, when the reader encounters the printed word meat, its phonological representation /mit/ is immediately activated from the orthographic features, and it is the phonological representation that activates candidate lexical entries such as ‘flesh’ and ‘gather’. However, before an activated lexical entry can be selected, it must pass a verification test or spelling check. The orthographic representation associated with the most active candidate

2 Rayner et al. (1998) do not provide as detailed an account of the verification model in their paper as we provide here. However, we believe that it is important to articulate their model and its predictions clearly if we are to critically evaluate their study, and we believe that our rendition of the model is entirely consistent with the model that they appear to be using to make their predictions and to explain their results. Our rendition is also consistent with the verification model described by Van Orden (1987).
phonology and meaning activation

is retrieved from memory and compared with the orthographic representation of the word being read. If a match occurs, the lexical entry is selected; if not, the verification process is repeated on the next most active candidate lexical entry. According to this model, the printed homophone meat may sometimes be mistaken for meet if the false candidate meet were made available to the orthographic verification procedure and the mismatch in spelling slipped by undetected. According to Rayner et al. (1998), two factors influence the likelihood of detecting a spelling error when the false candidate of a homophone undergoes orthographic verification. The first factor is the orthographic similarity of the homophone to its mate; the greater the orthographic similarity of the homophones, the harder it should be to detect the mismatch (e.g., the error word meat would be harder to detect than the error word chute because meat is more orthographically similar to meet than chute is to shoot). The second factor that influences orthographic verification is the predictability of the correct target word from the prior context; if the prior context strongly predicts a target word, top-down or expectancy-driven reading may lead to a relaxation or abortion of the verification process, allowing a mismatch to go undetected. Rayner et al. (1998) tested these predictions by manipulating orthographic similarity and contextual constraint. Orthographically similar homophone pairs shared the same first two letters (e.g., meat—meet) or the same first letter (e.g., rain—rein); orthographically dissimilar pairs did not share the same first letter (e.g., chute—shoot; right—write). In the high-constraint passages, the correct homophone target was highly predictable from the preceding context (mean = 0.86); in the low-constraint passages, the correct homophone target was not predictable (mean = 0.03). Whereas Rayner et al. (1998) argue that their data “appear to be compatible with” (p. 492) the verification model, we will argue that their high-constraint data are compatible but their low-constraint data are not.

Before describing the results of Rayner et al. (1998), it is worth highlighting an important difference between their model and ours with respect to the locus at which an error/inconsistency is detected. For the most part, Rayner et al. (1998) appear to be assigning error detection to the orthographic verification process rather than to the processes that attempt to integrate an activated word meaning with the preceding context as we do in our model. This locus is logically possible in the case of detecting homophonic errors (e.g., meat posing for meet); the phonology /mit/ activates the false candidate meet which is made available to the orthographic verification procedure and the mismatch in spelling between the retrieved orthography [meet] and the actual orthography of the printed word [meat] is successfully detected. However, an orthographic control error (such as moat posing for meet) cannot be detected by orthographic verification; in this case the phonological code /mot/ would have activated the semantic representation ‘trench’ and because its associated orthographic representation [moat] matches
the actual printed orthography moat, there is no mismatch to be detected. The only way to detect that the word moat is erroneous is to notice that the activated meaning ‘trench’ is inconsistent with the preceding context. Interestingly, Rayner et al. (1998) do not seem to have appreciated that their model calls for two different mechanisms for detecting homophonic versus nonhomophonic errors, or if they have, they do not make it explicit in their paper. A by-product of their emphasis on orthographic verification rather than semantic integration as the locus of error detection is that they favour first fixation durations as the most sensitive measures of early processing. Because we emphasize semantic integration and inconsistency detection, we favour gaze duration (the sum of all fixations on a word prior to moving to another word) because gaze duration has been shown to better reflect both meaning retrieval and semantic integration than does first fixation duration (Carpenter and Daneman, 1981; Inhoff, 1984; Just and Carpenter, 1987).

Rayner, Pollatsek, and Binder’s data and our response

Strongest support for phonologically mediated lexical access came from Rayner et al.’s (1998) high-constraint texts. As the verification model (Fig. 2) predicts, homophony interfered with the initial detection of homophone errors particularly if they were orthographically similar to their mates. In the orthographically similar condition (e.g., meat, meet), first fixation durations showed that homophone errors behaved the same as contextually correct homophones (Experiments 1, 2, 3), and both were less disruptive than were the orthographic control errors (Experiments 2 and 3). In the orthographically dissimilar condition (e.g., chute, shoot), first fixation durations showed that homophone errors were more disruptive than contextually correct homophones, but less disruptive than orthographic errors. These data are consistent with Rayner et al.’s (1998) verification model (Fig. 2) because the model predicts that homophone errors should be especially difficult to detect if the contextually correct homophone is highly predictable and similar in spelling to the homophone error.

It is when we come to Rayner et al.’s (1998) data for low-constraint texts (the texts more like ours) that the verification model runs into trouble. When the correct target word was not predictable, fixation durations showed a pattern similar to that for orthographically dissimilar homophones in high-constraint text. Regardless of the orthographic similarity condition, first fixation durations showed that homophone errors were more disruptive than contextually correct homophones but less disruptive than orthographic control errors (Experiment 3). The finding of a difference in fixation durations for incorrect homophones and correct homophones is consistent with Rayner et al.’s (1998) model which predicts that homophone errors should be more easily detected in low-constraint contexts than in high-constraint
contexts because the orthographic verification check is more efficient when the target word is not predictable. However, the model cannot account for the finding that orthographic similarity between the homophone error and its contextually correct mate did not affect the detectability of homophone errors.

Rayner et al. (1998) recognize that the lack of an orthographic similarity effect is a problem for a verification model such as theirs. They attempt to accommodate the finding within the framework of their model by arguing that “under low constraint conditions, the bottom-up verification process works more efficiently with no top-down expectation process to abort it, and thus it operates rapidly enough to affect first fixation durations even when the homophones are orthographically similar.” (p. 492). This explanation is of course totally unsatisfactory. The only way it could work is if the orthographic verification process were so efficient that it succeeded in detecting all homophone errors, and then there should have been no difference in the degree of disruption for homophone errors versus orthographic control errors. Although that was the pattern we found in our experiments, it was not the pattern found in the experiment of Rayner et al. (1998); first fixation durations for homophone errors were shorter than for orthographic errors, suggesting that some of the homophone errors had slipped by the orthographic verification process undetected. As long as some homophone errors were not being detected, there should have been a higher probability that they were the orthographically similar ones. It would appear that Rayner et al. (1998) are not entirely convinced by their own argument because immediately after making it they say “Although this is possible, it seems strange that there appeared to be no modulation of the size of the effect as a function of orthographic similarity” (p. 492).

The next attempt of Rayner et al. (1998) to deal with the lack of an orthographic similarity effect in low-constraint texts is even more puzzling. Instead of abandoning their orthographic verification model, they seem to be proposing a second model for lexical access and error detection in low-constraint conditions. Issues of parsimony aside, we will argue that this model cannot co-exist with their orthographic verification model because the two are logically incompatible. Rayner et al. (1998) argue that the low-constraint condition “may not be particularly diagnostic because the initial detection of anomaly may be largely based on the degree of anomaly of the actual words to the preceding context rather than to any orthographic or phonological similarity to the correct homophone” (p. 492). In other words, because they could not find an effect of orthographic similarity for low-constraint texts, they want to argue that phonological identity and orthographic similarity are important factors in error detection in high-constraint texts, whereas semantic inconsistency is the locus of error detection in low-constraint texts. Of course, we have argued all along that semantic inconsistency detection is the mechanism for detecting all
error words in this paradigm. However, Rayner et al. (1998) still want to maintain the phonological route to meaning activation and so they propose that homophony has two opposing effects on error detection in low-constraint texts: “(a) activation of a semantic code that is a good continuation, which delays (or possibly aborts) judgments that the word does not make sense in context, but (b) activation of a semantic code that is a ‘model’ of good continuation that makes detection of the meaning of the wrong homophone easier and faster” (Rayner et al., 1998, p. 495). We have no problems with mechanism (a); after all, the paradigm was designed to investigate the possibility that homophony might lead to the activation of the false but contextually consistent meaning, making the homophone error word difficult to detect. However, we are puzzled by their proposal of mechanism (b), the idea that phonological activation of both homophone meanings results in homophone errors being more inconsistent and hence easier to detect. This proposal is puzzling on empirical grounds because it predicts that homophone errors will be more conspicuous than orthographic errors, a pattern neither they nor we have found. Of course, by suggesting that “under certain circumstances, mechanisms (a) and (b) may be equally potent” (p. 495) Rayner et al. (1998) give themselves a way to account for the lack of difference between homophone errors and control errors we obtained in our experiments; however, trade-off explanations of null effects are complex and not particularly satisfying. In any case, there is reason to dismiss Rayner et al.’s (1998) proposal for low-constraint texts on logical grounds because it is incompatible with their verification model for high-constraint texts. As illustrated earlier (see Fig. 2), according to the verification model, context only has its effect at the orthographic verification phase which is after phonology has activated meaning. Consequently, mechanisms (a) and (b) should be equally possible in the high-constraint condition as in the low-constraint condition. And yet Rayner et al. (1998) invoke these two mechanisms for detecting semantic inconsistencies in conditions of low contextual constraint only, and we find this troublesome.

In the final analysis, Rayner et al. (1998) seem willing to ignore the low-constraint data and to base their strong claims for the phonological activation of word meaning on the data from high-constraint texts alone. We see two problems with this decision. First, words that are as highly predictable as Rayner et al.’s (1998) high-constraint words probably represent a small proportion of the content words that occur in natural complex prose; consequently, we think it unwise to assume

---

3 And Rayner et al.’s (1998) own data demonstrated that inconsistency detection is the mechanism for detecting error words in this paradigm. Remember that their first attempt to investigate error detection in low-constraint texts (Experiment 2) was a failure; there was no difference in first pass times for correct homophones, homophone errors, and orthographic control errors, precisely because the error words were consistent with the prior context when first encountered and only became inconsistent later in the sentence.
that the way readers process highly predictable words will generalize to the way they process most content words in natural texts. Second, even though Rayner et al.'s (1998) high-constraint data are consistent with phonologically mediated access and post-access verification, there is an equally plausible account of the findings of Rayner et al. that does not involve the use of phonological codes to activate meaning. This account is depicted in Figs. 3 and 4, and is an extension of our model for low-constraint texts depicted in Fig. 1.

Our alternative model for high-constraint text

Whereas our model for low-constraint texts emphasized bottom-up sources of activation (from print to orthography to meaning), the model in Figs. 3 and 4 includes top-down influences as well (see also McClelland, 1987). Fig. 3 illustrates the model for the homophone error meat; Fig. 4 illustrates the model for the orthographic control error moat. As seen in the top panel of Fig. 3, if the context for meet is highly constrained, it will begin to activate the semantic representation ‘gather’ and the semantic representation will itself activate the corresponding phonological representation /mit/ and orthographic representation [meet] before the reader has even fixated the printed word meat. When the reader does fixate the printed word (Fig. 3, bottom panel), the phonological representation /mit/ and semantic representation ‘gather’ will receive further activation. The combination of top-down activation by the preceding context and bottom-up activation by the printed word may be enough to select the phonological representation /mit/ and semantic representation ‘gather,’ allowing the homophone error meat to go unnoticed (see also Jared and Seidenberg, 1991). According to this model, it is meaning that first activates phonology and not phonology that first activates meaning. Fig. 4 illustrates why it is easier to detect the orthographic control error moat than the homophone error meat in highly constrained text. The first phase of activation before the printed word is encountered is of course identical for moat as for meat, with the semantic representation ‘gather’ and the phonological representation /mit/ receiving activation before the error word is encountered (Fig. 4; top panel). However, once the printed word moat is fixated, the contribution of bottom-up activation will be more diffuse because there will be bottom-up activation of an additional phonological representation /mot/ (Fig. 4; bottom panel). Consequently, there will be less convergence of activation on /mit/ and ‘gather’ than there was in the case of the printed word meat, and hence less of a probability that the orthographic control error moat goes undetected. We think that the model depicted in Figs. 3 and 4 provides a plausible account of the homophone interference effect found in Rayner et al.'s high-constraint texts without invoking exclusive bottom-up activation of meaning by phonology. The model is consistent with our model for processing words in the more natural low-constraint texts.
Before fixating 'meat'

While fixating 'meat'

Fig. 3. A model of meaning activation in highly predictable texts, as it applies to the homophone error meat. Bidirectional arrows depict activation in two directions. Dotted lines represent inhibitory processes.

Rayner et al.’s concerns about our research

So far we have focused on Rayner et al.’s (1998) interpretation of their findings and our concerns about their interpretation. Next we will briefly mention two issues they raise about our experiments and our response to these issues. Rayner et al. (1998) raise the issue of how to characterize the degree of contextual constraint in our experimental passages. Given that degree of contextual constraint had a large effect on the pattern of results in their study, it is difficult to compare the results from the two laboratories without having some idea about the contextual constraint levels in our texts. As Rayner et al. (1998) point out, it would appear that many of our correct
homophones were not predictable from the preceding context, making our natural texts more like their carefully constructed low-constraint passages. However, as they also point out, some of our correct homophones were highly predictable from the preceding context, and so a direct comparison across laboratories becomes tricky. Rayner et al. (1998) also question whether our homophone errors and our orthographic control errors are equally inconsistent with the preceding context. Whereas Rayner et al. (1998) tried to equate degree of inconsistency by collecting formal ratings for the two kinds of errors in their low-constraint passages, our assessment of degree of semantic inconsistency for our two kinds of errors was an informal one. As Rayner et al. (1998) point out, “the immediacy of detecting that
the actual word does not fit (and hence the degree to which it would affect early measures of processing such as first fixation duration and gaze duration) would depend on the degree to which it was anomalous with the prior context” (p. 495). If our homophone errors were more inconsistent than our control errors, this could have accounted for why they were detected easily and immediately.

The norming study

Our first step was to conduct a formal evaluation of these issues in a norming study. We used procedures similar to those used by Rayner et al. (1998) to obtain contextual constraint estimates and semantic consistency ratings for the three experimental texts we have used: the Russell Wood text used in Daneman and Reingold (1993) and in Experiment 2 of Daneman et al. (1995), and the Black Queen and Desjardins texts used in Experiments 1A and 1B of Daneman et al. (1995). For each of the three texts, contextual constraint scores were obtained by having three separate groups of 20 University of Toronto undergraduate students perform a modified cloze task in which they were presented the text up to but not including the target word, and they were asked to guess what the next word would be. For each of the three texts, semantic consistency scores were obtained by presenting three new groups of 30 undergraduate students with the text up to and including the target word and their task was to rate on a scale of 1 to 7 how well the target word fit into the preceding context. (A rating of 1 meant that the word was totally inconsistent with the context, whereas a rating of 7 meant that the word was consistent and a good continuation of the passage.) For the semantic consistency ratings for each of the three texts, each rater provided ratings for one third of the correct target words, one third of the homophone errors, and one third of the orthographic control errors. The contextual constraint scores (expressed as the proportion of completions with the correct homophone) and the semantic consistency ratings (1–7) for our three texts are provided in Table 1.

There are several points to note from the norming data in Table 1. First, the contextual constraint scores show that all three of our texts can be considered low-constraint texts when compared to the passages used by Rayner et al. (1998). However, it is also interesting to note that the average predictability of correct homophone target words in the Black Queen text was higher (0.28) than in the Russell Wood (0.04) and Desjardins (0.08) texts, and there was also a wider range of predictability scores for correct target words in the Black Queen text (scores ranged from 0.00 to 1.00, SD = 0.33). The Black Queen text was the text that showed a hint of a phonological effect in Daneman et al. (1995, Experiment 1B). We had attributed the weak phonological effect to the fact that the homophone errors in the Black Queen text were lower-frequency words than the homophone errors in the Desjardins text. However, it is possible that the weak effect was due to the
Table 1

Contextual constraint scores (in a proportion) and semantic consistency ratings (1–7) from norming study

<table>
<thead>
<tr>
<th>Text</th>
<th>Contextual consistency ratings</th>
<th>Semantic consistency ratings</th>
<th>Correct word</th>
<th>Homophone error</th>
<th>Orthographic error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>Russell Wood</td>
<td>0.04  0.06</td>
<td>6.74  0.60</td>
<td>1.53  0.85</td>
<td>1.34  0.52</td>
<td></td>
</tr>
<tr>
<td>Desjardins</td>
<td>0.08  0.13</td>
<td>6.82  0.27</td>
<td>1.55  0.23</td>
<td>1.36  0.27</td>
<td></td>
</tr>
<tr>
<td>Black Queen</td>
<td>0.28  0.33</td>
<td>6.78  0.25</td>
<td>1.51  0.26</td>
<td>1.10  0.10</td>
<td></td>
</tr>
</tbody>
</table>

The presence of some highly predictable target words in the Black Queen text. This is an issue we explore in the replication study reported next. And finally, the semantic consistency scores show that our results were unlikely to have been contaminated by differences in semantic inconsistency between homophone errors and orthographic control errors. As seen in Table 1, our homophone errors were not judged to be more inconsistent with the preceding context than were our orthographic control errors; if anything there were judged to be a little less inconsistent. Consequently, it was not a failure to equate for semantic inconsistency that produced our pattern of results in which homophone errors were as easy to detect as orthographic control errors. What the norms in Table 1 show is that at least two of our texts, Russell Wood and Desjardins, are comparable to the Rayner et al. (1998) low-constraint passages in terms of degree of contextual constraint and degree of semantic inconsistency. This means that neither contextual constraint nor semantic inconsistency can explain why we found no evidence for early phonological involvement in accessing word meanings in our low-constraint texts whereas they found some evidence for early phonological involvement.

The replication study

In addition to collecting norms, we also ran another experiment in which we attempted a replication of the experiment with the Desjardins and Black Queen texts (Experiment 1B, Daneman et al., 1995). In the original Daneman et al. study there were only 15 participants who read the Desjardins text and 15 participants who read the Black Queen text. By collecting eye movement data from a further 18 participants on the Black Queen text, the replication allowed us to see whether the previous hint of a phonological effect for the Black Queen text was a real effect and also whether differences in the probability of detecting the homophone errors in the Black Queen text were correlated with differences in the degree to which their correct homo-
phone mates were constrained by the preceding context. In addition, by collecting eye movement data from a further 18 participants on the Desjardins text, the replication experiment allowed us to see whether we could once again find the pattern of homophone errors being detected as easily as control errors in the low-constraint Desjardins text, the pattern that has been challenged by Rayner et al. (1998).

The materials for our replication experiment were identical to those for Experiment 1B in Daneman et al. (1995), so we will describe them only briefly here. The experimental manipulation involved 30 homophone word pairs with asymmetric word frequencies (e.g., hair–hare; meet–meat, peer–pier, and rain–rein). The mean Kučeřa and Francis (1967) frequency count for the higher-frequency member of the pair (e.g., hair, meet, peer, and rain) was 336 occurrences per million (Mdn = 162, SD = 520); the mean frequency count for the lower frequency member of the pair (e.g., hare, meat, pier, and rein) was 43 occurrences per million (Mdn = 14, SD = 85). Although the members of a homophone pair differed in frequency, they were orthographically similar to one another in that each member of the pair was spelled with the same initial letter and was the same length as the other. The orthographic control word for each homophone pair shared the same consonant sounds as the correct word and its homophone mate but differed only in the vowel sound (e.g., moat for meet–meat and ruin for rain–rein). Because the controls were matched for orthographic similarity rather than for frequency, we used the same orthographic control for both members of a homophone pair. The 30 lower-frequency homophones all appeared in (or were edited into) the short story The Desjardins by Duncan Campbell Scott (1988); the 30 higher-frequency homophones all appeared in (or were edited into) the short story The Black Queen by Barry Callaghan (1988). Participants read a version of the Desjardins or Black Queen text in which 10 of the target words appeared in their correct form, 10 as homophone errors, and 10 as orthographic control errors. Counterbalancing was accomplished by creating three versions of each text; a target word appeared in a different form (correct word, homophone error, and orthographic error) in each version. In the original Daneman et al. (1995) study, 15 participants read the Desjardins text and 15 participants read the Black Queen text. In the replication study, we had 18 participants read Desjardins and another 18 read the Black Queen. Participants in both studies were students at the University of Toronto.

The procedure for the replication study was similar to that for the original Experiment 1B of Daneman et al. (1995). Participants were told that they would

---

4 Note that our frequency manipulation involved a manipulation of the relative word frequency of the text word and the error replacement rather than a manipulation of their absolute frequency. Even our lower-frequency homophones were sufficiently common that readers would be likely to know their meaning and spelling (e.g., meet–meat and rain–rein, but not pigeon–pidgin or bridal–bridle as in Jared and Seidenberg, 1991).
be presented a short story on successive screens of a computer monitor. They were instructed to read the story silently at their own pace, making sure they understood it well enough to answer questions about its content later. The text was presented in black (brightness = 4 cd/m²) on a white background (brightness = 68 cd/m²). Proportional spaced fonts were used with an average of 2.2 characters per degree of visual angle and an average of 10 lines per screen. Displays were generated using an S3 VGA card and a 17” ViewSonic 17PS monitor. At the 60-cm viewing distance the display subtended a visual angle of 30° horizontally and 22.5° vertically. When participants finished reading a screen they pressed a button to proceed to the next screen. In the original Daneman et al. (1995) study, eye fixations were recorded by an Iscan (Model RK-416) video-based eye tracking system. In the present replication study, the eye tracker employed was the SR Research Ltd. EyeLink system. This system has high spatial resolution (0.005°), and a sampling rate of 250 Hz (4 ms temporal resolution). The EyeLink headband has three cameras, allowing the simultaneous tracking of both eyes and of head position for head-motion compensation. By default, only the participant’s dominant eye was tracked in our study. The EyeLink system uses an Ethernet link between the eye tracker and display computers for real-time saccade and gaze position data transfer. The system also performs saccade and blink detection on-line. In the present study, the configurable acceleration and velocity thresholds were set to detect saccades of 0.5° or greater. A 9-point calibration was performed at the start of the experiment, followed by a 9-point calibration accuracy test. Calibration was repeated if any point was in error by more than 1°, or if the average error for all points was greater than 0.5°.

The eye fixation results of the replication study are presented in Table 2. As discussed earlier, our preferred measure of initial meaning retrieval and semantic integration/inconsistency detection is gaze duration on the target word (the sum of all fixations on the target word prior to moving to another word); however, first fixation duration on the target word (the duration of the first fixation on the target word independent of the number of fixations on that word) and single fixation duration on the target word (the fixation duration when only one fixation is made on the word) are also presented in Table 2 because these are measures analyzed in Rayner et al. (1998). In addition to these three measures which are sensitive to early processes, total time on the target word (the sum of all fixations on the target word, including regressions) is presented as a measure that is sensitive to the later processes of error recovery as well. The corresponding data from Experiment 1B of Daneman et al. (1995) are also provided in Table 2 for comparison purposes, as are the combined data for the original and replication studies.

---

5 Note that single fixation durations were not reported in Daneman et al. (1995).
Table 2

Mean reading times (in milliseconds) in the replication experiment, Daneman et al.’s (1995) Experiment 1B, and averaged across the two experiments

<table>
<thead>
<tr>
<th></th>
<th>Replication</th>
<th>Daneman et al. (1995)</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>Hom. error</td>
<td>Orthog. error</td>
</tr>
<tr>
<td></td>
<td>word (18)</td>
<td>(18)</td>
<td>(18)</td>
</tr>
<tr>
<td>Desjardins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaze duration</td>
<td>280</td>
<td>314</td>
<td>319</td>
</tr>
<tr>
<td>First fixation duration</td>
<td>232</td>
<td>240</td>
<td>247</td>
</tr>
<tr>
<td>Single fixation duration</td>
<td>235</td>
<td>253</td>
<td>254</td>
</tr>
<tr>
<td>Total time</td>
<td>346</td>
<td>532</td>
<td>591</td>
</tr>
<tr>
<td>Black Queen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaze duration</td>
<td>252</td>
<td>299</td>
<td>332</td>
</tr>
<tr>
<td>First fixation duration</td>
<td>219</td>
<td>236</td>
<td>237</td>
</tr>
<tr>
<td>Single fixation duration</td>
<td>223</td>
<td>242</td>
<td>234</td>
</tr>
<tr>
<td>Total time</td>
<td>296</td>
<td>477</td>
<td>579</td>
</tr>
</tbody>
</table>

Hom. error = homophone error; orthog. error = orthographic control error.

As Table 2 shows, the eye fixation data from the new study closely replicated the eye fixation data from the original Daneman et al. (1995) Experiment 1B. Indeed, across-experiment statistical analyses showed no main effect of Experiment (original vs. replication) and no interactions between Experiment and any of the experimental manipulations (all ps > 0.05), so we will focus our discussion on the results averaged across the two experiments.

Take first the results for the Desjardins text. Remember, this is the text in which the correct words were the lower-frequency homophones (e.g., meat, rein), and the homophone errors were the higher-frequency homophones (e.g., meet substituted for meat; hair substituted for hare). It is also a text that was judged to have low contextual constraint in our norming study (the mean predictability of the correct homophone target words was 0.08). And it is the text that showed clear evidence for the orthographic route to meaning access in Experiment 1B of Daneman et al. (1995). In that experiment, the Desjardins text showed no evidence for the early engagement of phonological processes in meaning retrieval because homophone errors were as disruptive as orthographic control errors when initially encountered. However, there was evidence for the delayed involvement of phonology in error recovery in that readers spent less time regressing to homophone errors than control errors. This pattern of early orthographic influences and later phonological influences was closely replicated in the new study.

If we look at the results averaged across all 33 participants who read the
Desjardins text, the gaze durations in Table 2 show that readers initially took 41 ms longer to process a homophone error than its contextually correct mate ($t(32) = 3.53$, $p < 0.003$) and they took 42 ms longer to process an orthographic control error than the contextually correct word ($t(32) = 4.38$, $p < 0.001$). However, there was no significant difference in initial processing time for homophone versus orthographic control errors ($t(32) = 0.08$, $p > 0.90$). This lack of difference in initial processing time for homophone versus orthographic control errors was also observed in first fixation durations ($t(32) = 0.99$, $p > 0.30$) and single fixation durations ($t(32) = 0.03$, $p > 0.95$). The lack of difference suggests that homophone errors were detected as easily as orthographic control errors, and so orthography rather than phonology was used to activate word meanings. We did find evidence for a delayed involvement of phonological codes when we examined total time on the target words, a measure that includes later regressions to the target word. As Table 2 shows, readers spent 216 ms longer reading and rereading a homophone error than its contextually correct mate, and they spent 307 ms longer reading and rereading an orthographic control error than the correct word; however, now the time spent on homophone errors was significantly less than on the orthographic errors (all $p$-values $<0.01$), presumably because for those cases in which readers had successfully detected the inconsistent imposter word (e.g., *meat* substituted for *meet*) they could capitalize on the shared phonology (*/mɪt/*) as a route to recovering the correct alternative (*meet*). The finding of a delayed involvement of phonology in error recovery is noncontentious because Rayner et al. (1998) found it too. However, the finding that homophone errors were as disruptive as orthographic control errors for the low-constraint Desjardins text is inconsistent with Rayner et al.’s finding for low-constraint texts, because Rayner et al. (1998) reported a pattern in which homophone errors were less disruptive than orthographic control errors. Nevertheless, the pattern of results for our Desjardins text is entirely consistent with the pattern of results for our other low-constraint text, Russell Wood (Daneman and Reingold, 1993, Experiment 2; Daneman et al., 1995, Experiment 2) and so we continue to argue that the orthographic route is the dominant route to meaning retrieval when words are not highly predictable as is the case for most content words in natural complex prose.

The results for the Black Queen text differed somewhat from those for Desjardins. Remember, Black Queen is the text in which the correct words were the higher-frequency homophones (e.g., *meet*, *rain*), and the homophone errors were the lower-frequency homophones (e.g., *meat* substituted for *meet; hare* substituted for *hair*). It is also the text that was judged to be higher in contextual constraint than the Desjardins and Russell Wood texts; the mean predictability of the correct homophone target words was 0.28 in Black Queen. Although a mean contextual constraint score of 0.28 is much lower than the 0.86 mean predictability in Rayner et al.’s (1998) high-constraint texts, the predictability scores for some of the words in
the Black Queen were high (0.80–1.00) and there was a wide range of predictability scores across the 30 target words (0.00–1.00, SD = 0.33). In the Daneman et al. (1995) Experiment 1B, the results for the 15 participants who read the Black Queen text showed a hint (albeit a nonsignificant one) of an early phonological influence in that the homophone errors were less disruptive than their orthographic controls. We had attributed the weak phonological effect to the fact that the homophone errors in the Black Queen were words of lower frequency than the homophone errors in the Desjardins text. However, the results of the norming study suggest that the weak effect could be due to the presence of some highly predictable target words in the Black Queen text. By running a replication study with a further 18 participants, we could see whether the nonsignificant trend with 15 participants is indeed real and reliable. We could also examine the effect of contextual constraint by computing the correlation between a target word’s predictability and the degree to which the homophone impostor error was likely to be detected.

As Table 2 shows, the replication study produced a similar pattern of results to the original Daneman et al. (1995) Experiment 1B in that there was a nonsignificant trend of homophone errors being less disruptive than orthographic errors when initially encountered. When statistical analyses were computed for all 33 readers of the Black Queen text, this effect was significant for the gaze duration measure. As the gaze durations in Table 2 show, readers initially took 45 ms longer to process a homophone error than its contextually correct mate ($t(32) = 3.35, p < 0.003$) and they took 80 ms longer to process an orthographic error than the correct word ($t(32) = 6.53, p < 0.001$). Moreover, the 35 ms less spent in initial gaze durations on homophone errors than orthographic errors was also significant ($t(32) = 3.01, p < 0.006$). This finding suggests that homophone errors were less disruptive than orthographic errors, and hence detected less easily, a finding consistent with the view that phonology was used to activate some of the word meanings. One explanation for why there was an early phonological effect in the Black Queen text but not in the Desjardins text is that phonology may be more likely to be implicated in the processing of low-frequency words (Jared and Seidenberg, 1991), and the homophone errors were words of lower frequency in the Black Queen text than in the Desjardins text. A second possibility is that the phonological effect may have been due to the presence of some highly predictable homophone target words in the Black Queen text. If predictability of the correct homophone was related to the likelihood of an impostor homophone being missed rather than detected, then we should expect a correlation between the degree of predictability of a correct homophone and the disruptiveness of the homophone error, such that the higher the

---

6 The trend was there for first fixation durations and single fixation durations but the differences were not statistically significant, for both cases $p > 0.10$. 

predictability of the correct homophone the less disruptive the error. To test this hypothesis, we computed the correlation between the contextual constraint score for each of the 30 target words in the Black Queen text and the disruptiveness of the homophone error relative to the orthographic control error (i.e., orthographic error gaze duration minus homophone error gaze duration), averaged across the 33 participants who read Black Queen. This correlation was 0.47, \( p < 0.008 \), and suggested that the more predictable a target homophone, the more disruptive was the orthographic error relative to the homophone error. This finding is consistent with the findings of Rayner et al. (1998) to the extent that it shows that homophone interference effects (failure to detect homophone impostors) are more likely to occur when words are highly predictable. Presumably, Rayner et al. (1998) would argue that the finding is consistent with their verification model (depicted in Fig. 2) in which there is a bottom-up activation of meaning by phonology. However, as discussed earlier, we believe that an equally plausible alternative is the top-down model (depicted in Figs. 3 and 4) in which context activates (primes) the phonology before the printed word is even fixated. The advantage of this top-down account is that it is entirely consistent with our bottom-up model (depicted in Fig. 1) for processing words in the more natural low-constraint texts.

Concluding remarks

Rayner et al.’s (1998) article made us recognize the importance of taking the contextual predictability of a word into consideration when interpreting our data from the inconsistency detection paradigm. The data from our texts with low target word predictability (Russell Wood: Daneman and Reingold, 1993; Desjardins: Daneman et al., 1995, replication experiment) have consistently shown homophone errors to be as disruptive as orthographic errors, thus providing strong evidence that orthography rather than phonology is used to activate word meanings. However, these data are inconsistent with the data from Rayner et al.’s texts with low target word predictability. The data from our text that had a greater range of target word predictability (Black Queen: Daneman et al., 1995, replication experiment) are consistent with the view that phonology may be involved in activating word meanings, because they showed a pattern in which homophone errors were less disruptive than orthographic control errors, particularly if the correct homophones were highly predictable from the preceding context. These data are consistent with the data from Rayner et al.’s (1998) texts with high target word predictability, but we provide a different model to account for them.

We believe that our results for low-constraint text are best accommodated by the model depicted in Fig. 1. According to this model, when a word is not highly predictable from the preceding context, as is the case for most content words in
everyday texts, the dominant route from print to meaning is via the orthographic representation. However, the model leaves open the possibility of phonologically mediated lexical access, as may be the case when word recognition is slowed down by an unfamiliar or low-frequency word. A nice feature of our model is that it can be extended to account for top-down sources of activation in highly predictable contexts, as depicted in Figs. 3 and 4. According to the model depicted in Figs. 3 and 4, the homophone effects found by us and Rayner et al. (1998) for highly predictable words are a result of meaning activating phonology and not phonology activating meaning as Rayner et al. (1998) advocate in their verification model depicted in Fig. 2. However, we should stress that most content words are not highly predictable from their preceding context, and so top-down effects play a relatively small role in reading natural text. We certainly see no reason for invoking the verification model to account for the homophone effects found in the case of highly predictable words. The verification model does not account for our low-constraint data and it does not account for the low-constraint data of Rayner et al. (1998) either.

But how do we account for the contradictory findings for low-constraint text, for the fact that Rayner et al. (1998) did find evidence for the early involvement of phonology in meaning retrieval when target words had low predictability, whereas we have repeatedly not found such evidence? A recent paper by Jared, Levy and Rayner (1999) may provide some answers to this puzzle. Jared et al. (1999) used a very similar eye movement inconsistency detection task to ours, but they also examined whether factors such as reading skill and word frequency could account for the contradictory across-lab findings for low-constraint texts. Their results suggested that both factors may have played a role. Essentially, Jared et

---

7 Pollatsek et al.'s (1992)'s parafoveal previewing task is another on-line that has yielded data inconsistent with ours. In this chapter we have focused on research that has looked for evidence of phonological involvement during reading by showing homophone interference effects. However, there have also been some on-line studies that have looked for evidence of phonological involvement by showing homophone facilitation effects in preview or priming paradigms (e.g., Pollatsek et al., 1992; Rayner et al., 1995). For example, in the study of Pollatsek et al. (1992), processing of a target word (e.g., beech) was facilitated if a homophone of that target word (e.g., bench) had been presented as a preview in the parafovea more so than if a visually similar control (e.g., bench) had been presented as a preview in the parafovea. The Pollatsek et al. data are inconsistent with our data because they suggested that phonological codes are activated very early in the word identification process, even before the word in question is foveally fixated. Interestingly, however, the advantage of a homophone preview over a visually similar preview was only statistically reliable for the first fixation on the target word but not for the gaze duration on the target word (see Pollatsek et al., 1992, Experiment 2). We believe that even if the preview paradigm has provided evidence for very early (nonlexical) involvement of phonology in the word encoding process, it has not provided evidence for the involvement of phonological codes in the subsequent process of accessing a word’s meaning.
al. (1999) closely replicated our pattern of findings for good readers and for high-frequency words. Across three experiments, Jared et al.'s good readers showed the same pattern of data that we have continued to find; good readers were as disrupted by homophone errors as by orthographic control errors, suggesting that they used the orthographic route to retrieving word meanings. Jared et al.'s pattern of data for poor readers was similar to Rayner et al.'s (1998) data; poor readers were less disrupted by homophone errors than by orthographic errors, suggesting that they used the phonological route to meaning. Jared et al. also found an effect, albeit a less consistent one, for word frequency, such that lower-frequency words showed some evidence for phonological effects whereas higher-frequency words did not. Jared et al. (1999) suggested that "Daneman and colleagues may have tested primarily good readers using higher frequency words whereas Rayner et al. may have included readers with a wider range of abilities and used more lower frequency words" (p. 253). Thus, it would appear that factors such as predictability, word frequency, and reading skill may all have an effect on the extent to which phonology is implicated in activating a word’s meaning.

In any event, we believe that there is now considerable evidence against the strong early phonology position. Our findings, together with Jared et al.'s findings for good readers and high-frequency words, make a compelling case against any model that gives phonology an early and exclusive role in gaining access to a word's semantic representation.

Acknowledgements

Preparation of this paper was supported by grants to Meredyth Daneman and Eyal Reingold from the Natural Science and Engineering Research Council of Canada. We thank Robin Morris and Sandy Pollatsek for their reviews of the chapter. Correspondence should be addressed either to Meredyth Daneman, Department of Psychology, University of Toronto, Mississauga, Ontario, Canada, L5L 1C6 (E-mail: daneman@psych.utoronto.ca) or to Eyal Reingold, University of Toronto, Department of Psychology, 100 St. George Street, Toronto, Ontario, Canada, M5S 3G3 (E-mail: reingold@psych.toronto.edu).

References


Phonology and meaning activation


