The Use of Phonological Information in Skilled Silent Reading

Debra J. Jared

Department of Psychology
McGill University, Montreal

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

October, 1990

THE USE OF PHONOLOGICAL INFORMATION IN SKILLED SILENT READING

Abstract

Six experiments were conducted to address the role of phonological information in visual word recognition. A semantic decision task was used to ensure that word meanings were accessed. Experiments 1-4 showed that subjects make more false positive errors on homophone foils (e.g. living thing- FLEE) than on spelling controls (e.g. living thing-FLEX) only when both members of the homophone pair are uncommon and are similarly spelled. In Experiment 5, there was an increase in errors on low but not high frequency homophone category exemplars when they were preceded by a word related to the other member of the homophone pair (e.g. SHATTER-BRAKE). In Experiment 6, subjects produced longer decision latencies on homophone exemplars than on semantic controls only when they were low in frequency. These results indicate that, even in skilled readers, phonological information mediates the access of meaning for low frequency words, and that orthographic activation also contributes to the activation of their meanings.

LA FONCTION DES INDICES PHONOLOGIQUES LORS DE LECTURE À VOIX BASSE CHEZ DES LECTEURS EXPÉRIMENTÉS

Résumé

La fonction des indices phonologiques dans la reconnaissance visuelle des mots (anglais) a été examinée à l'aide de six expériences de décisions semantiques afin de s'assurer que le sens des mots étaient acquis Les études 1 à 4 ont démontré que les sujets ont fait plus d'erreurs (en répondant OUI alors que la réponse était NON) sur les homophones déjoués (e g. créature vivante-PAIR) que sur les contrôles orthographiques (e.g. créature vivante-PAIX) seulement quand les deux membres du couple homophoniques étaient des mots de basse fréquence et dont l'épellation était semblable. L'étude 5 a révélé qu'il y a augmentation des erreurs sur les exemples homophoniques de basse fréquence (et non sur les exemples fréquents) quand ces exemples sont précédés par un mot relié sémantiquement à l'autre membre du couple homophonique (e g COUPLE-PERE). Enfin, lors de l'étude 6, les sujets ont mis plus de temps pour prendre leurs décisions face aux exemples homophoniques que face aux contrôles sémantiques seulement quand il s'agissait de mots de basse fréquence. Ces résultats indiquent d'une part, que même chez les lecteurs expérimentés, les indices phonologiques jouent un rôle dans le processus d'accès sémantique pour les mots de basse fréquence, et d'autre part, que les indices orthographiques contribuent également à l'accès sémantique.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Mark Seidenberg, for his guidance and support throughout my years at McGill. He has given me the freedom to test my own ideas in the lab, and has challenged me to figure out how the details relate to the larger issues in reading and human cognition. His theories have been a major inspiration for my work.

I would also like to thank all of the undergraduates who participated in these and the other studies I have conducted while at McGill Their enthusiasm and effort in the face of often long lists of stimuli and repeated visits were much appreciated. Without them, this research would not have been possible.

Many others have made my stay at McGill an enjoyable experience, in particular, my fellow graduate students in the lab, Ken McRae, Phil Hetherington, and Michaela Hynie have provided good conversation that has made the journey more pleasant. In addition, Vic Ferreira wrote a number of programs that facilitated the setting up and analysis of the experiments, and Fernande Charron smoothed out the bumpy French in my translation of the abstract

I would also like to express my gratitude to my parents who have supported and encouraged me throughout my years at university, and who taught me to value a good education.

And, finally, I would like to express my deep appreciation to my husband, John Mitchell, for his constant support and love that have given me the strength to make it this far. I would also like to thank him for his willingness to let me monopolize the computer for the several months during which this thesis was written. To John, with love.

Table of Contents

List of Figures

List of Tables

Written languages differ in the ease with which the pronunciation of an unknown word can be derived. Historically, written languages have evolved towards more direct representations of the phonemes of the spoken language (Henderson, 1982; Hung & Tzeng, 1981). The reason for this trend is unclear, although a possible explanation is that phonological information may serve a useful function in skilled reading or learning to read (Seidenberg, Waters, Barnes, & Tanenhaus, 1984). The advantage of languages that encode phonology is that they offer readers two ways to access the meanings of words. In an alphabetic language such as English, the meaning of a printed word can be accessed directly on the basis of an analysis of the visual pattern, but another possibility is that letters can be translated into sounds, and meaning accessed on the basis of the spoken form. The extent to which readers of English translate letters into their phonological representations in order to access meaning has been a central issue in research on word recognition. A major difference in word recognition theories is in the relative importance they place on the two possible recognition processes.

Psychologists first began studying the role of phonological information in skilled silent reading over a century ago. Huey (1908/1968) devoted two chapters of his classic book to the topic of inner speech in reading. Recent interest in this issue stems from the publication of a seminal paper by Rubenstein, Lewis, and Rubenstein in 1971. They, and several subsequent theorists (Gough, 1972; Spoehr & Smith, 1973), argued that a printed stimulus is always recoded into its phonological representation in reading and that this is the only representation used to access meaning. Support for this view came from studies demonstrating an influence of phonology on word recognition, such as longer response times on homophones (e.g. BARE/BEAR) than on nonhomophones (e.g. BAKE) and an effect of the ease with which a words could be translated into their phonological representations (e.g. LOSE vs LATE). An opposing view claimed that phonological information plays no role in skilled word recognition and suggested instead that meaning is accessed directly on the basis of visual information (Becker, 1976, 1980; Brown, 1987; Johnson, 1975; Paap, Newsome, McDonald, & Schvaneveldt, 1982; Rumelhart & Siple, 1974; Smith, 1971). Support for this view came primarily from studies that failed to observe phonological effects (McCusker, Hillinger, & Bias, 1981).

In light of the considerable evidence for each of these positions, it is perhaps not surprising that many researchers ended up concluding that both visual and phonological pathways exist and operate in parallel (Coltheart, 1978, 1980; Forster & Chambers, 1973; Laberge & Samuels, 1974, McClelland & Rumelhart, 1981; Meyer, Schvaneveldt, & Ruddy, 1974; Paap, McDonald, Schvaneveldt, & Noel, 1987; Patterson & Morton, 1985; Seidenberg & McClelland, 1989; Shallice, Warrington, & McCarthy, 1983). Dual route theorists differ in their descriptions of the operation of the phonological route, and in their views of the relative timing of processing on the two routes. Coltheart (1978, 1980), for example, proposed that processing along the visual route is almost always faster than via the phonological route in skilled readers, whereas other researchers have suggested that processing along the visual route is faster only for common words (Andrews, 1982; Laberge & Samuels, 1974; Seidenberg et al., 1984). However, all of these theories have in common the assumption that processing occurs along both visual and phonological pathways when a printed word is presented. This dual route approach has dominated the field and is widely accepted.

Recently, Van Orden (1987; Van Orden, Johnston, & Hale, 1988) has questioned the evidence supporting the direct visual access route, and instead proposed that word candidates are activated exclusively by their phonological representations. A verification procedure then checks the spelling of candidates against the target stimulus. This view is essentially the same as the one proposed by Rubenstein et al. in 1971 that initiated interest in the role of phonological information in word recognition Thus, if Van Orden is correct, after nearly 20 years of research on the question of whether the access of meaning is phonologically mediated, we are back to where we started. The studies that have been conducted in the past two decades using other methodologies will be reviewed shortly to determine the extent of support for the view that the access of meaning is phonologically mediated, and then a series of experiments will be reported that further examine the performance of subjects on the task used by Van Orden to support his view

It is important to determine the role of phonological information in the skilled reading of English, because ideas about how best to teach children to read are often based on a theory of skilled reading (Crowder, 1982). There has been a long and stormy debate in reading education as to whether or not it is useful for beginning readers to be explicitly taught the spelling-sound relationships of English. Phonics programs were widely abandoned in favor of whole word methods earlier this century, and part of the justification for this was the theory that skilled readers recognize words on the basis of their overall shape (see Adams, 1990; Crowder, 1982). More recently, phonics programs have again fallen into disfavor, partly on the basis of arguments about skilled reading made by Goodman (1972, 1973) and Smith (1971). Goodman (1972, 1973) has argued that skilled readers use context to generate predictions about upcoming words and then visually inspect the text only as much as is needed to confirm their predictions. Smith (1971) has questioned the plausibility that spelling-sound rules are used because of the complexity of the rules needed to characterize English and the uncertainty as to the appropriate rule for a particular word. None of these positions view phonological information as playing a role in skilled word recognition. However, neither the word shape hypothesis nor Goodman's hypothesis of skilled reading have been supported by experimental work (Adams, 1979; McConkie & Zola, 1981), and Seidenberg and McClelland (1989) have shown how the pronunciations of words can be derived without making use of the spelling-sound rules objected to by Smith (1971). A demonstration that phonological information does play an important role in skilled word recognition would provide support for the view that spelling-sound translation lessons are a valuable component of beginning reading instruction (Adams, 1990).

Evidence that phonological information plays a role in skilled reading could also provide insight into the reading problems of the hearing impaired. Conrad (1977) estimated that only 4.5% of hearing impaired students leaving high school in England and Wales can read at a level appropriate for their age, and even fewer profoundly deaf read at this level. Half of the 15 and 16 year old hearing impaired students he studied scored below the fourth grade level on a sentence completion task. Studies conducted in the United States have found similar results on paragraph reading tests (DiFrancesca, 1972; Trybus & Karchmer, 1977). The decreased opportunity to benefit from the spelling-sound relationships present in English may be one reason for this difficulty (Treiman & Hirsh-Pasek, 1983).

It is important, then, to determine whether phonological information is used by skilled readers to access the meanings of words. The results of a number of recent studies strongly suggest that, in skilled readers, the phonological representations of words are available rapidly and automatically. Three studies have found evidence that phonological properties of a stimulus influence performance on a task in which the subjects are required only to name the ink color of the stimulus. Longer color naming latencies were found for pronounceable nonwords than for unpronounceable nonwords (Bakan & Alperson, 1967), for words sharing initial or final sounds with incongruent color words (e.g. BLOT or FLEW printed in green ink) than for unrelated words (Dalrymple-Alford, 1972), and for target words (e.g. FOOD) preceded by primes that were phonologically similar but orthographically different (e.g. RUDE) than for targets preceded by unrelated (e.g. WELL) primes (Tanenhaus, Flanigan, & Seidenberg, 1980). Since the activation of phonological information hindered performance on each of the tasks, these results suggest that it was not a conscious strategy of the subjects. Further, Humphreys, Evett, and Taylor (1982) have demonstrated phonological priming in a tachistoscopic recognition experiment under masking conditions that produced minimal identification of the prime itself And finally, Perfetti, Bell, and Delaney (1988) have provided evidence that phonological information is rapidly activated using a backward masking paradigm. Briefly presented targets (e.g. HOLE) masked by a pseudoword that sounded like the target (e.g. HOAL) were more accurately identified than targets masked by orthographic (e.g. HORL) or unrelated (e g. PLIS) pseudoword controls.

Thus, the phonological representations of words appear to be available quickly and automatically in skilled readers. However, this does not necessarily mean that such representations are used to access the meanings of words. They may instead be used only to generate a pronunciation when reading aloud or serve as a memory code for words pending comprehension of an entire phrase or sentence. The review of the literature below, and the studies that follow, attempt to determine whether or when phonological information is used to access the meanings of words in skilled silent reading. Most of the relevant studies in the literature have used the lexical decision task. A smaller number that have used the sentence evaluation task and the category decision task will also be reviewed, along with evidence from neurological patients.

The conclusion that is reached at the end of the review is that we still do not know for sure whether or not phonological information is used to access meanings of words. However, a critical examination of previous studies leads to a clearer picture of the sort of task and stimuli that would more definitively address the question. A series of experiments is then presented based on the approach that emerges from the review as the most likely one to yield an answer to the elusive issue of whether or not access to meaning is phonologically mediated in silent reading.

Lexical Decision Studies

As mentioned earlier, most of the experiments examining the role of phonological information in silent word recognition have used the lexical decision task. In this task, subjects are presented with a series of letter strings and must decide whether each string is a word or a nonword. Coltheart, Davelaar, Jonasson, and Besner (1977) have argued that results from studies using the lexical decision task are relevant to the issue of phonological mediation provided that the nonwords are not orthographically illegal (e.g. PVDHW). When illegal nonwords are used, subjects can base their decisions solely on the orthographic characteristics of the stimuli. Experimenters employing the lexical decision task have used several different strategies to determine whether phonological mediation occurs. These include comparing performance on homophonic and nonhomophonic stimuli, examining whether the ease of translation into a phonological representation affects performance, examining whether presentation of a phonologically related prime influences responding, and exploring whether performance deteriorates with concurrent articulation.

Homophones

One strategy used to determine whether or not words are phonologically recoded prior to the access of meaning has been to compare subjects' lexical decision performance on homophonic and nonhomophonic stimuli. The logic of this approach is that if letter strings are translated into their phonological representations, then homophonic stimuli should produce some confusion relative to nonhomophonic stimuli. On the other hand, if access to meaning occurs directly on a visual basis, the two types of stimuli should not differ.

Pseudowords. One group of studies has compared subjects' latencies to reject pseudowords that are homophonic to real words (e.g. GRONE) and those that are not homophonic (e.g. SLINT). If letter strings are phonologically recoded, performance on pseudohomophones should be slowed because their phonological representations activate representations of words and must be checked for spelling, whereas pseudowords do not. Rubenstein et al (1971) found that pseudohomophones produced significantly longer decisions and more errors than pseudowords. However, Meyer et al (1974) pointed out that pseudohomophones might produce longer decisions because they may look more like English words. The appropriate orthographic control for pseudohomophones has been a subject of debate among researchers. Some have observed a pseudohomophone effect (Barry, 1981; Besner & Davelaar, 1983; Besner, Twilley, McCann, & Seergobin, 1990, Coltheart et al., 1977), while others have not (Martin, 1982; Seidenberg & McClelland, 1989). Other work has provided evidence that a pseudohomophone effect is found when a small proportion of nonwords are pseudohomophones (McQuade, 1981), and in studies that include homophones among the word stimuli (Dennis, Besner, & Davelaar, 1985, but see McCann, Besner, & Davelaar, 1988).

These studies demonstrate that the pseudohomophone effect may occur in some circumstances. Coltheart et al. (1977) convincingly argue, however, that the finding of longer rejection latencies for pseudohomophones does not provide evidence about whether lexical access is phonologically mediated. This is because NO latencies are typically longer than YES latencies in a lexical decision task. It is possible that phonological recoding is a relatively slow process and that the meanings of words are usually activated before it has been completed. The extra time required to reject a pseudoword may be sufficient to allow this process to finish. Coltheart et al. (1977) conclude that in a decision task, only the finding of effects on YES responses would provide unequivocal evidence for phonological recoding prior to lexical access. Coltheart (1978) and Dennis et al. (1985) note, however, that the existence of a pseudohomophone effect demonstrates that the phonological code of a letter string is always generated despite the fact that it leads to slower responses on pseudohomophones.

Words. Another group of studies has compared lexical decision latencies for homophone and nonhomophone words. The logic of this approach is that if access to meaning occurs on the basis of phonological information, homophones (e.g. FLEA) will activate two entries (insect and run) whereas appropriately matched nonhomophones (e.g. MOTH) will activate only one (insect). Rubenstein et al. (1971) suggested that a spelling check would then be done, comparing the spelling associated with each meaning to the spelling of the stimulus presented. The highest frequency alternative is checked first, and if it does not match the stimulus, the other alternative is checked. Thus, on some occasions, the recognition of homophones would be slowed because the wrong alternative was checked first. Consistent with their hypothesis, Rubenstein et al. (1971) found that homophones produced longer decision latencies and more errors than nonhomophones, and they concluded that lexical access is phonologically mediated.

The results of Rubenstein et al.'s (1971) experiment were challenged by Clark (1973) and Coltheart et al. (1977). Clark (1973) demonstrated that their effect of homophony was not significant when items were treated as a random factor. A possible reason for the nonsignificant result in the item analysis for homophones is suggested by their posthoc analysis demonstrating that the effect of homophony occurred only for the homophones that were the less frequent member of the homophone pair. This finding is consistent with their hypothesis since lower frequency members are checked only after the higher frequency member has been found not to match the spelling of the stimulus.

Coltheart et al 's (1977) criticism was that the homophones and nonhomophones were not matched for frequency, number of letters, or part of speech. However, although the words may not have been explicitly matched on these factors, they did not differ greatly. The mean frequency (using the norms of Kucera and Francis, 1967) of the homophones was 18.6 and the mean frequency of the nonhomophones was 20.9. The mean length of the homophones was 4.2 letters and the mean length of the nonhomophones was 4.0 letters. Coltheart et al. (1977) compared lexical decision latencies for 39 homophones and 39 nonhomophones that were matched for frequency, length, and part of speech. To increase the likelihood of finding an effect of homophony according to Rubenstein et al.'s (1971) theory, all homophones were the less frequent member of a homophone pair. They failed to find an effect of homophony, and argued that this finding contradicted Rubenstein et al.'s (1971) view that words were necessarily phonologically recoded prior to lexical access. Coltheart et al. (1977) proposed instead that word recognition occurs on the basis of visual information. Their finding was replicated by Dennis et al. (1985, Expt. 4) who used the same words and found no effect of homophony in the latency data and an effect in the error data that was significant only by subjects. In contrast, however, they found in another experiment that a subset of 25 of the 39 homophones produced a significant effect on decision latencies in the subject data and a significant effect in errors in both the subject and item data (Dennis et al., 1985, Expt. 3). Barry (1981) presented a subset of 16 of the 39 homophones and their controls tachistoscopically, once to each visual field, and found a significant effect of homophony on both decision latency and accuracy. The effect on latency was small and primarily in words presented to the right visual field, although effects may have been attenuated because the items were repeated.

A study by Davelaar, Coltheart, Besner, and Jonasson (1978) suggests that a homophone effect is found in a lexical decision task only when the pseudowords in the experiment do not sound like English words. Davelaar et al. (1978, Expts. 3 & 4) found an effect of homophony when the pseudowords in the experiment were items such as SLINT but not when they were pseudohomophones such as GRONE. This homophone effect occurred only for homophones that were the lower frequency member of a homophone pair. Rubenstein et al. (1971), Coltheart et al. (1977), Dennis et al. (1985), and Barry (1981) all included pseudohomophones among their pseudowords, and hence the results of this experiment provide a possible account for their negative or weak results

Davelaar et al (1978) suggested that their finding that the effect of homophony in lexical decision depends on the nature of the pseudowords included in the experiment indicates that phonological recoding is a strategy under the control of the subject. Naive subjects rely on the outcome of phonological recoding, but they abandon this strategy when it produces too many errors, as when pseudohomophones are included, and instead rely solely on a visual strategy. However, Henderson (1982) has pointed out a problem with this account. Davelaar et al (1978) claim that homophone effect found in the SLINT environment with the lower frequency member of a homophone pair occurs because a spelling check is made on the higher frequency member first, and when that fails a second spelling check is then done on the lower frequency member. This second spelling check is done "to avoid erroneous YES responses to pseudohomophones" (p. 400). However, they claim that when pseudohomophones are present, the subject makes many errors and so ignores the results of phonological processing. But such errors would only occur if subjects were responding before the second spelling check. There is no reason why they should not continue on as proposed in the case of the SLINT nonword environment and do a spelling check on other lexical alternatives. Thus, the same strategy they propose to account for homophone effects in the SLINT environment should lead to accurate performance and similar homophone effects in the GRONE environment. It is not clear, then, why subjects would abandon a phonological strategy when pseudohomophones are included but maintain it in the presence of homophones. If anything, the inclusion of pseudohomophones should strengthen the homophone effect if indeed it results from a spelling check procedure. When no pseudohomophones are included, the spelling check is not necessary since the subject can respond YES to a homophone as soon as either meaning is activated by its phonological representation. The presence of pseudohomophones should make the spelling check mandatory and thus produce a homophone effect.

Davelaar et al.'s (1978) explanation of their finding, then, cannot account for both the presence of a homophone effect in the SLINT environment and its absence in the GRONE environment. One possibility is that the finding itself is an artifact of their procedure. Different sets of homophones were used in the pseudoword and pseudohomophone environments. Since the homophone effect is rather elusive, it seems imperative to demonstrate that only the change in pseudoword environment produces the difference in the size of the effect. There is some evidence, however, that the Davelaar et al. (1978) result will replicate. The 29 homophones used in the SLINT environment were all included in the Coltheart et al. (1977) experiment in which 50% of the pseudowords were pseudohomophones. The means for this subset of homophones and their controls can be calculated from the item means presented in the appendix of the paper. The latency for homophones was 18 msec faster than for matched nonhomophones, a reversal of the effect found with these words in the SLINT environment of Davelaar et al.'s (1978) experiment. The homophone effect appears, then, to require a different explanation than that of Davelaar et al. (1978) and Rubenstein et al. (1971).

One other study using homophones appears to have found support for the strategic use of phonology (Hawkins, Reicher, Rogers, & Peterson, 1976). Hawkins et al. found an effect of homophony in a tachistoscopic recognition task when a low proportion of homophones were included but not when the stimulus list contained a high proportion of homophones. However, the tachistoscopic recognition task does not require access to the meanings of words, so it is unclear whether such flexibility is possible when the task requires access to meanings. It remains to be seen whether the proportion of homophones included in a lexical decision task will influence the size of the homophone effect.

In sum, the results of lexical decision studies using homophones have not provided unequivocal evidence that the access of meaning is phonologically mediated. There is a possibility that phonological mediation occurs only for some words, although these studies do not indicate the relevant characteristics of such words beyond the suggestion that homophone effects are more likely to occur for the lower frequency members of a homophone pair. There is also the possibility that the use of phonological mediation depends on the nature of pseudowords included in the experiment. If this turns out to be true, it would imply that the lexical decision task is of limited value in determining whether phonological mediation occurs in the normal reading since text does not include nonwords of either type.

Homographs

Seidenberg et al (1984) hypothesized that performance on homographs, words that have two pronunciations associated with different meanings (e.g. DOVE, WIND), should differ from nonhomographs if subjects recode the words into their phonological representations. In their study, lexical decision latencies were similar for homographs and matched regular words, although a significant difference had been found in naming. It is possible that an effect was not evident because the homographs and regular words were not matched for number of meanings. A disadvantage of having two pronunciations may have been cancelled out by an advantage of having more than one meaning (Chumbly & Balota, 1984; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Rubenstein, Garfield, & Milliken, 1970). Such an advantage is not found in naming (Balota & Chumbly, 1984; Chumbly & Balota, 1984) because the task does not require the access of meaning. It is not possible to determine whether the size of the homograph effect is modulated by frequency because there are so few homographs in English and most are fairly common words.

Ease of Translation

Another group of studies has attempted to determine whether access to meaning is phonologically mediated in silent reading by examining whether the ease with which a letter string can be translated into its phonological representation affects lexical decision latencies. This has been examined by comparing nonwords that differ in pronounceability, words that differ in length, and words that differ in the consistency of their pronunciations relative to similarly spelled words No effect of ease of translation would be expected if access to meaning occurs directly on a visual basis.

Nonword Pronounceability. Rubenstein et al. (1971) argued that if letter strings are phonologically recoded, rejection latencies for nonwords matched for orthographic legality should be affected by their pronounceability. They reported that orthographically illegal pronounceable strings (e.g. GRATF) took significantly longer to reject than illegal unpronounceable strings (e.g. LIKJ); however this difference was not significant in an analysis by items (Clark, 1973). In a further, more carefully controlled study, Rubenstein, Richter, and Kay (1975) found a reliable disadvantage for the pronounceable strings. Other studies (e.g. Richardson, 1976, Snodgrass & Jarvella, 1972, Stanners, Forbach, & Headley, 1971; Walker, 1973) have compared rejection latencies for pronounceable and unpronounceable nonwords but did not match for orthographic legality. However, since words, by definition, must be both pronounceable and orthographically legal it is unclear what implications these findings have for normal reading. Also the observations using this technique are made on NO latencies, and as Coltheart et al (1977) argue, they may produce phonological effects not seen on faster YES trials

Length. Frederiksen and Kroll (1976) and Richardson (1976) hypothesized that if phonological recoding occurs prior to the access of meaning, then factors that affect the latency to name words aloud, such as word length, should also influence performance on the lexical decision task. Instead, they both found that the effects of the number of letters in a word did not replicate in a lexical decision experiment, and concluded that phonological recoding is not a prerequisite for retrieval of meaning On the other hand, Balota and Chumbly (1984), Chumbly and Balota (1984), Forster and Chambers (1973), and Hudson and Bergman (1985, Expts 1 & 2) have found that length effects observed in naming did replicate on a lexical decision task. Further research is needed to resolve this inconsistency However, if length effects turn out to be absent in lexical decision, it would not necessarily imply that access to meaning does not involve the use of phonological information. Their absence could be explained by assuming that the length effects found in naming are a product of the derivation of an articulatory code used to make a naming response, and not the derivation of a phonological code (Coltheart, 1978; Davelaar et al., 1978, Henderson, 1982, McCusker et al., 1981). Henderson (1982) has argued that until it is shown that length effects are a necessary product of phonological recoding, their absence in a lexical decision task does not constitute evidence against phonological recoding in silent reading.

Two findings argue against articulatory-motor preparation as the locus for length effects in naming Richardson (1976) pointed out that word length effects are not found when subjects name objects (e.g. Oldfield & Wingfield, 1965; Carroll & White, 1973). Further, four experiments (Balota & Chumbly, 1985; Cosky, 1976; Eriksen, Pollack, & Montague, 1970; Jared, unpublished study) have reported that the length effect they had found in immediate naming disappeared in a delayed naming condition which has been assumed to measure the response preparation component of naming, particularly if the delay is unpredictable (Balota & Chumbly, 1985). These studies imply that the locus of the length effect in naming is quite early, and thus may reflect spelling-sound translation. The absence of a length effect in lexical decision would be evidence contrary to the phonological recoding view.

On the other hand, if length effects turn out to be present in both naming and lexical decision, this would not necessarily imply that phonological recoding occurs in both tasks It is possible that a length effect in naming reflects the spelling-sound translation process but that in lexical decision it is the result of a post-recognition spelling check Hudson and Bergman (1985) have argued for this position on the basis of a demonstration of a frequency effect but no effect of length in lexical decision when the nonwords were illegal and when they were legal but unlike any real words This argument rests on the assumption that frequency effects could arise only if the lexicon had been accessed. An alternative, though, is that subjects based their decisions on orthographic familiarity, and that this decision was faster for high frequency words (Balota & Chumbly, 1984)

No firm conclusions, then, can be made at present from this line of research. First, the existence or absence of a length effect in lexical decision must be established and an explanation given for the inconsistent results in the literature A suggestion comes from Cosky (1976) and Jared and Seidenberg (1990) who have shown that in naming, length effects are smaller for high frequency words than for low frequency words It is possible that length effects are attenuated in lexical decision and occur only for low frequency words. Secondly, if it is confirmed, the locus of the length effect needs to be established The length effect appears to arise early in processing for naming, but the findings for lexical decision are less clear. A later occurring effect in lexical decision would be evidence against the phonological recoding hypothesis And finally, even if both effects were shown to occur early, the results would only be suggestive of a similar cause, phonological recoding, and not conclusive evidence. Length effects could arise from an early visual processing stage common to both tasks

Another measure of length for which visual processing requirements can be equated is the number of syllables in a word. As with letter length effects, if syllable effects are found to arise early in naming but not in lexical decision, it would be evidence against the view that the access of meaning is phonologically mediated. The evidence so far is inconclusive. Four studies (Balota & Chumbly, 1984; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Richardson, 1976) failed to find an effect of number of syllables on lexical decision, but they did not find syllable effects on naming either. Several naming studies have found syllable effects (Butler & Hains, 1979; Klapp, 1971; 1974; Klapp, Anderson, & Berman, 1973) but others have not (Henderson, Coltheart, & Woodhouse, 1973; Mason, 1978). Jared and Seidenberg (1990) demonstrated an effect of number of syllables for words matched on number of letters only when words were low in frequency. They suggested that the longer naming latency associated with words having an extra syllable may be due to the fact that these words usually contain more vowels, which are a primary source of spelling-sound inconsistency in English. If this account is correct, and lexical access is phonologically mediated, then syllable effects for these words should not be present in a delayed naming task and should be found in lexical decision. Klapp et al. (1973) and Jared (unpublished study) have found that syllable effects present in immediate naming disappear in a delayed naming task, but it remains to be determined whether syllable effects are found with low frequency words in lexical decision.

Thus, although it is very difficult to determine whether an effect of length in letters actually reflects the spelling-sound translation process, the effect of length in syllables found in naming low frequency words probably does reflect this process. However, an effect of number of syllables has not been demonstrated in a silent reading task such as lexical decision. Thus, despite the initial claims of Frederiksen and Kroll (1976) and Richardson (1976), it cannot be determined from this research whether or not phonological information plays a role in the access of meaning.

Spelling-sound Consistency. A potentially strong piece of evidence that the access of the meanings of words is phonologically mediated would be to find an effect of the regularity or consistency of spelling-sound correspondences in a lexical decision task (Henderson, 1982). Such an effect would have the advantage that it would be evident on

YES decisions and that it is directly related to the conversion of print to sound. The particular words considered to be more difficult to translate from spelling to sound depends on the theory of the translation process. According to Coltheart (1978), evidence for phonological recoding prior to the access of meaning would consist of demonstrating that exception words, words for which grapheme-phoneme correspondence rules (e g. Venezky, 1970; Wijk, 1966) fail to yield a correct pronunciation (e.g HAVE), produce longer decision times and more errors than regular words (e.g. GAVE and MUST), whose pronunciations can be correctly derived using these rules. Other theories (Glushko, 1979, Kay & Marcel, 1981; Seidenberg & McClelland, 1989) suggest that the phonological representation of words containing an inconsistently pronounced letter pattern (e.g. -AVE as in HAVE and GAVE) will take longer to produce and will be more prone to errors than that of words containing a consistently pronounced pattern (e.g. -UST as in MUST), and that the degree of difficulty depends on the degree of inconsistency (Jared, McRae, & Seidenberg, in press). These categories are not mutually exclusive, exception words usually have a pronunciation that is inconsistent with most words with same spelling pattern (e.g. HAVE) but this is not always true (e.g. MILD). Some regular words are also inconsistent (e.g. GAVE) and some are consistent (e.g. MUST)

The results of lexical decision experiments have been mixed. Stanovich and Bauer (1978) found that exception words produced significantly longer decision latencies and more errors than regular words, but Coltheart, Besner, Jonasson, and Davelaar (1979) failed to observe effects of regularity in two separate lexical decision experiments on the same stimuli that had produced a regularity effect when named. Mason (1978) did not find an effect of regularity for either good or poor college readers, although two aspects of this experiment mitigate against finding an effect; the exception words used were higher in frequency than the regular words and subjects saw each stimulus twice. Bauer and

Stanovich (1980) used the same words as in one of the Coltheart et al. (1979) experiments and also did not obtain a regularity effect. They pointed out that this set of words contained a number of inconsistently pronounced words among the regular words. In a further experiment they selected nine exception, regular-inconsistent, and regular-consistent monosyllabic words from the Coltheart et al. (1979) words and found that both exception and regular-inconsistent words produced longer latencies and more errors than regular-consistent words, although the latency difference for exception words was not statistically significant. They constructed another list with 50 exception and 50 regular words and, in two separate experiments, found that exception words produced significantly longer decision latencies. This result is somewhat surprising given the results of their second experiment, since 25% of their regular words were inconsistent and 11% of their exception words were consistent (Parkin, 1982). Andrews (1982) factorially manipulated regularity and consistency and found that only spelling-sound consistency had an effect on lexical decision latencies, and then only for low frequency words.

Other researchers have also found an effect of spelling-sound translation difficulty only under certain conditions. Parkin (1982) found an effect only for exception words with sufficiently unusual pronunciations that they were listed in the Oxford paperback dictionary (e.g. QUAY). Parkin and Underwood (1983) examined a possible problem with this experiment, that the effects were due to differences in orthographic not phonological regularity, and reported that orthographically irregular and orthographically regular exception words both produced significantly longer decision latencies than regular words. Waters, Seidenberg, and Bruck (1984) found that both types of exception words produced significantly more errors than regular words, although these effects occurred for low frequency words only. Consistent with this finding, Parkin, McMullen, and Graystone (1986) failed to find an exception effect for orthographically regular, monosyllabic words that were of moderate frequency and Jared (1985) found only a small consistency effect in the error data for moderate frequency multisyllabic words. In contrast, even with low frequency words, Seidenberg et al. (1984) found that neither exception nor regular-inconsistent words produced longer decision latencies or more errors than matched regular words, although in another experiment they did find an exception effect for orthographically irregular ("strange") words.

Waters and Seidenberg (1985) provide evidence that reconciles these conflicting results. They demonstrated that whether or not an exception effect occurs for low frequency, orthographically regular words depends on the composition of the stimuli in the experiment. Exception words produced significantly longer decision latencies and more errors than regular words when strange words were also present in the experiment (although only in the subject analyses), which is consistent with the findings of Parkin and Underwood (1983) and Waters et al. (1984), but the latency effect disappeared when strange words were excluded. This latter finding is consistent with Seidenberg et al. (1984), and along with the frequency explanation, accounts for the results of Parkin et al. (1986) and Jared (1985). Waters and Seidenberg (1985) explained these results by suggesting that subjects adopt variable decision criteria depending on the composition of the stimuli. The inclusion of irregularly spelled words makes the word-nonword discrimination more difficult, making it more likely that phonological information will have time to accrue and enter into the decision process.

Several other researchers have attempted to determine whether strategic factors modulate spelling-sound effects in lexical decision. Stanovich and Bauer (1978) reported that their exception effect disappeared in an experiment in which subjects were required to respond faster than usual through the use of a response deadline procedure. They concluded that phonological recoding takes place either following direct lexical access or in parallel with, but slower than, access based on a visual representation. Waters and

Seidenberg (1985) conducted a response deadline experiment with their exception, regular, and strange words and also found that the exception effect disappeared, although strange words still produced longer decision times and more errors. Parkin (1982) found that the exception effect that he observed for unusually pronounced exception words remained under speeded response instructions; however since these words are analogous to the strange words this may be an orthographic effect. In another study, using the same words and a few additional ones, Parkin and Ellingham (1983) examined whether subjects could be discouraged from using phonological information by including pseudohomophones among the nonword stimuli. They found no difference in the size of the exception effect when half the nonwords were pseudohomophones than when all were nonhomophonic pseudowords, although overall latencies were slower. Andrews (1982) also included a condition in her experiment in which half of the nonwords were pseudohomophones and she too still found an effect of spelling-sound consistency for low frequency words. Surprisingly, subjects produced significantly faster latencies on low frequency words and fewer errors overall than when no pseudohomophones were included. Andrews (1982) suggested that the inclusion of pseudohomophones induced subjects to decrease their reliance on a phonological code.

Thus, the literature on regularity and consistency effects does not conclusively answer the question as to whether phonological representations are used to access meaning in silent reading. No effects of spelling-sound translation difficulty are evident for high frequency words, but this does not necessarily imply that meanings are accessed on a visual basis for these words In the Seidenberg and McClelland (1989) model, the pronunciations of high frequency exception words can be computed as easily as those of high frequency regular words because the model improves its ability to derive the pronunciation of a word each time it is encountered. There is some evidence that phonological mediation occurs for low frequency words but this may depend on the strategy of the subject. The results of several experiments suggest that effects are smaller when response latencies are faster (Bauer & Stanovich, 1978; Waters & Seidenberg, 1985; Andrews, 1982), although this was not seen in Parkin's work (Parkin, 1982; Parkin & Ellingham, 1983) perhaps because the exception words used in the response deadline and pseudohomophone conditions were orthographically irregular. The evidence that the presence or size of the effect depends on the instructions given the subject and the nature of the words and nonwords included in the experiment makes it difficult to assess whether or not phonological representations are used to access meaning in a normal reading situation.

In sum, studies that have attempted to determine whether or not phonological information is used to access meaning by examining whether the ease with which the stimulus can be translated into its phonological representation affects lexical decision latencies have not settled the issue. As in the case of studies using pseudohomophones, studies using pronounceable and unpronounceable nonwords require NO responses which are usually slower than YES responses and may reflect additional processes not operating for YES responses (Coltheart et al., 1977). Further, the use of orthographically illegal nonwords raises questions as to the relevance of the results for the processing of orthographically legal words. Studies examining effects of spelling-sound consistency examined YES responses, but as was the case for studies using homophones, effects of spelling-sound consistency were sensitive to the composition of the stimulus list which makes it difficult to determine whether or not phonological mediation occurs in normal reading. Also, as was the case for homophone studies, there was some evidence that phonological mediation might occur for low frequency words, although these studies cannot rule out the possibility that phonological mediation might also occur for high frequency words. Studies examining length effects also examined YES responses, although effects of length in letters cannot be definitely attributed to the spelling-sound translation process, and the possibility that number of syllables may affect lexical decisions for low frequency words has not been explored. Such a study would also not be able to rule out the possibility that phonological mediation occurs for high frequency words since it is possible that the phonological representations for high frequency words with different numbers of syllables are equally easy to calculate (Seidenberg & McClelland, 1989).

Priming

In still another approach using the lexical decision task, researchers have reasoned that if words are recognized on the basis of phonological representations, then decision latencies should be affected by the phonological relationship between a pair of words presented either simultaneously or sequentially. On the other hand, if access to meaning occurs directly via a visual representation, the phonological relationship between the two words should be irrelevant.

In the first study using this strategy, Meyer et al. (1974) simultaneously presented two letter strings, and the subject's task was to decide whether or not both stimuli were words. The most important comparisons involved word pairs that were orthographically and phonologically similar (e.g. BRIBE-TRIBE; FENCE-HENCE) and their controls (BRIBE-HENCE; FENCE-TRIBE) and word pairs that were orthographically similar but phonologically different (e.g. COUCH-TOUCH; FREAK-BREAK) and their controls (COUCH-BREAK; FREAK-TOUCH). They found slight but nonsignificant facilitation for phonologically similar pairs relative to controls and significant interference for phonologically dissimilar pairs. In a second study using sequential presentation the interference effect was attenuated but still significant. The interference effect has been replicated in three other studies using simultaneous presentation, two with the Meyer et al. words (Shulman, Hornak, & Sanders, 1978; Hanson & Fowler, 1987), and one with a different stimulus set (Hanson & Fowler, 1987), although in all three studies the facilitation effect was also significant. Meyer et al. (1974) concluded that the interference effect indicated that word recognition is mediated at least part of the time through phonological representations. They proposed an encoding bias hypothesis to explain these results. When a second string ends with the same letters as the first, there is a tendency to use the same grapheme-phoneme conversion rules on the second string, which causes difficulty if it does not rhyme. Another possible explanation for the interference effect, however, was suggested by Banks, Oka, and Shugarman (1981) Since a large proportion of the word pairs rhymed, and it is unlikely that many of the pairs containing a nonword did, subjects might adopt a strategy of checking for rhymes. If this was the case, these experiments have little to say about whether the access of meaning is phonologically mediated.

Hillinger (1980) also reported evidence that word recognition is phonologically mediated, although his results did not support the encoding-bias hypothesis He not only found significant facilitation for phonologically similar pairs that were orthographically similar (LATE-MATE), but importantly, also found facilitation for orthographically dissimilar pairs (EIGHT-MATE) using a sequential presentation paradigm While the former effect could be due either to orthographic or phonological similarity, the latter effect could only be a result of phonological similarity. These facilitation effects were found relative to both unrelated (VEIL-MATE) and neutral (****-MATE) controls The latencies for the two control conditions did not differ, which suggests that facilitation effects were not due to subjects anticipating rhymes. If they had been, there should have been a cost to anticipating incorrectly and unrelated pairs would have been slower than neutral controls. In addition, there was no facilitation for words preceded by orthographically similar pseudowords (JATE-MATE), contrary to the encoding-bias hypothesis and to the view that phonological recoding occurs prior to recognition. Hillinger suggested that word primes activate their phonological representations in a phonological access file (see Forster, 1976) and that this activation spreads to phonologically similar words.

Hillinger's (1980) results have, however, been challenged Besner, Dennis, and Davelaar (1985) demonstrated that lexical decision latencies to words could be facilitated by presenting a pseudohomophone of the word on the previous trial (e.g. GROCE-GROSS) instead of an unrelated pseudoword (e.g. BRULT-GROSS). The size of the priming effect was much larger than that for words preceded by orthographically similar pseudowords. This suggests that the activation of phonological representations is not restricted to words, and thus could be occurring prior to recognition.

On the other hand, Martin and Jensen (1988), failed to replicate Hillinger's facilitation effects for word primes using the same procedure and both Hillinger's stimuli and another stimulus set In addition, they also failed to find phonological priming effects with Hillinger's stimuli and a third stimulus set using a procedure in which the prime is presented for a short or long duration and the subject responds only to the target. They concluded that their observations were compatible with a view of the lexicon in which there is no spreading of activation among phonologically similar words, or with the view that lexical decisions are made without phonological mediation. However, since this conclusion is based on null results, it must be viewed cautiously. It is possible that their experiment failed to detect priming effects that actually do exist.

One reason that Martin and Jensen (1988) may not have found priming effects is that they may be modulated by word frequency. The majority of words in these experiments were common words, only 37.5% and 10% of the target words on Martin and Jensen's (1988) two lists had a Kucera & Francis (1967) frequency less than 10, and only 30% of the targets used by Hillinger and 20.8% of those used by Meyer et al. (1974) were in this range. Davelaar et al. (1978) and Columbo (1986) have presented evidence suggesting that word frequency is important. Davelaar et al. (1978) presented strings one at a time for lexical decision. They compared decision latencies for homophones that were preceded by the other member of the pair two trials earlier (e.g. ATE in the sequence EIGHT, DRUG, ATE) with latencies for nonhomophones of similar frequency (e.g. ROB in the sequence TRY, BIRD, ROB). They observed a significant facilitation effect for low frequency homophones and a significant interference effect for high frequency homophones. Davelaar et al. (1978) concluded that this was evidence for phonological recoding in lexical decision. The possibility that at least part of these effects are orthographic cannot be completely ruled out since most homophone pairs were visually similar (e.g. HORSE, HOARSE) and the nonhomophone pairs were not reported to be matched to the item two trials earlier for visual similarity. An analogous result was reported by Columbo (1986). She found no overall facilitation effect for words preceded by rhymes, but found that rhyme primes significantly facilitated decisions for low frequency targets and significantly interfered with decisions for high frequency targets. However, orthographic and phonological similarity were confounded in this experiment so it is impossible to determine which is responsible for the effects.

The results of the studies discussed above suggest that lexical decision latencies are affected by the phonological relationship between pairs of words. There is some evidence that for phonologically similar words, facilitation may occur for low frequency words and interference for high frequency words, although further research needs to examine whether the effects are truly phonological. The question remains, however, as to whether these effects indicate that phonological representations are used to access meaning or whether they reflect the use of phonology in the decision process after the access of meaning. One way to address this issue is to see whether it is possible to obtain semantic priming effects in the absence of phonological priming effects. Three lexical decision studies have appeared to have demonstrated such a result. Shulman et al. (1978) observed facilitation for both BRIBE-TRIBE and COUCH-TOUCH pairs relative to controls, indicating no phonological recoding, and also facilitation for semantically related pairs such as

STREET-ROAD relative to controls in an experiment with random letter nonwords and simultaneous presentation of strings. A problem with this study, however, is that evidence relating to phonological recoding was obtained from visually similar word pairs whereas evidence for semantic access was obtained from visually dissimilar pairs. Facilitation for the similarly spelled word pairs suggests that they may not have been processed to the same extent as dissimilarly spelled word pairs; perhaps subjects could base their decisions on orthographic information. The semantically related word pairs took approximately 40 ms longer to respond to than the orthographically similar pairs; it is possible that phonological information became available during this time.

This problem was avoided in an experiment by Martin and Jensen (1988) who preceded each target (e.g. SPOOL) with an orthographically dissimilar rhyme (e.g. RULE), an orthographically similar rhyme (e.g. FOOL), a semantically similar word (e.g. THREAD), and two control primes, an unrelated word (e.g. WALTZ) or a neutral string (e.g. XXXX). There was no facilitation relative to the controls for targets preceded by either type of rhyme, but there was a significant facilitation for words preceded by semantically related words. This suggests that the access of meaning is not phonologically mediated. However, the stimuli in this experiment were quite common; only 10% had frequencies less than 10, and so this does not necessarily indicate that all word meanings are accessed directly. Davidson (1986) presented sentence contexts word by word (e.g. ...she mixed up the ___), and then presented either a congruous word (e.g. DOUGH), its homophone (e.g. DOE), a control, or a pseudoword for lexical decision. Subjects produced significantly faster latencies for congruous words than for controls, but there was no facilitation for homophones of the congruous words. He concluded that context primes semantic but not phonological codes. Neither the words nor their frequencies were reported so it is unclear whether this pattern of results would be found with all stimuli.

Three other papers have examined phonological and semantic priming in tachistoscopic recognition tasks. The advantage of this technique is that it minimizes subjects' abilities to generate expectations about targets The prime can be presented so briefly that it cannot be reported, yet it still affects processing of the following word. Thus, it is quite unlikely that results are due to strategic effects that are found in lexical decision priming experiments For example, Tweedy, Lapinski, and Schvaneveldt (1977) have shown that the semantic priming effect in lexical decision is larger with a greater proportion of related trials Although tachistoscopic recognition does not require access to meaning, if a semantic priming effect can be demonstrated in the absence of phonological priming, this would suggest that the meanings of words can be activated before phonological information has time to accrue Evett and Humphreys (1981) observed a semantic priming effect in one experiment, and then failed to find a difference in the size of the facilitation effects for FILE-TILE and COUCH-TOUCH prime-target pairs in a second experiment using different stimuli. They concluded that phonological coding is not necessary for the access of meaning. Only 26.7% of targets had a frequency less than 10 and 60% had a frequency greater than 20 and so these results cannot be used to argue that phonological representations are never used to access meaning, since this may be true for high frequency words only. Further, the phonological similarity of the primes and targets may not have been sufficient. Humphreys et al. (1982) demonstrated that homophones were recognized more easily when preceded by the other member of the homophone pair (e g MAID-MADE) than when preceded by an orthographically similar words (e g MARK-MADE) or an unrelated word (e.g. SHIP-MADE). Almost half of the homophone targets (45 8%) had frequencies less than 10. No facilitation was found for words preceded by a pseudohomophones (e.g. MIAL-MILE), contrary to the lexical decision findings of Besner et al. (1985), and so they concluded that automatic access to phonology must occur after recognition. There are two aspects of their stimuli that may have produced this null finding; quite a few pseudoword primes were very unusual orthographically (e.g. KWYTE, BOYL, KUPH, NIPHE, TRYEL) and may not have been completely processed in the brief presentation, and a smaller percentage of targets were low in frequency (21.3% less than 10) In their final experiment, Humphreys et al. (1982) found a similar sized facilitation effect for regular and exception homophone targets They reasoned that if automatic access to phonology was computed using grapheme-phoneme correspondence rules, then in the case of exception word targets, the primes (e.g. PAIR) should have produced one pronunciation and the targets (e.g. PEAR) another (the regularized pronunciation, e.g PEER), and no facilitation should result. They concluded that automatic access to phonology occurs following recognition. This conclusion rests on the assumption that the computed phonological route could not produce the correct pronunciation of exception words. However, the Seidenberg and McClelland (1989) model does produce the correct pronunciations for exception words without addressing whole word representations. The phonological representations of high frequency exception words are computed as easily as those for high frequency regular words, and so no difference in priming effects would be expected.

In sum, the priming literature suggests that the meanings of high frequency words may become available before their phonological representations, but is not conclusive as to whether all, or most words are recognized on a visual basis. More work comparing performance on low and high frequency words needs to be done, with careful attention to separating phonological and orthographic effects. However, even if priming experiments reveal that phonological priming effects do occur without semantic priming effects for low frequency words, it would still not be clear whether phonological activation of meaning precedes visual activation of meaning in an unprimed situation. Monsell, Doyle, and Haggard (1989) have suggested that the phonological route may be too slow to have much effect on lexical decisions unless it has been primed.

Concurrent Articulation

Several researchers have examined the role of phonological recoding in word recognition by seeing whether performance deteriorates when the subject performs a concurrent task that is believed to disrupt phonological recoding. These researchers have assumed that phonological recoding involves the articulatory apparatus and so have used concurrent tasks of shadowing and articulatory suppression Shadowing involves repeating back a list of digits or message presented auditorily, and articulatory suppression consists of uttering an irrelevant phrase (e.g. hiya, hiya, hiya) over and over

Kleiman (1975) examined the effects of a concurrent shadowing task on subjects' abilities to make three types of decisions, graphemic (Do these two words look alike after the first letter? HEARD-BEARD, GRACE-PRICE), phonemic (Do these two words rhyme? TICKLE-PICKLE, LEMON-DEMON), and semantic (Are these two words synonyms? MOURN-GRIEVE, DEPART-COUPLE). He found that performance on phonemic decisions was much more impaired than for either graphemic or semantic decisions. Since synonym judgments were no more impaired than graphemic judgments, he concluded that subjects were able to retrieve the necessary information about the meanings of the words without using speech recoding.

Besner, Davies, and Daniels (1981) and Besner (1987) have argued, however, that the finding that rhyme judgments are impaired by shadowing or articulatory suppression does not necessarily imply that these concurrent tasks interfere with the phonological code used for the access of meaning. Rhyme judgments might be performed using a segmentation and deletion process carried out on a whole word phonological representation, and thus these concurrent tasks may interfere only with these post-recognition processes Similarly, Baddeley (1986) has pointed out that the rhyme and graphemic judgment tasks required subjects to hold the target word in memory, and that subjects making rhyme judgments may have been more likely to hold the target phonologically than when making other types of judgments.

In support of this view, several experiments have produced evidence of phonological recoding under concurrent articulation. Baddeley and Lewis (1981) found that concurrent articulation did not impair homophony judgments for words or nonwords, although Besner et al. (1981) found an effect in their error but not their reaction time data. These two groups of experimenters also found no effect of suppression on subjects' performance of a task in which they had to decide whether or not nonwords sounded like real words (e.g. PHOCKS). In addition, Besner and Davelaar (1982) found that an advantage of pseudohomophones over pseudowords in a serial recall task remained under suppression conditions even though effects of phonological similarity and word length disappeared. They concluded that there are at least two phonological codes, one used for meaning access and the second used by working memory to aid comprehension, and that only the latter is affected by concurrent articulation. Further, the interference it causes appears not to be due to specific phonemic interference (McCutchen & Perfetti, 1982) but rather to the additional general processing demands it requires (Waters, Komoda, & Arbuckle, 1985). Thus, since concurrent articulation does not seem to disrupt the derivation of the phonological code used for meaning access, the absence of an effect of suppression on semantic tasks such as synonym judgments does not reveal whether or not such tasks involve phonological recoding (Besner, 1987).

Summary of Lexical Decision Results

The large body of lexical decision experiments conducted to date has not yielded a clear answer as to whether or not access to meaning occurs via phonological representations, however several guidelines for further research have emerged from a review and analysis of the findings. The first is that studies must observe phonological effects on YES responses; experiments using pseudohomophones and nonwords varying in pronounceability may overestimate the influence of phonology since it may become available only for slower NO responses. Second, the studies, in particular the priming studies, strongly suggest that high and low frequency words need to be examined separately since phonological mediation may occur for low frequency words only. Phonological priming experiments, though, may overestimate the salience of the phonological code; on the other hand, a failure to observe phonological priming effects must be interpreted cautiously. Third, studies examining the ease of translation of a word into its phonological code will not be able to determine whether access to the meanings of high frequency words is phonologically mediated, since the translation process may simply be more efficient for these words. It is, however, a useful strategy to examine phonological recoding of low frequency words. Homophones, on the other hand, can be used to examine phonological recoding for both high and low frequency words. Fourth, evidence from concurrent articulation experiments is not relevant to the issue of the use of phonological information in the access of meaning since it appears to disrupt later memory processes instead.

A final point that emerges from these studies is that lexical decision is not a good task to use in order to study the access of meaning. Homophony and consistency studies have uncovered a serious problem with the lexical decision task, and that is that subjects' performance is very susceptible to the composition of the stimulus lists. The result is that different sorts of lists give different answers as to whether or not lexical access is phonologically mediated, and it is unclear what sort of list is appropriate in order to generalize results to actual reading. The likely explanation of list dependent effects is that lexical decision involves discriminating words from nonwords, and subjects set variable decision criteria depending on the difficulty of this discrimination on a particular list (Balota & Chumbly, 1984; Besner, Davelaar, Alcott, & Parry, 1984, Gordon, 1983, Seidenberg & McClelland, 1989; Waters & Seidenberg, 1985). According to this view, responses in lexical decision may be based solely on stimulus familiarity, and may not involve the access of meaning. Several studies have appeared to demonstrate that subjects do access the meanings of words when performing lexical decisions with pseudoword distractors. James (1975) demonstrated an effect of word concreteness on lexical decision latencies for low frequency words, and several studies have demonstrated effects of number of word meanings (Chumbly & Balota, 1984; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Rubenstein, Garfield, & Milliken, 1970), or other meaning variables (Balota & Chumbly, 1984; Whaley, 1978). However, Gernsbacher (1984) has argued that these effects may have occurred because concreteness and number of meanings were confounded with familiarity. In her experiments, there was no effect of concreteness or number of meanings when word familiarity was controlled. Thus, there is good reason to suspect that the lexical decision task does not require access to meaning. Baron (1985) has argued that "the lexical decision task is a laboratory model of no natural process. It does not help us learn how people pronounce words or how we extract their meaning for it requires neither" (p. 706). The more promising strategies that have been used with a lexical decision task, then, need to be applied to tasks that necessarily involve the access of meaning.

Phrase and Sentence Evaluation Studies

Another body of research has made use of a task that bears a more obvious relationship to the task of reading continuous prose for meaning than does the lexical decision task, and does not involve word-nonword discrimination. In a phrase evaluation task subjects are shown a short phrase or sentence and asked to indicate whether or not it makes sense. Some of the same research strategies have been used with the sentence evaluation task as were used with lexical decision, and these include the use of homophones, the use of words that vary in spelling-sound consistency, and priming.

Words. Baron (1973) compared response latencies to reject phrases that contained a homophone of a word that makes sense in the phrase, such as IN THE HAUL, with latencies to reject phrases such as NUT AND BOUT that contained an inappropriate word that was just visually similar to the appropriate one. He found no difference in decision latencies for the two types of phrases, and concluded that meaning could be accessed without phonological mediation. However, the error rate was significantly greater for phrases containing homophones and so this conclusion is not warranted. Baron also compared performance on the two appropriate phrases, IN THE HALL and NUT AND BOLT, and failed to find a difference in YES decision latency or accuracy. A problem with this study is that the phrases were each repeated 16 times, and so the items were highly familiar. This may have diminished the size of phonological effects. Treiman and Hirsh-Pasek (1983) avoided repeating sentences and also found significantly more errors on those containing homophones than on those containing visually matched controls, and no difference in decision latencies. Profoundly deaf subjects showed no homophone effects in either the latency or error data. However, Coltheart (1980) and Henderson (1982) argue that the finding of false positive errors on a phrase evaluation task for sentences containing homophones of the appropriate word does not necessarily indicate that a phonological representation was used to access meaning. Such errors could result instead from a phonological representation retrieved following recognition and used as an aid to comprehension.

Pseudowords. Doctor (1978, cited in Coltheart, 1980) attempted to distinguish between these two alternatives. She included all-word sentences such as Baron's and also included pairs of sentences that contained a pseudoword that was homophonic either to the appropriate word (e.g. WE WATE IN THE QUEUE) or to an inappropriate word (e.g WE GRAIT IN THE QUEUE). She argued that greater difficulty on pseudoword sentences that sounded correct would indicate that phonological recoding took place prior to recognition because this is the only way pseudowords can contact a semantic representation. Doctor found no differences in either rejection latencies or errors between the two types of pseudoword sentences, although she replicated Baron's (1973) finding of a higher error rate for all-word sentences containing homophones of the appropriate word. Unlike Baron, she also found an effect in YES decision accuracy. Sentences containing homophones produced significantly more errors than those that did not. Banks et al. (1981) also failed to find a difference in subjects' abilities to reject pseudoword sentences that did and did not sound correct in a study using homophonic and nonhomophonic pseudowords that were carefully matched for visual similarity. Both Coltheart (1980) and Banks et al. (1981) concluded that the access of meaning via phonology is not a significant process for skilled reading. Coltheart (1980) attributed the phonological effects found on all-word sentences to post-recognition phonology. Henderson (1982) has argued, however, that the way subjects process pseudowords may not necessarily indicate how they read real words.

Spelling-sound Consistency

Other researchers have attempted to determine whether or not the phonological representation of a word is used to access its meaning by comparing performance on all-word sentences that contain regular and exception words. The rationale for this approach is based on a hypothesis of dual-route models (e.g. Coltheart, 1978) that the correct phonological representation of an exception word can only be retrieved following access via the visual route; the phonological route applies grapheme-phoneme correspondence rules to the string and will produce a regularized pronunciation of exception words.

The experiments using exception and regular words have taken several different approaches. Treiman, Freyd, and Baron (1983) chose exception words (e.g. PLAID) that sound like another word (e.g. PLAYED) when incorrectly given a regular pronunciation.

They reasoned that if subjects use phonology to access meaning, they should have difficulty rejecting sentences such as THE CHILDREN PLAID IN THE STREET and accepting sentences such as HE WORE A PLAID SHIRT since PLAID will access the meaning corresponding to PLAYED. If the access of phonology occurs after recognition, subjects should have no more difficulty with those sentences than sentences containing matched regular words (e.g. HE WORE A PLAYED SHIRT; THE CHILDREN PLAYED IN THE STREET). They found that subjects made more errors on sentences containing an exception word than on regular word sentences, however this difference was small (8% vs. 7.1%), and while significant in the analysis by subjects, it only approached significance in the item analysis. There was a trend for latencies on exception word sentences to be longer, although it was not significant. One aspect of the design of this experiment may, however, have reduced the size of the effect seen in the exception word sentences. Subjects saw each word eight times in a single session, and so the items would have been quite familiar by the end. Another problem with this study is that the analyses were not conducted separately for correct and incorrect sentences.

A second approach, used by Coltheart, Laxton, Rikard, and Elton (1988), has been to choose exception words (e.g. NONE) that are homophonic to another word (e.g. NUN) when pronounced correctly. They reasoned that if subjects access the phonological representations of words after meaning, then sentences containing an exception word that is homophonic to an appropriate word (e.g. THE NONE SAYS HER PRAYERS) should be more difficult to reject than sentences containing a regular word (e.g. THE NINE WAS IN CHURCH). On the other hand, if subjects use phonology to access meaning, they should have no more difficulty rejecting sentences containing an exception homophone than control sentences, since the regularization of NONE will not sound any more like the appropriate word than NINE. They should, however, have more difficulty on sentences containing a pseudohomophone of the appropriate word than sentences containing a control nonword. Coltheart et al. (1988) concluded that they had found support for both pre- and post-recognition phonological recoding in skilled reading, since subjects made significantly more false positive errors on exception word homophone and pseudohomophone sentences that sounded correct than on control sentences. Neither group showed any effects of homophony on YES responses.

Coltheart et al.'s (1988) nonword results conflict with the findings of Doctor (1978, cited in Coltheart, 1980) and Banks et al. (1981) who failed to find any differences between nonword sentences that did and did not sound correct. However, the difference in false positive error rates on homophonic nonwords and controls in their study was quite small (4% vs. 1%), and it is not clear whether it was significant by items. Thus, the effect with nonwords does not appear to be very robust, and as pointed out earlier, may be of limited value anyway in determining how words are processed. There is also a problem with Coltheart et al.'s (1988) logic for the all-word sentences. Their conclusion that subjects' difficulty with homophonic exception word sentences is evidence for post-recognition retrieval of phonology depends on the assumption that the correct pronunciation of exception words cannot be produced prior to recognition. This is true if the translation from spelling to sound is through the use of grapheme-phoneme correspondence rules as posited by Coltheart (1978) but not via the system developed by Seidenberg and McClelland (1989). Thus, the findings in this experiment do not settle the issue of whether or not phonological representations are used to access the meanings of words. Treiman et al.'s (1983) experiment is not subject to this criticism since it depends on subjects producing the incorrect pronunciation of the exception word (e.g. PLAYED for PLAID) which, according to both theories, can only be done if the phonological representation is derived prior to the access of meaning.

Waters et al. (1984) used a slightly different procedure in their sentence verification task. Sentences were presented on the screen without the last word and subjects read them silently. Upon completion, they pressed a button and the final word, which was either an exception word or a regular word. was presented. Subjects had to decide as quickly as possible whether the word made sense given the preceding context. They found that low frequency exception words produced significantly slower decision latencies and more errors than matched regular words, but there was no exception effect for high frequency words. The words used in this experiment were the same regular and exception words used in Waters and Seidenberg's (1985) lexical decision experiment that did not produce a low frequency exception effect. These results imply that phonological information does influence the recognition of low frequency words by skilled readers Seidenberg (1985) suggested that these results should be interpreted with caution, however, since decision times were relatively long and it is possible that phonological information was activated during the interval the decision was made. While such a technique permits the measurement of decisions upon presentation of the word of interest, care must be taken to ensure that words in each group are equally predictable given their sentence frames.

Priming

Treiman et al. (1983) examined whether the interference effects for similarly spelled but differently pronounced words (COUCH-TOUCH) found in the lexical decision task (Meyer et al., 1974; Hanson & Fowler, 1987; Shulman et al., 1978) are also found in a sentence reading task. They used a slight variation of the sentence verification task; subjects were shown sentence fragments and then were given a choice of two words and they were required to indicate which word correctly completed the sentence. The critical comparisons were between sentences such as HE MADE A NASTY HASTY REMARK and its controls, and sentences such as I WILL NEVER SEVER OUR RELATIONSHIP and its controls. They found a significant interference effect for sentences containing similarly spelled but differently pronounced word pairs, which they argue indicates that subjects use phonological codes in reading sentences.

Several researchers (Baddeley & Hitch, 1974; Baddeley & Lewis, 1981; McCutchen & Perfetti, 1982; Treiman & Hirsh-Pasek, 1983) have demonstrated that tongue twister sentences (e.g. SHE CHOSE THREE SHOWS TO SEE AT THE THEATRE) take longer to read than semantically equivalent sentences (e.g. SHE PICKED TWO MOVIES TO SEE WITH HER FRIEND). However orthographic and phonological similarity are confounded in tongue twister sentences. Baddeley and Lewis (1981) have demonstrated that verification time for their phonologically similar sentences was significantly correlated with their orthographic similarity. On the other hand, McCutchen and Perfetti (1982) argue that visual similarity was not responsible for their effect, and suggest that these results indicate automatic activation of phonological information, although not necessarily prior to recognition.

In sum, results from sentence evaluation experiments have suggested that phonological information is used in performing the task, however it is difficult to determine from this research whether skilled readers ever access the meanings of words based on their phonological representations, or whether the phonological information is accessed after meaning and used in sentence comprehension. The most promising way to answer this question using a sentence evaluation task seems to be the one used by Treiman et al. (1983). However, only a demonstration that subjects have greater difficulty on Treiman et al.'s type of exception word sentences would be informative in this debate. No difference between exception and control sentences could occur because subjects could correctly derive the pronunciations of the exception words prior to the access of meaning or because they retrieved the pronunciations following the access of meaning. It is also possible that no difference will be evident because the sentence evaluation task may not be a very sensitive measure. Subjects may be able to resolve phonological confusions before they have finished reading the sentence.

Acquired Dyslexia Studies

Another line of evidence concerning the role of phonological information in word recognition comes from neuropsychological studies of adults whose ability to read has been impaired due to brain damage from accidents or strokes. The strongest argument that could be made from patient data would be that phonological information plays no role, and this could be made by the existence of a patient who showed good comprehension of words in silent reading but could not derive phonological representations prior to the access of meaning. It is more difficult to demonstrate that phonological recoding does play a role in word recognition using data from patients since impaired comprehension in the presence of poor phonological processing could either be caused by the phonological deficit, or both could be caused by a third impairment or two independent impairments

There have been several reported cases of patients who show an impaired ability to derive phonological representations, but who nonetheless appear to comprehend written stimuli well (Dérouesné & Beauvois, 1985; Patterson, 1982; Shallice & Warrington, 1980). These patients have been called phonological dyslexics, although the term was coined for patients who show good word naming and poor pseudoword naming without reference their ability to understand written words (Beauvois & Dérouesné, 1979; Dérouesné & Beauvois, 1979). Patients with the ability to name words cannot be assumed to have understood them since in some views (Bub, Cancelliere, & Kertesz, 1985; Funnell, 1983; Morton & Patterson, 1980; Schwartz, Saffran, & Marin, 1980) there is an additional word specific processing route from orthography to output phonology that is independent of the semantic system. In order to make statements about the role of phonology in the access of meaning, then, written word comprehension must be specifically tested.

Phonological dyslexics rarely show a complete lack of ability to derive phonological information, usually assumed to be an inability to name pseudowords that is not due to other problems such as an inability to produce the spoken forms of stimuli or an inability to hold items in working memory. The five patients tested by Dérouesné & Beauvois (1979, 1985; Beauvois & Dérouesné, 1979) and B.T.T. (Shallice & Warrington, 1980) were able to read half of the pseudowords from at least one list, and thus these cases thus do not allow strong conclusions about the role of phonology in word recognition.

More interesting for the present purposes are patients who show very poor oral reading of pseudowords. Of these, A. M. (Patterson, 1982) has been the most extensively tested. He read only 8% of pseudowords in the same session that he read 95% of a list of nouns, although he later read as many as 26% of nonhomophonic pseudowords. He had no more difficulty rejecting pseudohomophones than pseudowords on a lexical decision task, which suggests he was not accessing their phonological representations. On the other hand, given a printed pseudoword, A. M. was able to correctly choose its equivalent from a set of three spoken by the experimenter on 79% of trials in which the alternatives were distinct (e.g. FLEB, TREAN, MIDE) and on 60% of the trials in which the alternatives were phonologically similar. Patterson (1982) argued, though, that since "A. M. 's ability to assemble a phonological code was so minimal, so slow, so hesitant, and so error-prone that it is exceedingly unlikely to have played any role in his confident and accurate reading of real words" (p. 94).

However, pseudoword reading tests and untimed lexical decision tasks may underestimate a patient's phonological recoding abilities. It is possible that a phonological representation may be produced and facilitate word recognition even if patients are unable to consciously use it to derive the pronunciations of pseudowords. Recently, Hildebrandt, Sokol, Dymkowski, and Ruzecki (1990) have demonstrated that a deep dyslexic patient (see Coltheart, Patterson, & Marshall, 1980) who could read only 9/232 pseudowords correctly nonetheless still showed effects of spelling-sound regularity in timed naming and lexical decision tasks. This study reaffirms Henderson's (1981) view that both latency and accuracy measures need to be examined in order to be able to relate the performance of patients and normal subjects.

Even the most impaired patients, then, may have a minimal amount of phonological information contributing to the activation of word meanings, and so these patients do not provide conclusive evidence for the strong hypothesis that reading can occur without phonology. However, if we assume that very little phonological information is available to aid word recognition in patients such as A. M., and they demonstrate normal comprehension of written words, then we could conclude that phonological information plays, at best, a minor role in word recognition. But the written word comprehension ability of A. M. was not entirely normal despite intact auditory comprehension. He made no more errors than normals on lexical decision tasks that included low frequency words or function words, and performed only slightly poorer on suffixed words (no latency measure was reported), but this does not necessarily mean that the words were fully comprehended. Comprehension needs to be demonstrated in other tasks as well because it is possible that lexical decisions could be performed on the basis of orthographic familiarity or partially activated semantic representations. Patterson found that A. M. did have difficulties comprehending some function words and may have problems with derivational forms. His comprehension of abstract words on word-picture matching and synonym judgment tests was good, although it is unclear how broad a range of word classes and frequencies were included in these tests. Dérouesné and Beauvois' (1985) patient L. B., who was better at reading pseudowords than A. M., showed a similar pattern. He scored normally on a test of content word comprehension but had difficulty understanding verb inflections and function words.

Since comprehension deficits appear to accompany poor phonological recoding ability, the conclusions that can be made about the role of phonological recoding are weaker. It suggests that phonological recoding may play a role in the recognition of some words, although one must be extremely cautious and not infer from the correlation of phonological and comprehension deficits in a few patients that intact phonological processing is necessary for understanding such words. Patterson (1982) acknowledges that this is a problem but she presents a variety of other arguments in support of the view that phonological recoding is necessary in order to read grammatical morphemes. In addition, she has argued that patients who show normal reading of single-morpheme content words provide convincing evidence that the recognition of these words does not require phonological recoding (Patterson, 1981, 1982). Allport (1979) has also suggested that the comprehension of grammatical morphemes but not content words depends on phonological recoding.

Since at least some phonological information may be contributing to the activation of a word in patients such as A. M., it cannot be determined whether or not phonological information is required in order to access the meanings of content words. However, given the impairment of patients such as A. M., if it were playing an important role in word recognition, some deficit in reading these words would be expected. A possibility is that phonological coding facilitates quick and accurate identification of words, particularly those that are not very familiar. When testing patients, researchers seldom collect latency measures although often report that patients respond slower than normal subjects. A model that accommodates this role for phonological information is the Seidenberg and McClelland (1989) model. Presentation of a stimulus will activate both a phonological representation and a semantic representation. Activation will also spread from phonological units to semantic units In the case of a high frequency word, activation of the semantic units will be rapid and strong before activation from phonological units reaches them. On the other hand, when a low frequency word is presented, the activation of the semantic units will be slower and weaker, and in this case the additional activation from phonological units may allow a clear pattern of activation to emerge on the semantic units faster. The extent to which phonological information contributes to the activation of the meaning of a word will depend on the strength of the connections between the orthography and the semantic representation of the word. This model can be used to simulate word recognition by dyslexic patients, for example, by deleting some of the processing units (Patterson, Seidenberg, & McClelland, 1987), but as of yet a simulation of phonological dyslexia has not been attempted.

In sum, studies of neurological patients suggest that it may be possible to access the meanings of some words with minimal phonological information available, however, this does not necessarily mean that skilled silent reading makes no use of phonological information since it is possible that skilled readers use phonological information for rapid, accurate access to meaning. It is also possible, however, that the way neurological patients perform is by means of completely different strategies than that those used by intact skilled readers (Henderson, 1981; Seidenberg, 1988). Thus generalizations from data on brain damaged patients to skilled reading must be made very cautiously.

Semantic Decision Studies

Evidence to this point for the use of phonology in the access of meanings of words is not very strong. However, the three types of techniques used to examine the question that have been reviewed so far have all been problematic. Lexical decision may be performed without access to meaning, effects found in most sentence verification experiments may reflect processes occurring after the access of meaning and the technique may not be sensitive enough to examine single word reading, and there are problems generalizing from neuropsychological patients to skilled readers. A final technique avoids difficulties of the previous methods. In the semantic decision task, subjects must decide whether or not a word belongs to a previously specified semantic category. Thus this task more directly measures the access of meaning than the others discussed. It does not involve the word-nonword discriminations of the lexical decision task, it allows measurement on single words, it is used with skilled readers, and data can be collected on YES responses.

There have been only a few studies examining the use of phonological information in semantic decisions, and most of these have made use of homophones. In other studies, Green and Shallice (1976) found no effect of word length on either YES or NO semantic decision times using categories such as SPORT and OCCUPATION, and Klapp et al. (1973) failed to find an effect of number of syllables on YES or NO decision latencies using the categories ANIMAL and OBJECT. Monsell et al. (1989) found that neither length in letters nor syllables were significant predictors of PERSON WORD/THING WORD classification time. Lupker and Williams (1989) did not find faster REAL/ARTIFICIAL decision latencies for words (e.g. TRUCK) when the word on the previous trial rhymed (e.g. DUCK). These experiments, then, provide no evidence that phonological information plays a role in reading for meaning. However, lexical decision studies examining length and priming effects suggest that these effects may occur for low frequency stimuli only.

Hillinger and James (1977) used an ambiguity decision task in which subjects had to decide whether or not words had more than one meaning, and observed that latencies for homographs (e.g. DOVE, WIND) were longer than for words also having two meanings, but only a single pronunciation (e.g. BANK, FAIR). They argued that this result would not be expected if access to meaning was based solely on a visual representation.

The experiments with homophones also provide evidence for the use of phonological information. Meyer and Ruddy (1973) and Meyer and Gutschera (1975) reported that NO responses to the question "Is this word the name of a fruit?" were slower when the test word was a homophone of a fruit name (e.g. PAIR) than when it was an unrelated homophone (e.g. TAIL). A problem with these studies is that spelling controls were not included and so it is unclear whether these effects are phonological or orthographic. Ellison (1975, cited in McCusker et al., 1981) conducted a study that did include spelling controls. The decision that subjects made was whether or not a word was an animal, and targets were either homophone foils (e.g. BARE), orthographically similar (e.g. BEAT), phonologically similar (e.g. CARE), or unrelated (e.g. TREE). Homophone foils took significantly longer to reject than rhymes, which were slower than orthographically similar words, and the unrelated words were rejected the fastest. Since there was some interference for orthographically similar words relative to unrelated words, the interference effect for homophones is not entirely due to phonological recoding. Thus, homophones must be evaluated with respect to spelling controls in order to determine the extent of phonological recoding. Banks et al. (1981) used an ANIMAL decision task, and found that subjects made significantly more errors on homophone foils than on spelling controls but there were no rejection latency differences. In an experiment using the category BODY PARTS, homophones took slightly longer to reject and produced more errors than spelling controls but none of the differences were significant. Only four subjects participated in the experiment, however. In another experiment, Banks et al. (1981) demonstrated that effects for homophones were due to the fact that they sounded like a member of the category and not to some peculiar property of homophones. They had subjects make both ANIMAL and BODY PART decisions using the same set of words, and found a significant interaction between the type of homophone foil and type of decision. Decision latencies were longer for homophones when they were foils for a category (e.g. DEAR for the category ANIMAL) than when they were unrelated (e.g. DEER in the category BODY PART). A smaller, but also significant, interaction was found for visually similar nonwords; that is

DEEB took longer to reject when the category was ANIMAL than when the category was BODY PART. Error rates for homophones of category members did not differ from spelling controls These results suggest that both phonological and visual similarity can cause interference in a categorization task. The size of the homophone effect is probably underestimated in this experiment because spelling controls were more similar visually to targets than were homophones.

Coltheart (1980) points out that studies such as these do not distinguish between phonological information used to access meaning and post-recognition phonology. He cited an unpublished study by Midgley-West (1978) who attempted to address this question by comparing performance for regular and exception homophone foils. According to Coltheart, the phonological representation of exception words derived prior to recognition is incorrect, and thus any phonological effects on this task for exception words must be due to post-recognition phonology. There was no difference between regular and exception words for either NO responses to homophone foils or YES responses to targets, and so he concluded that there was no evidence that a phonological code is used to access meaning. However, as pointed out earlier, there is a possibility that the correct phonological representations of exception words are calculated prior to the access of meaning (Seidenberg & McClelland, 1989) and so this is not conclusive evidence that meaning is accessed on a visual basis.

Coltheart (1978; Coltheart et al., 1977) has suggested that the categorization task could be performed either by searching members of the specified category for a match to the input or by interrogating the semantic representation accessed by the target. He argues that only if the latter procedure is used would the semantic categorization task be useful to examine codes used in reading for meaning. Advance category information is not given in normal reading so a category search strategy is not possible. In addition, if a search procedure were used, NO decisions could only be made after all members of the category were searched and this means that NO responses would necessarily be slower than YES responses. Thus, slower phonological coding may not be available to influence YES responses but may influence NO responses. The reports of longer rejection latencies for homophone foils, then, would not necessarily indicate that phonological information is used to access the meanings of words. Coltheart (1978) thoroughly analyzed the data from Meyer and Gutschera's (1975) experiment and concluded that a search model was not adequate to explain their results, and instead advocated the view that subjects interrogated the semantic representation accessed by the target. From the low (15%) false positive error rate for homophone foils, Coltheart (1978) argued that decisions based on phonological access do not occur very often. He suggested that if a decision is based on phonological access of meaning when a homophone foil such as PAIR is presented, subjects would be more likely to produce a false positive error than a correct response since the PEAR meaning would be primed from the category name FRUIT. Thus, phonological access of meaning probably occurs less than 30% of the time.

Recently, Van Orden (1987; Van Orden, Johnston, & Hale, 1988) performed a series of semantic categorization experiments and concluded that access of meaning is always by means of a phonological code. To avoid difficulties of making conclusions based on NO responses, they compared false positive YES responses to homophone foils and matched spelling controls. Van Orden (1987) found that subjects made significantly more errors on homophone foils than on spelling controls, and that this effect was larger for foils that are spelled very similarly to the actual category member (e.g. MEAT-MEET) This similarity effect disappeared when words were presented briefly and then pattern masked Van Orden argued that these findings supported the phonological mediation hypothesis, and provide evidence against the view that the influence of phonological codes on word recognition is delayed relative to the influence of orthographic codes. He proposed that a verification type model (e.g. Rubenstein et al., 1971; Becker, 1976, 1980; Schvaneveldt & McDonald, 1981; Paap et al., 1982) could best explain his results. In his view, lexical entries are activated by their phonological representations and then a spelling check is conducted on these entries in order of their activation levels until the orthographic representation of one matches that of the stimulus word. A word is more likely to be a false candidate for the verification procedure if it is phonologically similar to the target, and a candidate is more likely to slip by the verification procedure if it is orthographically similar to the target. Thus, the particularly high error rates on similarly spelled homophone foils. Masking interrupts processing before the spelling check occurs, eliminating the effect of orthographic similarity.

In a third experiment, Van Orden tested this theory against a dual route account in which access is direct for common words and phonologically mediated for less common words (Andrews, 1982; Seidenberg et al., 1984). He argued that the dual route account predicts that the likelihood of phonological confusions would depend on the frequency of the stimulus the subject sees, that is, the homophone foil,whereas the verification account predicts that the likelihood of false positive errors depends on the frequency of the homophone that is the exemplar of the category. A false candidate it is less likely to pass the orthographic check if knowledge of its spelling is complete and readily available. He found an effect of exemplar frequency on false positive errors but did not find an effect of homophone foil frequency, and concluded that that the verification theory was supported. In a second paper, Van Orden et al. (1988) attempted to determine whether the phonological code responsible for the effects found in the first paper was computed prior to access or retrieved afterwards. They compared false positive error rates for both word and pseudoword homophone foils and their respective controls. Both types of foils produced significantly more errors than spelling controls, and error rates to the two types of foils did not differ. In addition, YES response latencies that were false positive errors (e.g. body part-HARE) were not significantly different than YES response latencies for control exemplars (e.g. body part-TOOTH) either for word or nonword foils. Van Orden et al (1988) concluded that computed phonology is available rapidly enough to influence positive responses in a categorization task, and that there was no evidence of an influence of an independent direct access route.

However, if Coltheart's (1978) analysis is applied to Van Orden's results, then he did find evidence for direct access. In the experiment where stimuli were pattern masked, and in his view the spelling check was prevented, the false positive error rate was 40% for similarly spelled homophones and 46% for less similarly spelled homophones. A rate of 50% would be expected if both homophones were available and subjects chose randomly between them. But according to Coltheart, the false positive error rate would be expected to be greater than 50% if the category name primes the exemplar and makes it more likely to be the basis of the response. This priming is likely given that some of the categories had a small number of exemplars (e.g. PART OF A HORSE'S HARNESS). At least some of the time, then, access appears to be via a direct visual route.

If category names do prime phonological representations of exemplars, then at least some false positive errors may be due to subjects responding on the basis of a match between the incoming phonological representation and the primed representation, rather than on the basis of a semantic representation activated by phonology. This would mean that false positive error data do not give an accurate picture of the extent to which phonology is used to access meaning. If priming were partially responsible for the increased false positive errors on homophone foils, then one might expect to find an influence of the strength of the association between a category name and the exemplar on false positive errors to the homophone foils, since several studies have found production frequency effects on YES categorization responses (Balota & Chumbly, 1984; Green &

Shallice, 1976; Rosch, 1975). Van Orden (1987) collected production frequency norms for each of the exemplars used in his third experiment by giving a group of subjects the category names and asking them to list as many exemplars as possible in 30 seconds. There was no correlation between the number of times an exemplar was produced and the rate of false positive errors on homophone foils. In addition, Van Orden et al. (1988) did not observe effects of either production frequency or typicality on false positive error rates. However, the range of values on the production frequency and typicality measures may not have been very large, and small stimulus sets were used so it is possible that these correlations did not detect the influence of these measures. One way to avoid possible priming effects is to use very broad category names such as LIVING THING and OBJECT that would be much less likely to strongly prime any particular word. Monsell et al. (1989) advocated the use of such categories to minimize typicality effects and post-recognition processing.

Van Orden (1987) did not observe an effect of the frequency of the homophone foils on false positive errors, which he interpreted as evidence against the view that high frequency words are recognized directly on a visual basis. One reason why he may not have found a lower error rate for high frequency homophone foils is that effects of foil frequency may have been cancelled out by effects of exemplar frequency. There was a confound in his stimuli in that the low frequency foils had primarily high frequency exemplars (which were associated with fewer errors) and the high frequency foils had primarily low frequency exemplars (which were associated with more errors). To examine the role of homophone foil frequency, it and exemplar frequency need to be manipulated factorially. A second reason for his failure to find an effect of foil frequency is that spelling controls were not used in his experiment; the controls he used were neither phonologically similar nor visually similar to the exemplar. Thus, errors on homophone foils could have been due either to visual or phonological similarity to the exemplar, and it is possible that errors on high frequency foils were due only to their visual similarity to the exemplar whereas errors on low frequency foils were due to both visual and phonological similarity to the foil. In order to attribute effects to phonological processes, then, it is imperative that performance on homophone foils be compared to that on spelling controls

In sum, research using the semantic decision task has provided stronger evidence that phonological information is used to access meanings than has previous research. This is important, because the task comes closer than any others in directly measuring the access of meaning of single words by skilled readers. However, a potential problem with the task was noted, and that is that performance may be affected by advance information from the category names. Thus, before Van Orden's conclusion that access to meaning is always phonologically mediated can be accepted, this possibility needs to be explored In addition, since the review of the literature uncovered very little other evidence that skilled readers use phonological information to access the meanings of high frequency words, performance on these words needs to be examined carefully with reference to an orthographic control. From the review of the lexical decision literature, it became evident that homophones are the best stimuli to use in examining phonological effects on high frequency words, since an absence of an effect of ease of translation (number of letters, number of syllables, spelling-sound consistency) may occur because the spelling-sound translation process is more efficient for these words.

The studies that follow compare responses for homophone and nonhomophone stimuli on the semantic decision task. In the first study, the frequency of the homophone foil and the frequency of the exemplar are factorially manipulated, and performance on the foils is compared to that on spelling controls, in order to determine whether evidence for phonological mediation of the meaning of high frequency words occurs using the same procedure as Van Orden (1987). In the second study, broad categories are used to examine whether the results obtained in Van Orden's work and in the first experiment are due to priming from the category name. The third experiment explores whether the use of phonological information is strategic by discouraging subjects from using phonology, and the fourth experiment examines whether foils must not only sound like but also look like the exemplar in order to produce more false positive errors. The results of these four experiments suggest that the effects Van Orden reported are more limited than his studies suggest. Two further studies examine phonological effects on YES responses to exemplars.

EXPERIMENTS WITH HOMOPHONE FOIL TARGETS

Experiment 1:

A Replication and Extension of Van Orden's Studies

The goal of the first experiment was to examine whether it was possible to replicate Van Orden's (1987; Van Orden et al., 1988) finding that subjects make more false positive categorization errors when stimuli are homophones of a category exemplar than when they are spelled similarly to a category exemplar. In addition, the study examined whether these effects depend on the frequencies of the exemplar and foil Van Orden (1987, Expt 3) found an effect of exemplar frequency but not foil frequency, however, foil and exemplar frequency were confounded in the study and control words were not matched to foils for orthographic similarity to the exemplar. Two levels of exemplar frequency (high and low) were crossed with three levels of foil frequency (high, low, and pseudohomophone), which produced six groups (see Table 1).

Two other category-target conditions were included in addition to the homophone foil (e.g. car part-BREAK) and the spelling control (e.g. car part-BRAVE) conditions mentioned above. The exemplar (e.g. car part-BRAKE) was included in order to be able to compare false positive response latencies on homophone foils to YES response latencies for the actual exemplars. Van Orden et al. (1988) argued that their failure to find faster YES latencies for exemplars was evidence that the orthographic representations of exemplars did not contribute to decision latencies, since false YES decision latencies to foils could not have benefitted from orthographic information. They compared false positive response latencies for homophone foils to a mean YES latency for six other exemplars of the category in their first experiment, and to a single other exemplar in their second experiment. The actual exemplar is a better comparison since it controls for variables such as typicality and production frequency that affect the ease of making the

Table 1 An illustration of the conditions used in Experiment 1.

Group		Category-Target Relation			
Foil Freq.	Exemplar Freq.	Exemplar	Foil	Spelling Control	Homophone Control
High	High	male relative-SON	male relative-SUN	male relative-SIN	air vehicle-SUN
High	Low	car part-BRAKE	car part-BREAK	car part-BRAVE	painter's equipment-BREAK
Low	High	child's toy-BALL	child's toy-BAWL	child's toy-BAIL	construction material-BAWL
Low	Low	parasite-FLEA	parasite-FLEE	parasite-FLEX	award-FLEE
Non	High	footwear-SHOES	footwear-SHEWS	footwear-SHOSS	water vehicle-SHEWS
Non	Low	canine-FOX	canine-FOCKS	canine-FOW	fastener-FOCKS

category decision, and it is also a homophone.

In a fourth condition, subjects had to decide whether the homophone foil was a member of a completely unrelated category (e.g. painter's equipment-BREAK). This allowed the comparison of rejection times and errors for homophones when they sounded like they belonged to the category and when they did not. Banks et al. (1981) included a similar condition to control for the possibility that homophones could take longer to reject and be more prone to errors than spelling controls because they are represented differently from nonhomophones, and not because they sound like a member of the category. This homophone control condition was not included in Van Orden's experiments.

Thus, there are several comparisons of interest in the experiment. The most important of these are the comparisons of the false positive error rate on homophone foils with that on spelling controls and on homophone controls. If more false positive errors are made on homophone foils than on either spelling controls or homophone controls, it would provide evidence that access to the meanings of words is phonologically mediated. Correct NO response latencies for homophone foils were compared to those for spelling controls and for homophone controls. Van Orden (1987) and Van Orden et al. (1988) argued that similar NO latency distributions provide evidence that all words undergo a spelling check, not just homophones. However, his theory suggests that on some occasions, the wrong member of the homophone pair will be selected for the spelling check first, and a second spelling check subsequently performed on the other member. In these cases, NO latencies for homophones should be longer than for spelling controls. A problem with NO latencies, though, is that they may include further extended semantic processing that is not done when a target is a member of the category, so the evidence from NO latencies is much less important than evidence from false positive errors. And finally, false positive response latencies on homophone foils were compared to correct YES latencies on exemplars to examine Van Orden et al.'s (1988) claim that exemplars do not benefit from having orthographic information consistent with the YES response. However, there is also a problem making inferences from this comparison, since correct responses on exemplars may occur for trials on which subjects finish processing the target before responding, whereas incorrect responses on foils may occur when subjects respond before processing has been completed. Thus, these results are also less important than the false positive error data. The effect of the frequency of the exemplar and the frequency of the foil on the magnitude of these effects was also examined. Van Orden's (1987) view predicts that phonological effects will occur for both high and low frequency words, whereas the dual route view predicts phonological effects for low frequency words only. A comparison of the size of phonological effects for the word groups and the pseudoword groups will provide evidence as to whether the effects arise prior to or after the access of meaning.

The inclusion of the exemplar and homophone control conditions required a change in design from Van Orden's experiments. He presented all stimuli in a single session. In order to avoid intra-list repetition effects, the present experiment was conducted in four sessions separated by at least one week. Only one member of a stimulus quadruple appeared in each session For example, the trials car part-BRAKE (exemplar), car part-BREAK (homophone foil), car part-BRAVE (spelling control), and painter's equipment-BREAK (homophone control) all appeared in different sessions. As in Van Orden's experiments, a large number of filler trials were included, since several researchers have suggested that subjects change their processing strategies when a large number of homophones are included in an experiment (Davelaar et al., 1978; Hawkins et al., 1976; McQuade, 1981).

Method

Subjects. Twelve McGill University undergraduates were paid $10 each to participate in the study. All were native speakers of English.

Stimuli. There were 288 experimental trials and 720 filler trials. The experimental trials consisted of 72 quadruples (see Appendix A). The first step in the construction of the quadruples was to choose 72 pairs of homophones. In order to be able to use the same stimuli with broad category names in Experiment 2, the homophone pairs were selected so that one member of each pair -the exemplar- was either a living thing or an object, and the other -the foil- was neither. Half of the exemplars were living things and half were objects. The 72 homophone pairs were chosen so that they fell into six groups of 12 pairs each. The six types of groups were produced by factorially manipulating the frequency of the exemplar (high, low) and the frequency of the foil (high, low, pseudohomophone). The mean frequencies of the exemplars and homophone foils for each of the six groups are presented in Table 2.

The second step in the construction of the stimulus quadruples was to choose a spelling control (e.g. car part-BRAVE) for each of the 72 homophone foils. The spelling controls needed to be as similar orthographically to the exemplars as the homophone foils were to the exemplars. To accomplish this, the orthographic similarity of each foil to its exemplar (e.g. BREAK/BRAKE) was calculated using Weber's (1970) graphic similarity measure (used by Van Orden, 1987; Van Orden et al., 1988), and then another stimulus was chosen that was as similar as possible to the exemplar (e.g. BRAVE/BRAKE) on this measure, and was similar to the foil in frequency. The mean similarity scores of the foils to the exemplars and of the spelling controls to the exemplars for each of the six groups are presented in Table 2, along with the mean frequencies of the spelling controls

In the final step, a category name was chosen for each of the 72 exemplars (e.g. car part-BRAKE). These category names were also used for the matched homophone foils (e.g. car part-BREAK) and spelling controls (e.g. car part-BRAVE). The fourth member of each quadruple, the homophone control, consisted of the homophone foil and an unrelated category name (e.g. painter's equipment-BREAK) To create these unrelated categories, the category names within each of the six groups were shuffled (e.g. car part-

Table 2: Mean word frequency and mean orthographic similarity to exemplar for the stimuli in the six experimental conditions in Experiments 1-3.

	Word Frequency			**Similarity to Exemplar**	
Group	**Exemplar**	**Foil**	**Spelling Control**	**Foil**	**Spelling Control**
HF Foil/ HF Exemplar	85.3	83.4	76.5	.63	.64
HF Foil/ LF Exemplar	7.3	86.4	88.4	.62	.63
LF Foil/ HF Exemplar	92.1	6.9	5.9	.66	.70
LF Foil/ LF Exemplar	5.0	4.5	4.5	.65	.66
PW Foil/ HF Exemplar	83.8	-	-	.66	.67
PW Foil/ LF Exemplar	6.8	-	-	.67	.68

Note Word frequency was calculated using the Kucera and Francis (1967) norms and orthographic similarity was calculated using Weber's (1970) measure. HF = high frequency; LF = low frequency; PW = pseudoword The statistics for the homophone control group are the same as for the foil group since the only difference in these conditions was in the preceding category name.

CELLAR, store personnel-LATTER, painter's equipment-BREAK).

Four lists containing 72 experimental trials were created, with each member of a quadruple on a different list. Eighteen items from each of the category-target conditions (exemplar, homophone foil, spelling control, and homophone control) appeared on each list. These 18 consisted of three from each of the six exemplar frequency/foil frequency groups. The homophone controls were placed on lists such that no category name appeared more than once on a list. The number of experimental YES trials on each list was 18 (the exemplars), and the number of experimental NO trials on each list was 54 (18 homophone foils, 18 spelling controls, and 18 homophone controls). The homophonic experimental stimuli were 21.5% of the stimuli on the entire list (16.7% of the list were word homophones and 4.8% were pseudohomophones); one third of these were foil trials.

In addition to the experimental stimuli, 720 filler trials were included, 180 on each list. One hundred and eighty categories were created, and each category appeared once on each list. (Half of these were living thing categories and half were object categories as required by Experiment 2). Four words were chosen for each category, some of them were exemplars and some were not exemplars. Of the 180 filler words on each list, 108 were exemplars of their categories and 72 were not exemplars. This ensured that across each entire list (experimental trials plus fillers) there were the same number (126) of YES and NO trials. The NO filler words were from a variety of grammatical classes so that the homophone foils, which were not all nouns, would not be unusual. (The NO fillers were chosen so that the correct response would still be NO when the categories were changed to LIVING THING and OBJECT in Experiment 2).

In sum, there were four lists, each having an equal number of YES and NO trials. The same 252 category names appeared on each list; the target appearing with a given category name was different on each list. The only targets that appeared twice were the 72 homophone foils that appeared with two different category names. The order of presentation was randomized on each list such that no more than three YES or NO trials appeared in succession.

An additional 16 categories and 16 stimuli were chosen for a practice list. Half of the stimuli were exemplars of their categories and half were not.

Procedure The subjects completed four experimental sessions, each lasting 30 minutes. The sessions were separated by at least one week. In a session, the subjects first saw the 16 practice trials and then the 252 experimental trials from one of the lists Subjects were given feedback on each practice trial to ensure that they understood the task. The order of presentation of the experimental lists was counterbalanced across subjects such that each list was presented three times in each of the four session positions. No subject saw any of the experimental lists twice.

On each trial the category name appeared on a computer screen for 2 s, a fixation point (*) appeared for the next 500 ms, and then the target stimulus appeared and remained until the subject responded Subjects were instructed to indicate whether or not the target stimulus was a member of the indicated category by depressing one of two telegraph keys as quickly as possible The left key was used to indicate NO responses and the right key was used to indicate YES responses. The intertrial interval was 1.5 s

Stimuli were presented in lower case letters in the centre of an IBM monitor (model 5154) attached to an IBM AT computer A real-time clock in the computer calculated response times in milliseconds from the time the target stimulus appeared on the screen to the time the subject depressed one of two telegraph keys connected to the computer. The computer also recorded which key was depressed.

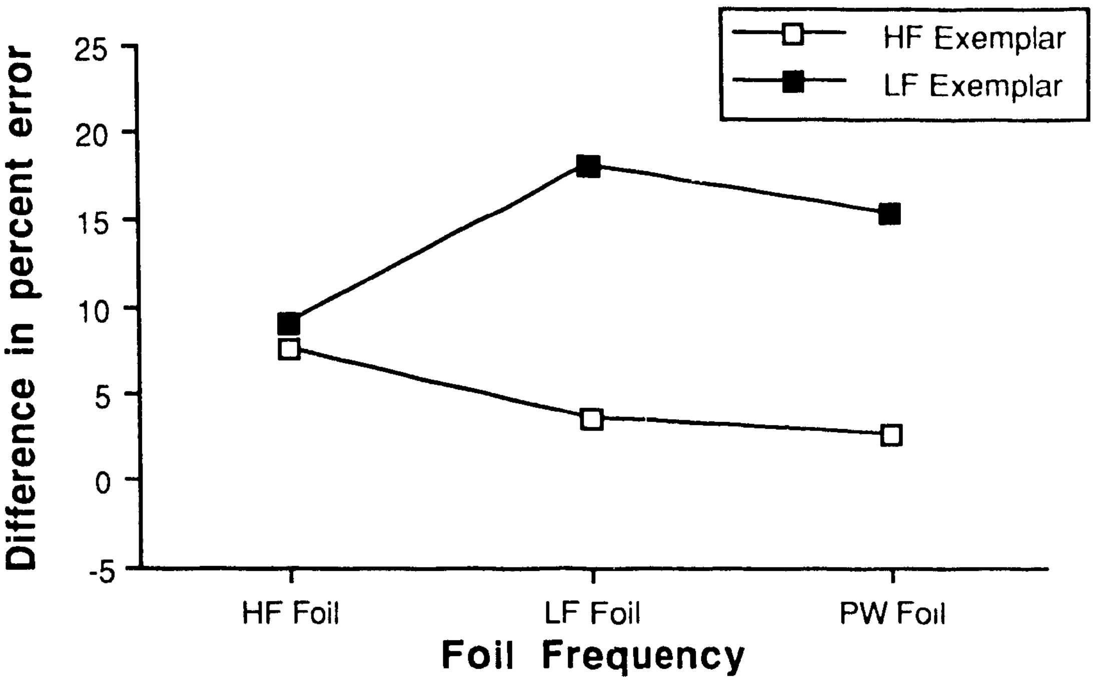
Results

Three sets of analyses were conducted. The first set examined error rates (false YES responses) on homophone foils, spelling controls, and homophone controls. A second set examined correct NO latencies on these three groups. A third set compared erroneous YES latencies on homophone foils and correct YES responses on exemplars. Analyses of variance using both subject and item means (Clark, 1973) were used in the first two sets of analyses. There were three factors in each analysis: category-target relation (homophone foil, spelling control, homophone control), target frequency (high, low, pseudoword), and frequency of the exemplar (high, low) The factors were treated as within-subject factors in the analyses using subject means and were analyzed as between-subject factors in the analyses using item means Planned comparisons were performed to test for significant differences between pairs of means that were theoretically relevant In this and subsequent experiments, subject means are reported in the text and figures

False Positive Error Data. Percent errors were arcsine transformed prior to analysis, although the untransformed data produced essentially the same results The untransformed percentages are reported in the text The differences in percent errors between homophone foils and spelling controls are presented in Figure 1, and the differences between homophone foils and homophone controls are presented in Figure 2 There was a main effect of category-target relation, $F(2,22) = 38.96$, $p < .001$ by subjects and $F(2,198) = 19.17$, $p < .001$ by items Subjects made 16.2% errors on homophone foils (e.g. car part-BREAK), 6.8% on spelling controls (e.g car part-BRAVE), and 3.1% on homophone controls (e g painter's equipment-BREAK) Planned comparisons indicated that significantly more errors were made on homophone foils than on either spelling controls, $F(1,11) = 48.10$, $p < .001$ by subjects and $F(1,198) = 16.75$, $p < .001$ by items, or homophone controls, $F(1,11) = 45.39$, $p < .001$ by subjects and $F(1,198) = 36.82$, $p < .001$ by items

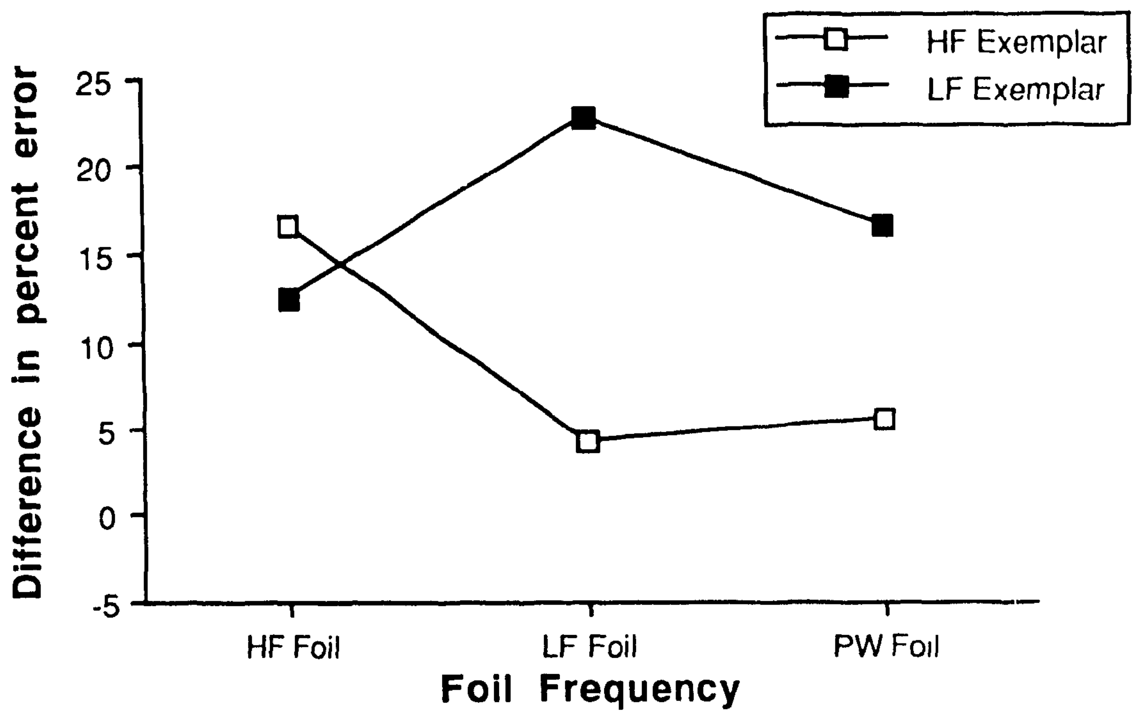
The main effect of target frequency was significant in the subject analysis, $F(2,22) = 9.94$, $p < .001$, and approached significance in the item analysis, $F(2,198) = 2.93$, $p < .06$. Subjects made more errors on high frequency words (11.5%) than on low frequency words (8.5%) and pseudowords (6.2%). The interaction between category-target relation and target frequency was not significant either by subjects, both F's < 1 The differences

Figure 1: The difference in mean percent errors between homophone foils and spelling controls in Experiment 1.



HF Exemplar
LF Exemplar
Difference in percent error
25
20
15
10
5
0
-5
HF Foil
LF Foil
PW Foil
Foil Frequency

Figure 2: The difference in mean percent errors between homophone foils and homophone controls in Experiment 1.


HF Exemplar
LF Exemplar
Difference in percent error
25
20
15
10
5
0
-5
HF Foil
LF Foil
PW Foil
Foil Frequency

between homophone foils and spelling controls were of similar magnitude for high frequency words (8.3%), low frequency words (10.6%), and pseudowords (9.0%). According to Myers (1979), planned tests can be performed even if the overall F is not significant. Consistent with Van Orden's (1987) view, and contrary to a prediction of dual route theory, planned comparisons indicated that the difference for high frequency words was significant both by subjects, $F(1,11) = 20.03$, $p < .001$, and items, $F(1,198) = 4.49$, $p < .05$. The differences between homophone foils and homophone controls were also of similar magnitude for high frequency words (14.6%), low frequency words (13.6%), and pseudowords (11.1%).

The main effect of exemplar frequency was not significant, $F(1,11) = 1.00$, $p > .05$ by subjects and $F < 1$ by items. The interaction of category-target relation and exemplar frequency was significant by subjects, $F(2,22) = 9.05$, $p < .01$, but not by items, $F(2,198) = 2.24$, $p > .05$. The difference in percent errors between homophone foils and spelling controls was 14.1% for those with low frequency exemplars, and 4.6% for those with high frequency exemplars. Consistent with Van Orden's (1987) view, the planned comparisons indicated that the difference for foils with low frequency exemplars was significant, $F(1,11) = 101.50$, $p < .001$ by subjects and $F(1,198) = 17.96$, $p < .001$ by items, but the difference for foils with high frequency foils was not significant. The difference between homophone foils and homophone controls was 17.4% for those with low frequency exemplars and 8.8% for those with high frequency exemplars.

The triple interaction was significant by subjects, $F(4,44) = 3.09$, $p < .05$, but not by items, $F(4,198) = 1.00$, $p > .05$. Planned comparisons performed to examine the difference in error rate between homophone foils (e.g. car part-BREAK) and spelling controls (e.g. car part-BRAVE) in each of the six groups revealed that the difference was significant by both subjects and items only for low frequency foils with low frequency exemplars, $F(1,11) = 26.25$, $p < .001$ by subjects and $F(1,198) = 8.44$, $p < .01$ by items, and pseudohomophones with low frequency exemplars, $F(1,11) = 34.33$, $p <$ .001 by subjects and $F(1,198) = 7.69$, $p < .01$ by items. The differences in mean error rates between homophone foils and spelling controls for the six groups are presented in Figure 1.

Planned comparisons were also performed on the differences between homophone foils (e.g. car part-BREAK) and homophone controls (e.g. painter's equipment-BREAK) Planned comparisons indicated that these differences were significant by subjects and items in four of the six groups: high frequency foils with high frequency exemplars, and high frequency foils, low frequency foils, and pseudohomophones with low frequency exemplars, all p's < .02. The differences in percent error between homophone foils and homophone controls for the six groups are presented in Figure 2

NO Reaction Time Data. This analysis examined correct NO latencies on homophone foils (e.g. car part-BREAK), spelling controls (e.g car part-BRAVE), and homophone controls (e.g. painter's equipment-BREAK). A subject's response latency on a trial was only included in the analyses if the subject responded correctly on that trial, and also responded correctly to the other three members of the stimulus quadruple (e g responses on car part-BRAKE, car part-BREAK, car part-BRAVE, painter's equipment-BREAK all had to be correct for them to be included) This procedure ensured that the same number of scores were included in each of the three category-target relation conditions in the analysis (Van Orden, 1987; Van Orden et al , 1988) The additional constraint that a trial was included only if the subject correctly responded to the related exemplar ensured that a NO response to the foil (e.g. car part-BREAK) occurred because subjects correctly avoided confusion with the exemplar, and not because subjects activated the representation of the exemplar (e.g. BRAKE) via phonology, but thought that the exemplar was a nonmember of the category (e g that a BRAKE was not a car part) These criteria were met by 69.6% of responses. Another 21 6% of responses were correct but were discarded because the subject made an error on another member of the quadruple. Essentially the same results were obtained when all correct responses were included in the analyses. Nine response times greater than 1500 ms were replaced with times of 1500 ms

There was a main effect of category target relation in the NO latency data, $F(2,22) = 18.61$, $p < .001$ by subjects and $F(2,195) = 10.53$, $p < .001$ by items. Planned comparisons revealed that homophone foils (702 ms) produced significantly longer latencies than spelling controls (665 ms), $F(1,11) = 15.92$, $p < .01$ by subjects and $F(1,195) = 7.61$, $p < .01$ by items, and significantly longer latencies than homophone controls (641 ms), $F(1,11) = 37.72$, $p < .001$ by subjects and $F(1,195) = 20.77$, $p < .001$ by items None of the other main effects or interactions were significant.

Contrary to what Van Orden (1987) and Van Orden et al. (1988) observed in their experiments, the main effect of category-target relation in the NO latency data was not due to a few outlying scores. A second analysis was performed with all scores greater than 1000 ms removed, along with their corresponding scores in the other two category-target relation conditions. This resulted in the removal of 10.5% of the scores included in the previous analyses. The main effect of category-target relation was still significant, $F(2,22) = 21.14$, $p < .001$ by subjects and $F(2,195) = 13.59$, $p < .001$ by items. Planned comparisons revealed that homophone foils produced significantly longer latencies than either spelling controls (19 ms), $F(1,11) = 15.82$, $p < .01$ by subjects and $F(1,195) = 7.40$, $p < .01$ by items, or homophone controls (46 ms), $F(1,11) = 31.18$, $p < .001$ by subjects and $F(1,195) = 27.19$, $p < .001$ by items.

YES Response Latencies. In a final set of analyses, correct YES response latencies on exemplars were compared with false positive response latencies on homophone foils to examine whether exemplars benefit from having orthographic information consistent with the YES response. One-tailed *t* tests of subject and item means were used because too few errors were made to perform an analysis of variance with the foil frequency and exemplar frequency factors An item was included in these analyses only if a subject made both a false positive error on the homophone foil and correctly accepted the matched exemplar. Nine out of 140 false positive errors on homophone foils were not included because the subject failed to respond correctly to the corresponding exemplar. Subjects responded YES significantly faster to exemplars (546 ms) than to the homophone foils (620 ms), $t(11) = 3.61$, $p < .005$ by subjects and $t(88) = 2.33$, $p < .02$ by items. This tendency occurred in each of the six groups, however, there were too few errors in each group (the range was from 8-33) to examine the effects of exemplar and foil frequency. The difference between exemplars and homophone foils was exaggerated by a few extreme scores. When 12 (9.2%) false positive homophone foil responses longer than 900 ms and the corresponding exemplar response were removed, the difference between correct YES responses (540 ms) and false positive response latencies (566 ms) only approached significance by subjects, $t(11) = 1.78$, $p < .06$, and was not significant by items, $t(82) = 1.00$, $p > .05$.

Discussion

The results of the error analyses were similar to those of Van Orden (1987, Van Orden et al., 1988). He found that subjects produced significantly more false positive errors on homophone foils than on spelling controls, and this finding was replicated here Further, Van Orden (1987, Expt. 3) did not observe an effect of foil frequency on the magnitude of the difference between homophone foils and controls when all of the stimuli were words, and Van Orden et al. (1988) observed a similar sized difference between foils and spelling controls for words and pseudowords Consistent with this, in Experiment 1 the size of the difference between homophone foils and spelling controls was similar for high frequency words, low frequency words, and pseudohomophones. Van Orden (1987) found that the magnitude of the difference in false positive error rates between homophone foils and spelling controls was influenced by the frequency of the exemplar. Consistent with this, in Experiment 1 the difference between homophone foils and spelling controls was larger when homophone foils had low frequency exemplars than when they had high frequency exemplars.

In the NO latency data, an overall difference between homophone foils and spelling controls was observed, as in the Van Orden (1987) and Van Orden et al. (1988) studies. They argued that the difference in their latency data between the two groups was due to a small number of outliers in the homophone foil group, and that otherwise the NO latency distributions for the two groups were essentially the same. However, the difference between homophone foils and spelling controls remained in Experiment 1 when outliers were excluded from the analyses. Van Orden (1987) did not report the percentage of scores that were removed from the analysis that produced similar means for the two groups, but Van Orden et al. (1988) reported that the means did not differ when about 30% of scores were removed. It is debatable whether these could all be considered outliers. Here outliers were considered to be the upper 10% of scores. The data of Experiment 1 were consistent with Van Orden's (1987, Expt. 3) observation that target and exemplar frequency did not affect the magnitude of the NO latency difference between homophone foils and controls.

Van Orden et al. (1988, Expt. 1) found similar latencies for false positive responses on homophone foils and for correct YES responses on category exemplars. In the second experiment in the paper, the exemplars were matched more closely to the foils and this time the YES latencies were faster for exemplars than for foils (49 ms faster for words, 63 ms faster for pseudowords), although the differences were not significant. Van Orden et al. argued that the differences were due to outlier latencies for homophone foils, and demonstrated that the means for foils and exemplars were the same when 30% of scores were removed In Experiment 1, YES latencies collapsed across group were significantly faster for exemplars than for homophone foils, but there was no difference when about 10% of trials with the longest latencies were removed. This suggests that the exemplars benefitted little from having orthographic information consistent with the YES response. However, this conclusion is based on a small number of responses and, as mentioned earlier, it may not be valid because erroneous YES responses may be made before foils are fully processed.

In sum, the overall results of Experiment 1 replicated Van Orden's results quite well, with the exception of the NO latency distributions. In addition, Experiment 1 has demonstrated that the higher error rate for homophone foils than for spelling controls is not due to a more general difficulty in processing homophones, since homophone foils produced significantly more errors than homophone controls. The difficulty with homophones on a semantic decision task arises only when they sound like an exemplar of a category.

Van Orden (1987; Van Orden et al., 1988) argued that the results of his experiments provided evidence that subjects only access the meanings of words using a phonological representation. The findings cited in support of this conclusion were the higher error rate on homophone foils than on spelling controls regardless of frequency, the similar YES latencies on exemplars and foils, and the similar false positive error rates on homophone and pseudohomophone foils. This last finding provides evidence that the phonological effects arise prior to the access of meaning, rather than following access. Similar NO latency distributions for homophone foils and spelling controls suggested to him that subjects were making use of a spelling check for both types of stimuli, although problems with this hypothesis were mentioned earlier. He interpreted the effect of exemplar frequency on homophone foil error rates as additional support for the verification hypothesis since better knowledge of the exemplar would facilitate a spelling check. Dual route theory, he claimed, predicts that a difference between homophone foils and spelling controls should be observed for low frequency foils but not high frequency foils. A high frequency foil should activate its meaning directly and thus not be influenced by its phonological representation, but a low frequency word is accessed via the phonological route and so should be more susceptible to phonological confusion errors. However, in neither Van Orden's (1987) experiment, nor in Experiment 1 was the difference between homophone foils and spelling controls affected by foil frequency. In addition, Van Orden (1987) pointed out that dual route theory predicts that the difference between homophone foils and spelling controls should be greater when homophone foils have high frequency exemplars than when they have low frequency exemplars. This is because the high frequency exemplars would reach maximum levels of activation sooner than low frequency words (Morton, 1969), and thus would be more likely to be mistakenly selected. In contrast, in Van Orden's (1987) experiment and in Experiment 1, fewer errors were made on homophone foils with high frequency exemplars than on those with low frequency exemplars

The overall results of Experiment 1 appear to support Van Orden's view. Closer inspection of the data, however, reveal that the evidence for the use of phonological information in accessing the meanings of high frequency words is not strong. Although high frequency foils produced significantly more errors than spelling controls, when the high frequency foils with high frequency exemplars and high frequency foils with low frequency exemplars were examined separately, the differences for each group were only significant in the subjects analysis, which suggests the effect is limited to only some of the words. In addition, the magnitude of the difference between homophone foils and spelling controls was similar for high frequency foils with high frequency exemplars and high frequency foils with low frequency exemplars. This finding is not consistent with Van Orden's proposal that all words undergo a spelling check, since in his view subjects should have more complete knowledge of the spelling of a high frequency exemplar and thus should be better able to detect its homophone foil in the spelling check

Dual route theory, on the other hand, has difficulty explaining the absence of a false positive error effect for low frequency and pseudohomophone foils with high frequency exemplars. These stimuli should have accessed semantic representations on the basis of phonological information, and since the semantic representation of the exemplar is also activated with homophone foils, the foils should have been more susceptible to false positive errors than spelling controls.

The weak evidence for the use of phonological mediation in accessing the meaning of high frequency words may be due to the small number of stimuli used in each of the groups in the experiment (although more were used here than in Van Orden's experiment), and perhaps if more high frequency homophone stimuli were available, stronger effects would be found. However, another possibility is that the effect of homophone foils versus spelling controls was exaggerated in Experiment 1 and Van Orden's experiments, particularly for high frequency words, because of priming from the category name. Experiment 2 explores this possibility.

Experiment 2:

The Effect of Category Specificity

Experiment 2 addressed whether the false positive error rate in the category decision task is affected by task-specific strategies. Specifically, false positive errors may arise, in part, from subjects' attempts to generate potential targets. Balota and Chumbly (1984), Forster (1988), and Monsell et al (1989) have suggested that category names may prime exemplars in a category decision task Becker (1976) specifically proposed in his verification model that subjects generate a semantic candidate set when shown a prime. Several studies have demonstrated priming by category names in a lexical decision task (for a review see Neely, in press) Using a variation of the semantic decision task, Rosch (1975) found that prior presentation of the category name facilitated judgments of whether a pair of words belonged to the same category relative to a neutral prime (the word BLANK), and that the size of the facilitation effect was similar for pairs of words that were good and poor exemplars of the category Further evidence for the priming of exemplars by category names in a semantic decision task comes from the pilot work of Van Orden (1987, Expt 2) who observed that subjects needed to view an exemplar target for less time than a nonexemplar target in order to be able to report it.

Once subjects are shown a category name, they may begin to generate possible semantic candidates, and these candidates may themselves be phonologically recoded, much as they are when a spoken response is prepared. Consequently, when the target stimulus appears on the screen, there may already be considerable activation in both the semantic and phonological systems. The phonological code of the target may be generated automatically shortly after the target is presented The combination of activation of the phonological representation generated by the prime and the phonological information generated by the target may be enough to trigger a YES response, which would be a false positive response in the case of a homophone foil. That is, the subject may respond on some occasions before the target has activated a semantic representation This means that at least some false positive errors on homophone foils may not reflect phonological mediation of meaning. The orthographic to semantic conversion may normally be faster than processing along the phonological route, which requires an orthographic to phonological conversion and a phonological to semantic conversion However, the orthographic to semantic conversion may not always be faster than the orthographic to phonological conversion alone

Van Orden's (1987) strong view that phonological mediation is obligatory predicts that the effects that he observed should not be dependent on the nature of the category names used. That is, more errors should still be observed on homophone foils than on spelling controls for both high and low frequency targets when broad categories are used His position would be called into question if the effects were highly dependent on this task specific aspect of the studies.

To reduce the likelihood that phonological representations are activated by the category name prior to the presentation of the target, two broad categories, LIVING THING and OBJECT, were used in Experiment 2 instead of the more specific categories used in Experiment 1. Exactly the same target stimuli as in Experiment 1 were used in order to ensure that any differences between the experiments could be attributed to effects of category specificity. If the false positive errors found in Experiment 1 reflect only the access of meaning by phonological representations, then results of Experiment 2 should be similar. If on the other hand, priming from the category name resulted in subjects responding prior to the activation of a semantic representation of the target, then fewer, if any, false positive errors should be observed in Experiment 2.

Method

Subjects. Twelve McGill University undergraduates were paid $10 each to participate in the study. All were native speakers of English.

Stimuli. The target stimuli were the same 1008 as those used in Experiment 1 (see Appendix A) The category names used were LIVING THING and OBJECT. Half of the experimental stimuli and half of the filler stimuli were preceded with the category LIVING THING and half of each were preceded by the category OBJECT. The category name that preceded a target was chosen so that the correct response to a target stimulus was the same as in Experiment 1 Half of the items in each category had a correct response of YES and half had a correct response of NO The target items appeared on the same one of four lists and in the same order as in Expt 1 No more than three trials with the same category name or the same correct response appeared in succession

Procedure. The viewing conditions and procedure were exactly the same as in Experiment 1

Results

The data were analyzed in the same manner as in Experiment 1. As before, three sets of analyses were conducted. The first set examined error rates (false YES responses) on homophone foils, spelling controls, and homophone controls, the second set examined correct NO latencies on these three groups, and the third set compared erroneous YES latencies on homophone foils and correct YES responses on exemplars The scores on six words in the homophone control condition (SUN, MAIL, BEACH, CELLAR, PRINTS, and BUOY) were not included in the analyses The homophone control condition consisted of pairing the homophone foil words with an unrelated category name, such that the correct response was also NO (e.g in Experiment 1, a foil trial was car part-BREAK and the homophone control was painter's equipment-BREAK). However, although these six words are foils for the category LIVING THING, they are also exemplars of the only other category used in the experiment (OBJECT), and thus could not be presented in an unrelated category These words were included in the experiment anyway because of the difficulty of finding enough pairs of homophones in which one was living thing or an object and the other one was neither. Since the main contrast of interest is between homophone foils and spelling controls the exclusion of these items has little effect on the conclusions.

False Positive Error Data. The overall error rate (9.8%) was similar to that in Experiment 1 (8.7%). The differences in percent errors between homophone foils and spelling controls are presented in Figure 3, and the differences between homophone foils and homophone controls are presented in Figure 4. There was a main effect of category-target relation, $F(2,22) = 12.57$, $p < .001$ by subjects, and $F(2,192) = 6.30$, $p < .01$ by items. Subjects made 14.2% errors on homophone foils (e.g. object-BREAK), 8.8% on spelling controls (e.g. object-BRAVE), and 6.5% on homophone controls (e.g. living thing-BREAK). Planned comparisons indicated that significantly more errors were made on homophone foils than on either spelling controls, $F(1,11) = 15.47$, $p < .01$ by subjects and $F(1,192) = 6.96$, $p < .01$ by items, or homophone controls, $F(1, 11) = 13.67$, $p < .01$ by subjects and $F(1,192) = 11.39$, $p < .001$ by items.

The main effect of target frequency was not significant, $F(2,22) = 1.53$, $p > .05$ by subjects and $F < 1$ by items. Subjects made 11.2% errors on high frequency words, 9.2% on low frequency words, and 9.1% on pseudowords. The interaction of category-target relation and target frequency was also not significant, $F(4,44) = 2.39$, $p > .05$ by subjects and $F < 1$ by items. The difference between homophone foils and spelling controls was 0.4% for high frequency words, 8.0% for low frequency words, and 8.0% for pseudowords. The difference between homophone foils and homophone controls was 4.6% for high frequency words, 10.3% for low frequency words, and 16.6% for pseudowords.

The main effect of exemplar frequency was not significant, $F(1,11) = 3.06$, $p > .05$ by subjects and $F < 1$ by items. The interaction of category-target relation and exemplar frequency was marginally significant by subjects, $F(2,22) = 3.32$, $p = .05$, but not by items, $F < 1$. Planned comparisons indicated that there was a significant difference

Figure 3. The difference in mean percent errors between homophone foils and spelling controls in Experiment 2.

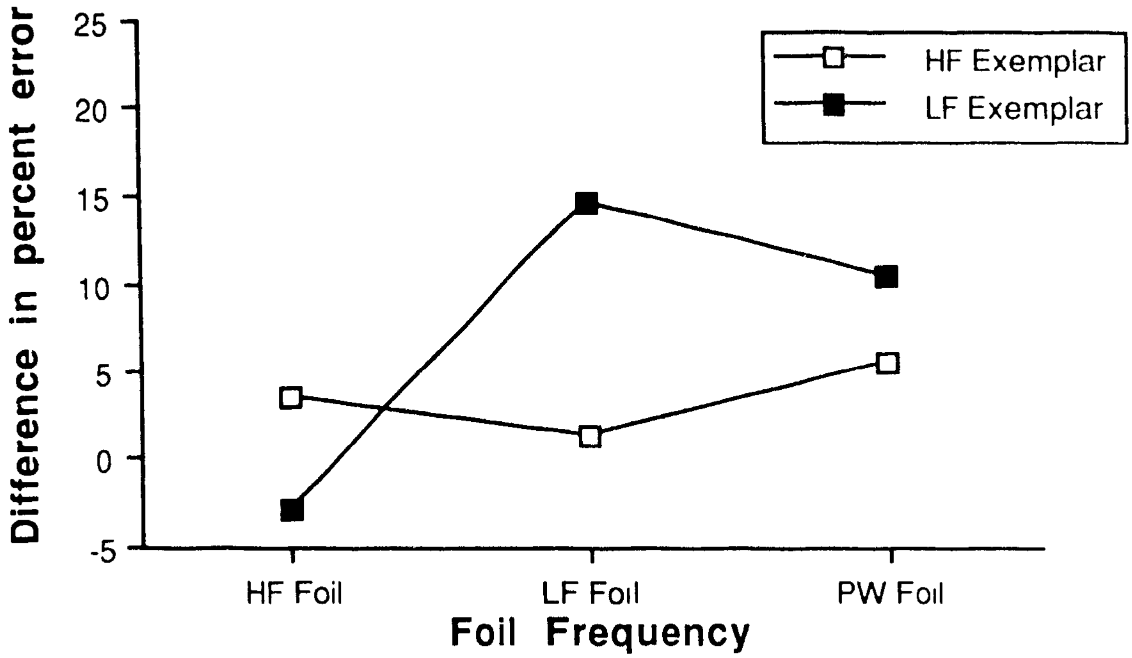
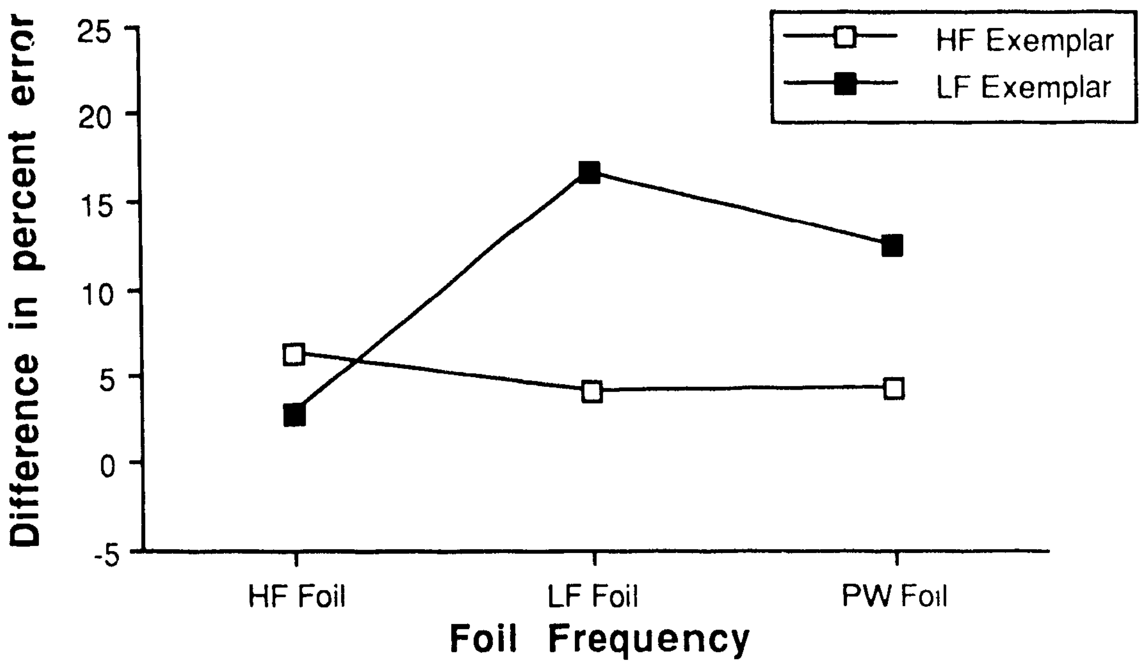

Difference in percent error
25
20
15
10
5
0
-5
HF Exemplar
LF Exemplar
HF Foil
LF Foil
PW Foil
Foil Frequency

Figure 4: The difference in mean percent errors between homophone foils and homophone controls in Experiment 2.


HF Exemplar
LF Exemplar
Difference in percent error
25
20
15
10
5
0
-5
HF Foil
LF Foil
PW Foil
Foil Frequency

between homophone foils with low frequency exemplars and spelling controls (7.3%), $F(1,11) = 14.42, p < .01$ by subjects and $F(1,192) = 4.69, p < .05$ by items, but the difference between homophone foils with high frequency exemplars and spelling controls (3.5%) was not significant. The difference between homophone foils and homophone controls was 10.6% for those with low frequency exemplars and 4.8% for those with high frequency exemplars.

The triple interaction was significant by subjects, $F(4,44) = 2.77, p < .05$, but not by items, $F < 1$. Planned comparisons were performed to examine the difference in error rate between homophone foils (e.g. object-BREAK) and spelling controls (e.g. object-BRAVE) in each of the six exemplar frequency/foil frequency groups. These differences are presented in Figure 3. The tests revealed that the differences between homophone foils and spelling controls was significant for low frequency words with low frequency exemplars, $F(1,11) = 15.75, p < .01$ by subjects and $F(1,192) = 4.93, p < .05$ by items, and was significant for pseudowords with low frequency exemplars in the subjects analysis, $F(1,11) = 15.26, p < .01$, and approached significance in this group by items, $F(1,192) = 3.37, p < .07$. The differences did not approach significance for the high frequency word groups, all p's $> .20$

Also of interest are comparisons between homophone foils (e.g. object-BREAK) and homophone controls (e.g. living thing-BREAK). The data are presented in Figure 4. Planned comparisons indicated that this difference was significant for the same two groups, the low frequency words with low frequency exemplars, $F(1,11) = 38.74, p < .001$ by subjects and $F(1,192) = 7.15, p < .01$ by items, and the pseudowords with low frequency exemplars, $F(1,11) = 15.26, p < .01$ by subjects and $F(1,192) = 5.12, p < .05$ by items.

NO Reaction Time Data. The correct NO decision latencies on homophone foils, spelling controls, and homophone controls were included in this analysis. As in

Experiment 1, a subject's response latency for an item was only included if the subject responded NO correctly to it, and responded correctly to the other three members of the stimulus quadruple (e g. object-BRAKE, object-BREAK, object-BRAVE, living thing BREAK). This criterion was met by 67 4% of responses Another 22.2% of responses were correct but were discarded because the subject made an error on another member of the quadruple Essentially the same results were found when all correct reaction times were included. Twenty-five response times greater than 1500 ms were replaced with times of 1500 ms.

The main effect of category-target relation in the NO latency data was significant by subjects, $F(2,22) = 3.91$, $p < .05$, but not by items, $F(2,192) = 1.46$, $p > .05$ The difference between homophone foils and spelling controls was 11 ms, and the difference between homophone foils and homophone controls was 34 ms The main effect of target frequency approached significance by subjects, $F(2,22) = 3.26$, $p < .06$ and was significant by items, $F(2,192) = 4.33$, $p < .02$ Subjects responded more quickly to high frequency words (763 ms) than to low frequency words (786 ms) and pseudowords (798 ms). None of the other interactions even approached significance, all F's < 1.

YES Response Latencies. An item was included in these analyses only if a subject made both a false positive error on the homophone foil and correctly accepted the matched exemplar. Ten out of 123 false positive errors on homophone foils were not included because the subject failed to respond correctly to the corresponding exemplar Subjects responded YES significantly faster to exemplars (678 ms) than to the homophone foils (775 ms) in the analysis by subjects, $t(10) = 2.94$, $p < .02$, but the difference was not significant by items, $t(80) = .99$, $p > .05$ This trend occurred in each of the six groups. As in Experiment 1, there were too few errors in each group (the range was from 11-29) to include frequency of exemplar and frequency of foil in the analysis

Discussion

A significant overall difference in the false positive error rate between homophone foils and spelling controls was found in Experiment 2, replicating the findings of Experiment 1 and those of Van Orden (1987, Van Orden et al., 1988). In addition, homophone foils produced significantly more errors than homophone controls, which, as in Experiment 1, indicates that the elevated rate of errors on homophone foils arises because they sound like a member of the category, not because they are generally harder to process.

However, unlike Experiment 1 and Van Orden's (1987, Expt. 3) experiment, high frequency homophone foils did not produce more false positive errors than spelling controls. Thus, the presence of an effect for high frequency words depends on the type of category given to the subject. When the category was specific (Experiment 1), such as car part, subjects made significantly more errors (8.3%) on homophone foils than on spelling controls. However, when the category was broad (Experiment 2), such as object, the difference was only 0.4% on exactly the same words. This suggests that in Experiment 1, the effects for high frequency words were inflated by priming from the category name. Subjects may have responded on the basis of a match between a phonological candidate activated by the category name and phonological activation generated by the target prior to activation of a semantic representation. The results of Experiment 2 suggest, then, that the use of specific categories in a category decision task should be avoided when studying the access of meaning in single word reading so that subjects cannot make strong predictions about subsequent targets.

The failure to find a difference in false positive error rates between homophone foils and spelling controls for high frequency words suggests that meaning is not phonologically mediated for these words. If high frequency words had been phonologically recoded, two meanings would have been available for the homophones whereas only one would have been available for the spelling controls. On at least some trials, subjects could be expected to erroneously choose the wrong homophone alternative, and make a false positive error. However, there was no evidence that the decision was more difficult for homophones than for spelling controls.

Although Experiment 2 provided no evidence that the access of meanings of high frequency words is phonologically mediated, the results do suggest that phonological mediation occurs for low frequency words. More false positive errors were made on homophone foils than on spelling controls for low frequency words and for pseudowords, which suggests that two meanings were available for the homophonic stimuli. The size of the difference was 8.0% for both groups. Van Orden et al. (1988) argued that similar sized effects for words and pseudowords indicated that phonological information mediated the access of meaning, instead of being activated following access. They also claimed that orthographic information did not contribute to the activation of meaning since exemplars did not benefit from having orthography consistent with a YES response. In Experiment 2, the difference between YES latencies for exemplars and homophone foils was significant by subjects. Since the YES analyses are based on a small number of responses, it would not, therefore, be safe to conclude from this data that orthography never contributes to semantic activation.

Two other findings were consistent with Van Orden's account of a spelling check procedure for phonologically activated candidates. The size of the difference in the error data was larger for foils with low frequency exemplars. According to his view, the lower error rate on foils with high frequency exemplars occurs because subjects have better knowledge of their spellings and thus are more able detect homophone foils The second finding was that there was no difference between homophone foils and spelling controls in NO latencies. Van Orden (1987) claimed that the lack of difference in the NO latency data indicates that all phonologically activated candidates undergo a spelling check, not just homophones. However, as pointed out earlier, his view actually predicts that homophone NO latencies should sometimes be longer than spelling control latencies.

Another piece of evidence that Van Orden (1987) used to support his view that candidates are activated exclusively via their phonological representations comes from his tachistoscope experiment. In that experiment he found a large difference in errors between foils and spelling controls that was not dependant on the orthographic similarity of the foil to its exemplar. He argued that the mask that followed the brief presentation of the stimuli prevented the spelling check from being performed on phonologically activated candidates.

The finding that priming from specific category names exaggerated errors on homophone foils in Experiment 1 suggests that when the spelling check was prevented by masked tachistoscopic presentation, the rate of false positive errors on foils should have been considerably larger than 50% if candidates are activated exclusively via phonological information (Coltheart, 1978). That is, given two candidates activated by a phonological representation (the foil and the exemplar) and no orthographic information to tell them apart, subjects should be more likely to choose the exemplar because it is primed by the category name. Van Orden (1987, Expt. 2) found an error rate of only 43% on homophone foils, which indicates that there is some activation of candidates via orthographic information.

Experiment 3:

The Effect of the Proportion of Homophones

The results of Experiment 2 suggest that skilled readers make use of phonological information in accessing the meanings of low frequency words but not high frequency words. Experiment 3 explores the possibility that the use of phonology in the access of the meanings of low frequency words is not obligatory, but rather is a strategy under the control of the subject.

Several authors (Coltheart, 1978; Davelaar et al., 1978; Hawkins et al., 1976; McQuade, 1981) have proposed that the use of phonological information is strategic, and that subjects will avoid the use of phonology when it impairs performance on a task. They attempted to observe strategic use of phonology by varying the proportion of homophonic stimuli in an experiment. The logic of the approach is that if stimuli are recoded phonologically, subjects should make more errors or take longer to make decisions on homophonic (e.g. BARE, GRONE) than nonhomophonic stimuli (e.g. BAKE, GROBE). However, when many homophonic stimuli are included, subjects would notice that the phonological strategy was causing them to make a large number of errors, and so they would abandon it in favor of a visual strategy. No effect of homophony would be expected if a visual strategy is used. Using a lexical decision task, Davelaar et al. (1978) found an effect of homophony for low frequency words when the pseudoword distractors did not sound like English words (e.g. SLINT), but the homophone effect disappeared when the distractors were pseudohomophones (e.g. GRONE). Also using lexical decision, McQuade (1981) found a larger pseudohomophone effect when a low proportion of the stimuli were pseudohomophones. Hawkins et al. (1976) found an effect of homophony in a tachistoscopic word recognition task only when the stimulus list contained a low proportion of homophones. These authors all concluded that subjects make use of a phonological strategy when this strategy causes few errors (i.e. when there are few homophonic stimuli) but abandon it and use a visual strategy when it leads to many errors (i.e. when a large number of homophonic stimuli are included).

It is unclear whether the proportion of homophones influences performance on a semantic decision task which requires subjects to focus on the meaning of the stimuli. In Van Orden's (1987; Van Orden et al., 1988) experiments and in Experiments 1 and 2, the proportion of homophones was kept low by including large numbers of filler trials The proportion of homophonic stimuli was 16.4% in Van Orden's first two experiments and was 10% in his third experiment and in the two Van Orden et al. (1988) experiments. In Experiments 1 and 2 the proportion of homophonic stimuli was slightly higher at 21.4% In Experiment 3, exactly the same experimental stimuli were used as in Experiments 1 and 2, but the 180 nonhomophonic filler trials per list were replaced by 36 homophone filler trials per list, so that 83.3% of the stimuli were homophones. The spelling controls were the only nonhomophonic stimuli. If subjects can strategically control their use of phonological recoding they should be much less likely to use it in Experiment 3, and thus show little difference in the number of false positive errors between homophone foils and spelling controls.

Method

Subjects. Twelve McGill University undergraduates were paid $8 each to participate in the study. All were native speakers of English.

Stimuli. As in Experiment 2, LIVING THING and OBJECT category names were used. The 288 experimental stimuli were the same as those used in Experiments 1 and 2 (see Appendix A). The filler words used in the previous experiments were removed from the four lists and replaced by 36 homophone fillers on each list. Half of these were LIVING THINGS and half were OBJECTS. Since it was impossible to find enough homophone words so that the 72 LIVING THING and 72 OBJECT filler homophones needed for the experiment were only used once, instead 36 of each were found and every filler appeared on two lists. They were distributed among the four lists such that a list had only 12 homophone fillers in common with any other list.

In sum, each of the four lists used in the experiment had 108 targets. Half of each were preceded with the category LIVING THING and half of each were preceded by the category OBJECT, and within each category, half of the targets were exemplars and half were not. No more than three trials with the same category name or the same correct response appeared in succession. A practice list of 16 trials, 14 of which had homophone targets, was also developed.

Procedure. The viewing conditions and procedure were exactly the same as in Experiments 1 and 2. Because fewer stimuli were used here than in the previous experiments, the experimental session lasted approximately 20 minutes.

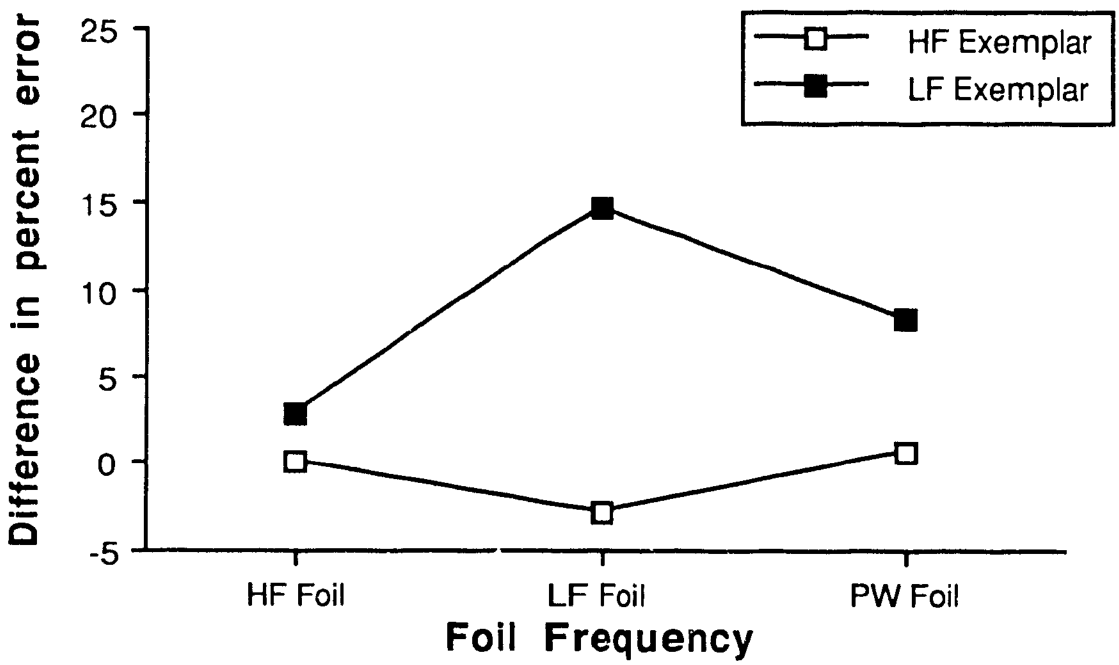
Results

Three sets of analyses were conducted, as in the previous two experiments. The first set examined error rates (false YES responses) on homophone foils, spelling controls, and homophone controls, the second set examined correct NO latencies on these three groups, and the third set compared erroneous YES latencies on homophone foils and correct YES responses on exemplars. As in Experiment 2, the scores on six words in the homophone control condition (SUN, MAIL, BEACH, CELLAR, PRINTS, and BUOY) were not included in the analyses.

False Positive Error Data. The overall error rate (6.7%) was lower than that in Experiment 1 (8.7%) and Experiment 2 (9.8%). The differences in percent errors between homophone foils and spelling controls are presented in Figure 5, and the differences between homophone foils and homophone controls in Figure 6.

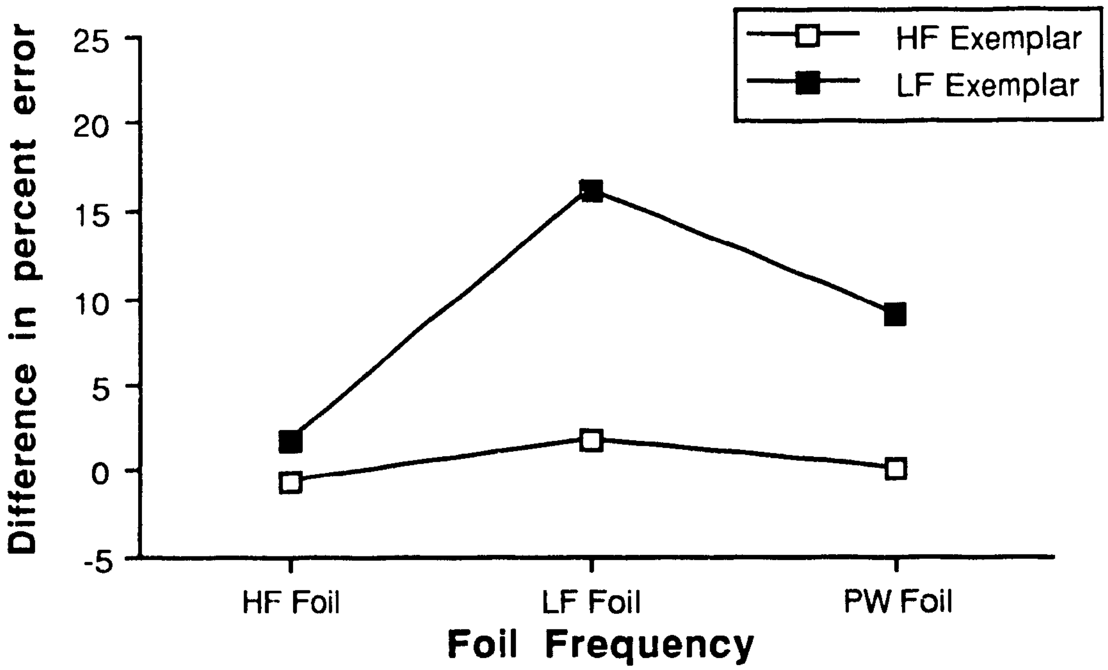
The main effect of category-target relation was significant by subjects, $F(2,22) = 4.58$, $p < .05$, and was marginally significant by items, $F(2,192) = 2.96$, $p = .054$. Planned comparisons indicated that the overall difference between homophone foils and spelling

Figure 5: The difference in mean percent errors between homophone foils and spelling controls in Experiment 3.



Difference in percent error
25
20
15
10
5
0
-5
HF Exemplar
LF Exemplar
HF Foil
LF Foil
PW Foil
Foil Frequency

Figure 6: The difference in mean percent errors between homophone foils and homophone controls in Experiment 3.



HF Exemplar
LF Exemplar
Difference in percent error
25
20
15
10
5
0
-5
HF Foil
LF Foil
PW Foil
Foil Frequency

controls only approached significance by subjects, $F(1,11) = 4.03$, $p < .07$, and by items, $F(1,192) = 2.95$, $p < .09$. The difference between homophone foils and homophone controls was significant in both analyses, $F(1,11) = 7.53$, $p < .02$ by subjects and $F(1,192) = 5.52$, $p < .02$ by items Subjects made 9.6% errors on homophone foils (e.g. object-BREAK), 5.7% on spelling controls (e g object-BRAVE), and 4.9% on homophone controls (e.g. living thing-BREAK). In Experiment 2, subjects made 14.2%, 8.8%, and 6.5% errors on these three groups respectively

The main effect of target frequency was significant, $F(2,22) = 10.54$, $p < .001$ by subjects and $F(2,192) = 4.44$, $p < .02$ by items Planned comparisons indicated that subjects made fewer errors on pseudowords (3.2%) than on either low frequency words (8.3%), $F(1,11) = 16.84$, $p < .01$ by subjects and $F(1,192) = 6.40$, $p < .02$ by items, or high frequency words (8.7%), $F(1,11) = 7.04$, $p < .01$ by subjects and $F(1,192) = 15.84$, $p < .01$ by items This is in contrast to the absence of a target frequency effect in Experiment 2. The interaction of category-target relation and stimulus frequency was not significant by subjects, $F(4,44) = 2.06$, $p > .05$, or by items, $F < 1$. The difference between homophone foils and spelling controls was 1.0% for high frequency words, 6.0% for low frequency words, and 4.5% for pseudowords Planned comparisons indicated that only the difference for pseudowords was significant by subjects, $F(1,11) = 6.57$, $p < .05$, and none of these differences were significant by items. In Experiment 2 these differences were 0.4%, 8.0%, and 8.0% respectively, and the latter two were significant. The difference between homophone foils and homophone controls was 0.6% for high frequency words, 8.9% for low frequency words, and 4.5% for pseudowords.

The main effect of exemplar frequency was significant by subjects, $F(1,11) = 5.33$, $p < .05$, but not by items, $F(1,192) = 1.09$, $p > .05$ by items The interaction of category-target relation and exemplar frequency and was significant by subjects, $F(2,22) = 10.37$, $p < .001$, and approached significance by items, $F(2,192) = 2.46$, $p < .09$. Planned comparisons indicated that homophone foils with low frequency exemplars produced significantly more errors (8.3%) than spelling controls, $F(1,11) = 16.82$, $p < .01$ by subjects and $F(1,192) = 7.12$, $p < .01$ by items, but there was no significant difference for homophone foils with high frequency exemplars (-0.6%). The difference between homophone foils and homophone controls was 8.7% for those with low frequency exemplars and 0.4% for those with high frequency exemplars. In Experiment 2, significant effects were also observed only for homophone foils with low frequency exemplars.

The triple interaction was not significant by subjects, $F(4,44) = 2.11$, $p > .05$, or by items, $F < 1$. Planned comparisons performed to examine the difference in error rate between homophone foils (e.g. object-BREAK) and spelling controls (e.g. object-BRAVE) revealed that the difference was significant for low frequency words with low frequency exemplars (14.6%), $F(1,11) = 16.25$, $p < .01$ by subjects and $F(1,192) = 5.95$, $p < .02$ by items, and was significant by subjects for pseudowords with low frequency exemplars (8.3%), $F(1,11) = 7.07$, $p < .05$, but not by items, $F(1,192) = 2.62$, $p > .05$. These were the two groups that produced the largest differences in Experiment 2. In Experiment 2, the item analysis for pseudowords approached significance ($p < .07$); here it was not significant ($p = .11$). None of the other differences approached significance. The differences in mean error rates between homophone foils and spelling controls for the six groups are presented in Figure 5.

Planned comparisons between homophone foils (e.g. object-BREAK) and homophone controls (e.g. living thing-BREAK) indicated that significantly more false positive errors were made on homophones foils than on homophone controls for low frequency words with low frequency exemplars, $F(1,11) = 24.01$, $p < .001$ by subjects and $F(1,192) = 7.67$, $p < .01$ by items, and the difference was significant by subjects for pseudowords with low frequency exemplars, $F(1,11) = 7.07$, $p < .05$, and approached significance by items, $F(1,192) = 3.13$, $p < .08$. These were the two groups that produced significant differences in Experiment 2. The differences in mean error rates between homophone foils and homophone controls for the six groups are presented in Figure 6

In sum, the overall size of the difference between homophone foils and spelling controls was smaller (by 1 5%) than in Experiment 2 and only approached significance here. The differences were largest in the same two groups as in Experiment 2. The overall size of the difference between homophone foils and homophone controls was also smaller (by 3%) than in Experiment 2 but remained significant. The difference was also largest in the same two groups as in Experiment 2. In contrast to the absence of a target frequency effect in Experiment 2, here fewer errors were made on pseudowords than on high or low frequency words.

NO Reaction Time Data. As in Experiments 1 and 2, a subject's response latency for an item was only included in the analyses if the subject responded correctly to it and the other three members of the stimulus quadruple (e g. object-BRAKE, object-BREAK, object-BRAVE, living thing-BREAK). This criterion was met by 73 7% of responses. Another 19.4% of responses were correct but were discarded because the subject made an error on another member of the quadruple. Essentially the same results were found when all correct reaction times were included. Five percent of response times were greater than 1500 ms and were replaced with times of 1500 ms. Subjects responded more slowly in Experiment 3 than in Experiment 2 by an average of 124 ms.

The main effect of category-target relation in the NO latency data was not significant, $F(2,22) = 2.94$, $p > .05$ by subjects and $F < 1$ by items. Decision latencies were 919 ms for homophone foils, 915 ms for spelling controls, and 896 ms for homophone controls. In Experiment 2 the difference between homophone foils and spelling controls was 4 ms and the difference between homophone foils and homophone controls was 34 ms.

The main effect of target frequency was significant by subjects, $F(2,22) = 13.88$, $p < .001$, and by items, $F(2,192) = 13.73$, $p < .001$. Subjects responded more quickly to pseudowords (865 ms) than to high frequency words (935 ms) and low frequency words (930 ms). This is in contrast to the results of Experiment 2 where high frequency targets were responded to faster than low frequency words and pseudowords

No other differences were found in Experiment 2, and here there were two further effects that were significant in the subject analysis only. The interaction of category-target relation and target frequency was significant by subjects, $F(4,44) = 7.72$, $p < .001$, but not by items, $F(4,192) = 1.50$, $p > .05$, and the interaction of category-target relation and exemplar frequency was significant by subjects, $F(2,22) = 4.89$, $p < .02$, but not by items, $F < 1$ Neither the main effect of exemplar frequency, nor the triple interaction were significant by subjects or items.

YES Response Latencies. An item was included in these analyses only if a subject made both a false positive error on the homophone foil and correctly accepted the matched exemplar Seven out of 82 false positive errors on homophone foils were not included because the subject failed to respond correctly to the corresponding exemplar. Subjects responded YES 12 ms faster to exemplars than to homophone foils, but the difference was not significant by subjects, $t < 1$, or by items, $t(62) = 1.28$, $p > .05$. The difference was 94 ms for low frequency foils with low frequency exemplars and their matched exemplars, but this difference was not significant, probably because only 24 false positive errors were made on words in this group. Fewer were made in each of the other groups.

Discussion

The subjects in Experiment 3 responded more cautiously than subjects in Experiment 2, trading off speed for accuracy. They made 3.1% fewer errors than subjects in Experiment 2, but their decision latencies were 124 ms slower on average.

The evidence for phonological mediation was weaker than in Experiment 2. The overall difference between homophones and spelling controls was 1.5% smaller than in

Experiment 2 and only approached significance. The difference for low frequency words was 2% smaller and the difference for pseudowords was 3.5% smaller, and both were no longer significant. However, low frequency foils with low frequency exemplars still produced significantly more errors than spelling controls, and in fact, the numerical difference was the same as in Experiment 2 (14.6%). The difference between pseudohomophone foils with low frequency exemplars and their spelling controls dropped by 2.1% to 8.3%, and although the difference was still significant by subjects, it no longer approached significance by items.

Van Orden et al. (1988) have argued that evidence for phonological mediation of the access of meaning is the finding of a similar-sized homophone foil effect for words and pseudowords. One possible interpretation of the stronger effect for low frequency words here is that it may reflect, at least in part, the retrieval of phonological information following the access of meaning. However, several aspects of the data suggest instead that it was the pseudowords that were processed rather differently in Experiment 3. While the overall drop in error rate from Experiment 2 to Experiment 3 was 3.1%, the drop was 5.9% in errors made on pseudowords and only 0.8% for low frequency words. In Experiment 3, subjects produced significantly fewer errors and faster latencies on pseudowords than on either high or low frequency words. In contrast, in Experiment 2, there was no effect of frequency on errors, and in the latency data, high frequency words were responded to most quickly, low frequency words had longer latencies, and pseudowords had the longest latencies. These observations suggest that the manner in which the task was performed changed in a way that facilitated the detection of pseudowords. One possibility is that under conditions where decisions are difficult, subjects may also use orthographic familiarity information to detect pseudowords. This additional information would be useful only for pseudowords since the correct response for a pseudoword is always NO but the correct response for a word depends on its meaning.

The results of Experiment 3 were again consistent with Van Orden's account of a spelling check procedure, since subjects made more errors on homophone foils when they had low frequency exemplars than when they had high frequency exemplars. The prolonged NO decision latencies, particularly for words, suggest that subjects were performing this check more carefully than in Experiment 2 because of the higher proportion of homophones. The failure to observe a significant difference between YES latencies on exemplars and false positive latencies on foils is also consistent with Van Orden et al.'s (1988) observations. They claimed that this was evidence that orthographic information does not contribute to YES responses. However, there was a 94 ms difference between low frequency foils and matched low frequency exemplars that was very likely not significant because it was based on only 24 responses. This observation suggests that it would be premature to make the conclusion that there is no influence of orthography from the present data. Further evidence from a much larger number of false positive errors is needed.

In sum, the results of Experiment 3 suggest that when a high proportion of homophones are included in a semantic decision task, it is performed more slowly and performance on pseudowords is better relative to words, but the pattern of responses on words changes very little. These results suggest that phonological recoding cannot be strategically controlled, contrary to previous claims (Coltheart, 1978; Davelaar et al., 1978; Hawkins et al, 1976; McQuade, 1981). Coltheart (1978) has argued that the visual route is usually faster than the phonological route. If including a high proportion of homophones in the experiment forces subjects to abandon phonological recoding, Coltheart's view suggests that latencies should have become faster, not slower. In addition, if subjects were using a visual recognition strategy, no effect of homophony would be expected; however this effect was observed for low frequency foils with low frequency exemplars. The results of Experiment 3 suggest that the effect of including many homophones was not to change the code subjects used to access meaning, but rather it extended the orthographic checking process. This, perhaps in combination with an orthographic familiarity strategy, made pseudowords more likely to be detected

Prior evidence for the ability to avoid the use of phonology came from tasks that do not require that meaning be accessed. Hawkins et al. (1976) employed a tachistoscopic task which can be performed using the orthographic pattern. Davelaar et al. (1978) and McQuade (1981) used lexical decision which, as discussed in the introduction, may be performed on the basis of orthographic familiarity (Balota & Chumbly, 1984; Besner et al., 1984; Gordon, 1983; Seidenberg & McClelland, 1989; Waters & Seidenberg, 1985). Thus the lack of phonological effects found in these studies when a high proportion of homophonic stimuli were included may not reflect the absence of phonological mediation of the access of meaning but rather may reflect a switch to a superficial orthographic strategy. No firm conclusion regarding the use of phonology in the access of meaning can be made, then, from these studies. The category decision task, on the other hand, requires that the meanings of words be accessed, and thus is more suited to provide evidence about phonological mediation. The results of Experiment 3 suggest that phonological mediation is not an optional strategy for low frequency words.

Experiment 4:
The Effect of Spelling Similarity

The finding in Experiment 3 that subjects did not abandon phonological recoding when the experiment contained a high proportion of homophones was surprising given previous claims that subjects can strategically control their use of phonological information. The conclusion that phonological recoding was occurring was based on the observation that subjects produced more false positive errors on low frequency foils with low frequency exemplars than on spelling controls. Experiment 4 was designed to replicate this result with a larger number of stimuli. Only 12 words were included in each group in Experiments 1-3, primarily because of the difficulty finding high frequency homophones. More homophones are available if only low frequency pairs are required. Fourteen new homophones were added in Experiment 4. Two words, FLEA and POLE, that had produced more errors than other homophone foils in Experiments 1-3 were not included to ensure that these words were not unduly affecting the results.

In addition, Experiment 4 included a spelling similarity manipulation. That is, it was also designed to determine whether more false positive errors are made on homophone foils such as ALTER that are very similar in spelling to their exemplars (ALTAR) than on homophone foils such as SLAY that are much less similar to their exemplars (SLEIGH). Van Orden (1987) argued that if a spelling check is performed on phonologically activated candidates, then more errors should be made when exemplars and foils are similarly spelled. A phonological impostor (the exemplar) should be more likely to slip by the spelling check if its spelling is similar to the target foil than if its spelling is dissimilar. He found a larger difference between homophone foils and spelling controls when exemplars were spelled similarly to their foils. This effect should be especially strong in Experiment 4 if the consequence of including a large proportion of homophones is to force subjects to perform this check more carefully.

In sum, if the results of Experiment 3 replicate and phonological recoding cannot be prevented, a significant difference between homophone foil and spelling control errors should be found. And, if a spelling check is performed, more errors should be made on foils that are similar to their exemplars than on dissimilar foils. The homophone control condition was not included in Experiment 4 because all three previous experiments demonstrated that the higher error on homophone foils could not be attributed to a more general difficulty in processing homophones.

Method

Subjects. Twelve McGill University undergraduates were paid $5 each to participate in the study. All were native speakers of English.

Stimuli. As in Experiments 2 and 3, the category names used were LIVING THING and OBJECT. There were 72 experimental words (these are presented in Appendix B) and 198 filler words. Twenty-four of the experimental words were homophone exemplars, 24 were the other member of the homophone pair and served as homophone foils, and 24 were spelling controls. All were low frequency words. Twelve of the exemplar-foil-spelling control triples had foils and spelling controls that were spelled similarly to their exemplars (e.g. ALTAR-ALTER-AJAR). The mean similarity of the foils to the exemplars (using Weber's, 1970, measure) was .72 and the mean similarity of the spelling controls to the exemplars was .76. The remaining 12 triples had foils and spelling controls that had spellings that were not similar to their exemplars (e.g. SLEIGH-SLAY-SLAM). The mean similarity of the foils to the exemplars was .52 and the mean similarity of the spelling controls to the exemplars was .50. (Theoretically, values on this measure range from 0 to 1, but few homophone pairs have similarities less than .40 or greater than .80). Exemplars, foils, and spelling controls were matched for frequency and length. The mean frequencies, respectively, were 4.3, 4.6, and 6.3 for the similarly spelled triples and 5.3, 5.8, and 5.9 for the dissimilarly spelled triples. Half of the exemplars in each group were members of the category LIVING THING and half were members of the category OBJECT. Since the homophone foil condition was not included in this experiment only three lists were needed. Each member of a stimulus triple (exemplar, foil, spelling control) was placed on a different list, with the result that 24 experimental words, eight of each category-target relation, appeared on each list.

Sixty-six filler words also appeared on each list, and 56 of these were homophones. Sixteen of the homophone fillers were LIVING THINGS, 16 were OBJECTS, and 24 were not exemplars of their categories. Five of the nonhomophone fillers were members of their categories and five were not members. Thus, of the 90 words on each list, 45 were preceded by the category LIVING THING and 45 were preceded by the category OBJECT, and within each category, half of the words were exemplars and half were not. Each list was composed of 80% homophones and 20% nonhomophones. No word appeared more than once in the experiment, and no pseudowords were included. Sixteen other words were chosen for a practice list. Half of the words were preceded by the category LIVING THING and half were preceded by the category OBJECT, and half of each were members of their categories.

Procedure. The viewing conditions and procedure were exactly the same as in Experiments 1-3 except that the experiment required only three 15 minute sessions.

Results

One word triple (object: CORD, CHORD, COD) in the high similarity group was removed from the analysis because 10 of the 12 subjects made an error on the spelling control (COD) and the two others had latencies of longer than 1500 ms. This item was ambiguous because it is a LIVING THING, but most people see it in a frozen package at the supermarket and so it could also be an OBJECT.

There were two factors in the analyses of the NO data, orthographic similarity (similar vs. dissimilar) and category-target relationship (foil vs. spelling control). The data were analyzed as in the previous experiments.

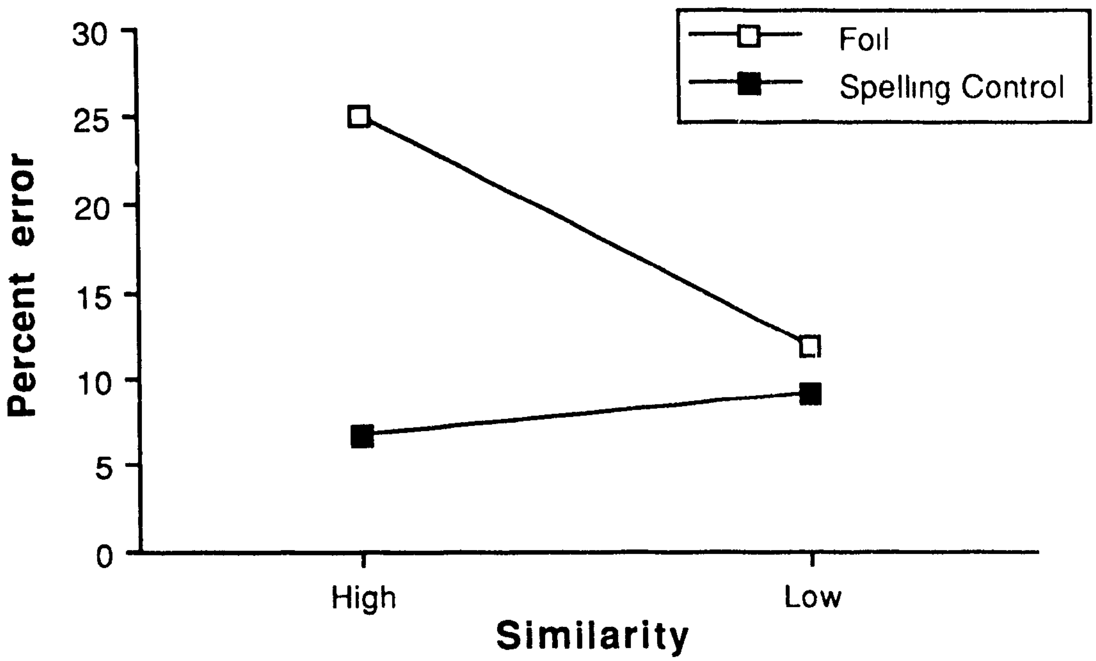
False Positive Error Data. The differences in percent errors between homophone foils and spelling controls are presented in Figure 7. There was a main effect of category-target relation that was significant by subjects, $F(1,11) = 9.46$, $p < .01$, and was marginally significant by items, $F(1,42) = 3.88$, $p = .056$ Subjects made more errors on homophone foils (18.4%) than on spelling controls (7.9%). The interaction of category-target relation and similarity was significant by subjects, $F(1,11) = 6.95$, $p < .05$, but not by items, $F(1,42) = 1.88$, $p > .05$. Simple main effects revealed a significant difference between homophone foils and spelling controls for foils that are spelled similarly to their exemplars (18.2%), $F(1,11) = 13.12$, $p < .01$ by subjects and $F(1,42) = 5.35$, $p < .05$ by items, but not for dissimilarly spelled words (2.8%), both F's < 1.

NO Reaction Time Data. A subject's response latency was only included in the analyses if the subject responded NO correctly to it and the other two members of the stimulus triple. This criterion was met by 66.7% of responses Another 19.0% of responses were correct but were discarded because the subjects made an error on another member of the stimulus triple. A total of 5.6% of response times were greater than 1500 ms and were replaced with times of 1500 ms.

The main effect category-target relationship was not significant, both F's < 1, nor was the interaction of category-target relationship and similarity, both F's < 1. Subjects took 20 ms longer to make decisions about homophone foils than spelling controls when they were spelled similarly, and 35 ms longer when they were spelled dissimilarly.

YES Response Latencies. An item was included in these analyses only if a subject made both a false positive error on the homophone foil and correctly accepted the corresponding exemplar. One out of 40 false positive errors on homophone foils was not included because the subject failed to respond correctly to the corresponding exemplar. There was no difference between YES latencies for homophone foils (898 ms) and exemplars (896 ms), both t's < 1. The difference for the similarly spelled group alone

Figure 7: Mean percent errors for foils and spelling controls in Experiment 4 as a function of their similarity to corresponding exemplars.



Foil
Spelling Control
Percent error
30
25
20
15
10
5
0
High
Low
Similarity

was not significant either, both t's < 1.

Discussion

Low frequency homophone foils with low frequency exemplars produced more errors than spelling controls in Experiment 4. Because additional words were included in this experiment, it suggests that the effect found for this group in Experiment 3 was not specific to the small set of words used. This is further evidence for the claim that the use of phonological information cannot be strategically controlled. If subjects had been able to prevent phonological recoding, then there should have been no difference between homophone foils and spelling controls. Experiment 4 also demonstrated that the difference in error rates on homophone foils and spelling controls is larger when the foil and exemplar are similarly spelled. This suggests that subjects were performing a spelling check on phonologically activated candidates, and that the spelling check was less likely to catch a phonological impostor if its spelling was similar to the target homophone foil. Thus, Experiments 1-4 have demonstrated that homophone foils do not in general produce more errors than spelling controls. Rather, the effect is limited to low frequency homophone foils that are spelled similarly to their low frequency exemplars.

Implications of Experiments 1-4 for Word Recognition Theories

According to Van Orden (1987), targets are phonologically recoded, and the phonological representation activates candidate lexical entries The most active of the candidates is submitted to a spelling check Its spelling is compared to that of the target. False candidates are more likely to pass the spelling check if their spelling is unfamiliar and/or they are spelled similarly to the target. This proposal predicts that more false positive errors should be made on homophone foils than on spelling controls when foils have low frequency exemplars and/or when exemplars are spelled similarly to the target foils. The results of Experiment 1-4 support these predictions The observation that the use of phonological information is not strategically controlled by the subject is also consistent with his view.

A further claim that Van Orden (1987) made was not supported by the results of the present experiments. He argued that both high and low frequency words are phonologically recoded, yet no evidence for the phonological mediation of the meanings of high frequency words was found in Experiments 2 and 3 when priming from the category name was avoided. However, when examined more closely, his view does not predict a difference in the false positive error rate between high frequency homophone foils and spelling controls, although it does predict a small difference in the NO latency data In the case of high frequency foils with low frequency exemplars, the foil would be submitted to the spelling check first, and because its spelling is familiar, it should easily pass. Thus there should be no more errors and no longer latencies than for spelling controls In the case of high frequency foils with high frequency exemplars, the exemplar would be submitted to the spelling check first when it was higher in frequency, but because its spelling is familiar, the spelling check should easily detect that it is not the target presented The foil would be submitted next, and again an error should not be made because its spelling is familiar. Thus for this group, the theory would predict longer latencies for foils than for spelling controls and no effect in the error data. The effect in the latency data would be weak because only the exemplars that are higher frequency than the foil would be chosen for the spelling check first. The only case where Van Orden's view predicts more errors on foils than on spelling controls is when the spelling of both the foil and the target were not familiar. This is what was observed in Experiments 2 and 3.

A consequence of the above analysis is that a comparison of false positive error rates on homophone foils and spelling controls is not capable of determining whether the meanings of high frequency foils are phonologically mediated and the spelling check is efficient, or whether they are accessed on a visual basis. Some evidence might be obtained from the NO latency data, although a problem with making inferences from this data was pointed out earlier. NO latencies may include post-recognition semantic processing that may mask small recognition effects. There was no difference between high frequency homophone foils with high frequency exemplars and their spelling controls in the NO latency data either in Experiment 2 or 3. Van Orden's view would also predict a delay in decision latencies for low frequency foils with high frequency exemplars, and this was also not observed. In both Experiments 2 and 3. latencies for these foils were actually slightly faster than for spelling controls.

Van Orden et al. (1988) argued that the observation of similar YES latencies on exemplars and homophone foils indicated that the exemplars did not benefit from having orthographic information consistent with the YES response. In Experiments 2-4 there was no significant difference between foil and exemplar YES latencies, however it was argued that strong conclusions could not be drawn from this data because incorrect YES latencies may arise from an earlier point in processing than correct YES latencies, and because the analyses were based on too few items. Van Orden et al.'s (1988) analyses were also based on a small number of responses. In their Experiment 1, the analysis of word latencies was based on an average of 2.1 scores per subject, and in their Experiment 2, it was based on an average of 3.3 scores per subject. Thus, it is unclear from these data whether there is orthographic activation of meaning. There was a hint from Experiment 2 that some orthographic information contributed to the activation of word meanings in Van Orden's (1987, Expt 2) tachistoscope experiment. Evidence that specific category names prime exemplars suggests that the false positive error rate should have been much higher in his experiment if no orthographic information was available.

In sum, the results of Experiments 1-4 have provided some support for Van Orden's (1987) theory. There was evidence for phonological mediation for low frequency words and there was evidence for the use of a spelling check. Further data on high frequency words is required to determine whether he is correct in saying that their meanings are phonologically mediated. Also, more evidence is required to determine whether orthographic information directly activates word meanings If such evidence were found, it would contradict Van Orden's (1987) hypothesis that bottom-up activation is exclusively phonological.

The dual route view, in contrast, assumes that skilled readers quickly activate the meanings of high frequency words on a visual basis, and therefore correctly predicts no difference between high frequency homophone foils and spelling controls For low frequency words, activation from orthography to meaning is presumed to be slower and/or weaker (although not in Coltheart's version of the theory), and so phonological information can influence the activation of the semantic representation (Seidenberg, 1985, Seidenberg & McClelland, 1989). This view correctly predicts a difference between homophone foils and spelling controls for low frequency words Because the meanings of low frequency words are activated from orthography and phonology, it can also account for the finding that more errors are made on homophones foils that are spelled similarly to their exemplars. When low frequency foils are presented (e.g. ALTER, SLAY), the meanings associated with similarly spelled exemplars (e.g ALTAR) will receive activation from both orthographic (e.g. A,L,T,R) and phonological codes but dissimilarly spelled exemplars (e.g. SLEIGH) will receive activation primarily from phonological codes. Thus similarly spelled exemplars will be better able to compete with the foil. Finally, dual route theory provides an explanation for Van Orden's (1987, Expt. 2) tachistoscope experiment results False positive errors on homophone foils would be less than expected given priming of the exemplar if there is activation of the foil's meaning from orthography.

Dual route theory does, however, have difficulty accounting for the larger number of errors made on foils with low frequency exemplars than those with high frequency exemplars. High frequency exemplars should be activated more strongly than low frequency exemplars, and thus should be better able to compete with the foil. To account for effects of exemplar frequency, a spelling check could be added that is used when two or more words are highly activated. This could involve waiting until more evidence accumulates from the orthographic route, examining the output of the orthographic route alone, or examining the orthographic code associated with each meaning to see if it matches the input orthographic code. The spelling check would be more effective for high frequency exemplars than for low frequency exemplars. In each case, however, a delay in processing would be expected for homophones relative to spelling controls when they are low in frequency, and no such effect was observed. The finding that subjects cannot strategically control the use of phonological information contradicts Coltheart's (1978) view, but is consistent with Seidenberg and McClelland's (1989) dual route model.

In sum, dual route theory also accounts for some of the evidence in Experiments 1-4, but to account for the effect of exemplar frequency, a spelling check had to be added on. Dual route theories do not assume that a spelling check normally occurs in word recognition because activation of meaning by orthography is usually sufficient to distinguish between members of a homophone pair.

Neither Van Orden's (1987) theory nor dual route theory can account for all of the findings of Experiments 1-4. Further evidence from high frequency words is examined in

Experiments 5 and 6 to test the contradictory views these theories have about phonological mediation of the meanings of high frequency words. In addition, Experiments 5 and 6 examine whether the spelling check is a necessary component of word recognition as Experiments 1-4 and Van Orden's (1987) theory suggest, or whether it is a special strategy as dual route theory suggests.

EXPERIMENTS WITH HOMOPHONE EXEMPLAR TARGETS

Experiment 5:

The Effect of Priming the Foil

The results of Experiments 1-4 suggest that access to the meanings of low frequency words is phonologically mediated, and that a spelling check is performed on phonologically activated candidates. However, it is possible that the spelling check is a strategy used by subjects when the experiment contains homophone foil trials. Orthographic information from the foil target may give its semantic or lexical representation a small advantage over that of the exemplar, but when subjects become aware that some targets sound like an exemplar of the category but are not exemplars, they may check their spellings carefully to prevent errors. Van Orden (1987) acknowledged that his findings could not determine whether the spelling check is always performed or is only performed when the experiment contains homophone foils. If a spelling check is performed only when the experiment contains homophone foils, this would imply that there is usually enough orthographic information available to distinguish between members of a homophone pair.

Experiment 5 was designed to examine whether subjects use a spelling check when no homophone foil trials are included in the experiment. Instead of examining performance on homophone foils, responses on homophone exemplars were examined. Van Orden's view still predicts that the spelling check will occur, and that subjects will make fewer errors on exemplars when the other member of the homophone pair is high in frequency. When the spelling of the other member is familiar, the spelling check should be better able to detect that it does not match the exemplar target than when its spelling is unfamiliar. The dual route view, on the other hand, predicts that subjects should be more likely to make an error on a homophone exemplar if the other member is high in frequency. The meaning of a high frequency foil would be activated more strongly than that of a low frequency foil, and thus would be better able to compete with the meaning of the target exemplar. (I'll continue to use the term foil to refer to the member of the homophone pair that is not an exemplar of the category).

In order to be able to examine effects of the frequency of the foil in skilled readers who make few errors on target words that are correct exemplars of the category, Experiment 5 compared performance on homophone exemplars with performance on these exemplars when the foil was made more salient by priming it with a semantically related word. That is, in one condition exemplars were preceded by a word semantically related to the foil (e.g. SHATTER-BRAKE) and in the other they were preceded by an unrelated word (e.g. BOLD-BRAKE). If a spelling check is carried out, priming the foil should make it more likely to be checked first, and foils with unfamiliar spellings are more likely to erroneously slip by. According to dual route theory, increasing the activation of the foil should cause subjects to make more errors or produce longer decision latencies on homophone exemplars with high frequency foils.

Another advantage of examining YES responses on homophone exemplars is that this method may be more likely to reveal an effect of phonology for high frequency words, if such an effect exists. In the previous experiments, evidence for the use of phonology came primarily from error responses, and subjects did not make many errors on high frequency words. It was argued that the failure to find effects for high frequency words could occur either because access to their meanings was phonologically mediated but the spellings were well known so the spelling check did not err, or because their meanings were accessed directly on the basis of visual information. There were no clear effects of homophony for high frequency words in the NO latency data, but time to decide that a word is not an exemplar of a category may include considerable processing after the meaning of the word had been obtained. Thus, correct YES responses should be a more sensitive measure of effects of phonology. If both the correct and foil meanings of high frequency homophones are activated to some extent, then increasing the activation of the foil meaning should cause subjects to produce longer decision latencies or make more errors on the homophone exemplars. If subjects quickly determine the meanings of homophones before semantic representations receive activation from phonological representations, then there should be no effect of the priming manipulation.

Underwood and Thwaites (1982) conducted a lexical decision experiment that also made use of pairs of words in which one was a homophone (e.g. WAIST) and the other was a word semantically related to the other member of the homophone pair (e.g. RUBBISH). The two words were presented simultaneously; the semantically related words appeared in the same location as the fixation point had been, and the homophones were presented to the right of fixation and were pattern masked. The subjects' task was to decide whether the stimulus that appeared centrally (e.g. RUBBISH) was a word or not. They found that response latencies to central words were slowed when the word in the periphery was a homophone (e.g. WAIST) whose other member was related to the target as compared to a condition when the word in the periphery was an unrelated homophone. Underwood and Thwaites attributed this effect to the use of phonological information in decision processes since the peripherally presented word would usually be processed after the centrally presented target word. In Experiment 5, the word semantically related to the foil was presented prior to the homophone and thus could influence early processing of the homophone.

Preliminary Study

Before the main Experiment 5 was conducted, a preliminary study was run. This was needed to ascertain that the semantically related words used actually did prime the foil (e.g., that SHATTER primes BREAK). An experimenter could choose words that are semantically related to the foil, but these might not necessarily cause subjects to respond faster to to the foil when they precede it. In the preliminary study, two semantically related words were chosen for each foil (e.g. SHATTER, FRACTURE). Foils were presented to subjects three times, in separate sessions, once preceded by each semantically related word, and once preceded by an unrelated word. The semantically related word that caused the greatest decrease in response latency for each foil was chosen as the prime for the main experiment. An analysis was then done to ensure that these primes did indeed result in significantly faster response latencies on foils than when foils were preceded by an unrelated word.

Neely, Keefe, and Ross (1989) have argued that semantic priming effects could be due either to automatic spreading activation or to strategies such as the generation of expectancies and post-lexical semantic matching, depending on how the task is set up. To reduce the likelihood that priming effects were due to subject strategies, a continuous semantic decision task was used. Words were presented one at a time and subjects made decisions to each item. Primes appeared as a normal trial prior to the target trial. Subjects were not told that some successive trials were related. Further, the proportion of related trials was kept low by including filler words unrelated to preceding words (see Neely et al., 1989; Seidenberg, Waters, Sanders, & Langer, 1984; Tweedy et al., 1977), and none of the primes were high associates of the targets (see Fischler, 1977).

Method

Subjects. Eighteen McGill University undergraduates were paid $8 each to participate in the study. Another 15 subjects volunteered to fill out a questionnaire. All were native speakers of English.

Stimuli. The target stimuli were the homophone word foils that were used in Experiments 1-3 (e.g. object-BREAK). The pseudohomophone foils (e.g. object-SHEWS) were not included because in the main study the exemplars were presented, and their exemplars are not homophones (e.g. SHOES). For each of the 48 foils, two semantically related words were chosen to serve as primes (e.g. SHATTER and

FRACTURE for BREAK). Three experimental lists were created with each LIVING THING foil appearing once on each list, and three lists were created with each OBJECT foil appearing once on each list. Across the three lists, foils appeared once preceded by an unrelated word, once preceded by one of the semantically related words, and once preceded by the second semantically related word. These trials were distributed among the lists such that on 8 of the 24 experimental trials on each list the foil was preceded by an unrelated word, and on 16 of the trials the foil was preceded by a semantically related word. An additional 112 fillers were included on each list; 80 were exemplars of the category and 32 were not exemplars. The fillers reduced the percentage of related trials to 10% of the trials on a list. Thus, each of the six lists contained 80 YES trials and 80 NO trials. The fillers were placed randomly on each list, although not between prime-target pairs. The category names did not appear on the lists since the same decision was to be made about all words on a given list. Two practice lists were created, one for the LIVING THING category and one for the OBJECT category. Half of the items on each list were members of the category and half were not.

A questionnaire was created that listed each of the 96 semantically related prime words with a blank line beside each one. Instructions on the top of the page asked subjects to write down the first related word that came to mind for each of the words on the list. This data was collected to ensure that the targets were not highly associatively related to the primes (e.g. BREAD-BUTTER).

Procedure. The questionnaire was distributed to 15 subjects who were asked to fill in the first related word that came to mind for each word on the list. They were instructed not to spend too much time on any item. The 18 other subjects completed three experimental sessions, each lasting 30 minutes. The sessions were separated by at least one week. In a session, the subjects were told the name of a category and saw the practice trials followed by a list of 160 trials for that category, and then were told the name of the other category and saw the practice trials and a list of 160 trials for that category. Subjects were given feedback on each practice trial to ensure that they understood the task. Subject saw each of the six lists once. The order of presentation of the experimental lists was counterbalanced across subjects such that each list was presented in each of the six list positions three times. Words were presented on the screen one at a time and remained until the subject responded. Subjects were not told that words on some trials were semantically related to words on the subsequent trial. The intertrial interval was 1s. Other aspects of the procedure are the same as in the previous experiments.

Results and Discussion

The 2.8% of response times that were greater than 1500 ms were replaced with times of 1500 ms. Item means were then calculated for foil targets (e.g. object-BREAK) preceded by each of their semantically related primes (e.g. SHATTER, FRACTURE). The prime that was associated with the shortest response time on the succeeding foil was chosen. These are presented in Appendix C. The questionnaires were examined to see how often subjects produced the foil target given the chosen prime word. Targets were produced on 32 occasions out of a possible 720 (4.4%). In one case (ENCOUNTER) 6/15 subjects produced the foil (MEET), in three cases 4/15 subjects produced the foil, and in the nine other cases three or fewer of the subjects produced the foil. Thus, the chosen semantically related prime words are not highly associatively related to the foil targets.

Responses on foil targets when they were preceded by an unrelated word and when they were preceded by the chosen prime word were submitted to analyses of variance There were three factors in the analyses, prime condition (primed vs unprimed), foil frequency (high vs. low), and exemplar frequency (high vs. low). They were all treated as within factors in the analyses by subjects. In the analysis by items, prime condition was treated as a within factor (since exactly the same words were used in both conditions) and the other two factors were treated as between factors. Percent errors were arcsine transformed prior to analysis.

There was a significant effect of prime condition in the error data, $F(1,17) = 13.04$, $p < .01$ by subjects and $F(1,44) = 10.65$, $p < .01$ by items. Subjects produced fewer errors on foils when they were preceded by a semantically related prime (6.6%) than when they were preceded by an unrelated prime (10.8%). There was no significant interaction between prime condition and frequency of the foil, $F(1,17) = 1.83$, $p > .05$ by subjects and $F(1,44) = 2.58$, $p > .05$ by items, between prime condition and frequency of the exemplar, both F's < 1, and no significant triple interaction, $F(1,18) = 1.71$, $p > .05$ by subjects and $F(1,44) = 1.94$, $p > .05$ by items. The size of the priming effect was 3.7% for high frequency foil with high frequency exemplars, 3.7% for high frequency foils with low frequency exemplars, 1.0% for low frequency foils with high frequency exemplars, and 8.3% for low frequency foils with low frequency exemplars.

There was a significant main effect of prime condition in the decision latency data, $F(1,17) = 18.89$, $p < .001$ by subjects and $F(1,88) = 7.30$, $p < .01$ by items. Subjects produced faster responses on foils when they were preceded by a semantically related prime (709 ms) than when they were preceded by an unrelated prime (756 ms). There was no interaction between prime condition and frequency of the foil, between prime condition and frequency of the exemplar, and no significant triple interaction, all F's < 1. The size of the priming effect was 48 ms for high frequency foils with high frequency exemplars, 45 ms for high frequency foils with low frequency exemplars, 48 ms for low frequency foils with high frequency exemplars, and 45 ms for low frequency foils with low frequency exemplars.

The preliminary study served to select a semantically related word (e.g. SHATTER) for each homophone foil (object-BREAK), and has demonstrated that these words do in fact prime the foil in a continuous semantic decision task. Since prime and target trials were not presented as pairs, and a low proportion of the trials were related, it is very likely that these effects were not due to task-specific strategies.

The words found to prime the foils in the preliminary study were then used as primes for the exemplars in a similar continuous semantic decision task in the main experiment. For example, in the preliminary study, subjects had to judge whether SHATTER was an object (NO) and then whether BREAK was an object (NO). In the main experiment, subjects had to judge whether SHATTER was an object (NO) and then whether BRAKE was an object (YES). If access to meaning is phonologically mediated, then both meanings of a homophone should be available, and priming the incorrect meaning should increase decision errors and/or latencies on exemplars. Further, according to Van Orden's account, priming should have a greater effect on exemplars with low frequency foils, whereas dual route theory predicts that priming should have a greater effect on exemplars with high frequency foils. On the other hand, if access to meaning occurs prior to the influence of phonology, preceding exemplars with primes related to the foil should have no effect.

Main Experiment

Method

Subjects. Twenty-four McGill University undergraduates were paid $5 each to participate in the study. None had participated in the pilot study. All were native speakers of English.

Stimuli. The target homophones were the 48 homophone exemplars of the categories LIVING THING and OBJECT that were used in Experiments 1-3 (e.g object-BRAKE). The 48 words chosen in the pilot study that were semantically related to the foils served as primes. An additional 48 words, each unrelated to one of the exemplars, were chosen to serve as unrelated controls. All of the primes and unrelated controls required a NO response. The experimental words are presented in Appendix C.

Two lists were created for each category; both lists for a category contained the 24 homophone exemplars belonging to that category. On one list they were preceded by their prime and on the other they were preceded by their unrelated control. On each list, half of the exemplars were preceded by their prime and the other half were preceded by their unrelated control. An additional 112 filler trials were included on each list, 56 were exemplars of the category and 56 were not. They were placed randomly on the lists but not between a prime-target pair. Thus, there were two LIVING THING lists and two OBJECT lists, each with 160 trials, 80 of which were YES trials and 80 of which were NO trials. Only 7.5% of the words on a list were homophone exemplars preceded by a word semantically related to the other member of the homophone pair. Two practice lists were created, one for the LIVING THING category and one for the OBJECT category. Half of the items on each list were members of the category and half were not.

Procedure. The subjects completed two experimental sessions, each lasting 30 minutes. The sessions were separated by at least one week. In a session, the subjects were told the name of a category and saw the practice trials followed by a list of 160 trials for that category, and then were told the name of the other category and saw the practice trials and a list of 160 trials for that category. Subjects were given feedback on each practice trial to ensure that they understood the task. Subject saw each of the four lists once. The order of presentation of the experimental lists was counterbalanced across subjects such that each list was presented in each of the four list positions six times. Words were presented on the screen one at a time and remained until the subject responded. Subjects were not told that words on some trials were semantically related to words on the subsequent trial. The intertrial interval was 1s. Other aspects of the procedure are the same as in the previous experiments.

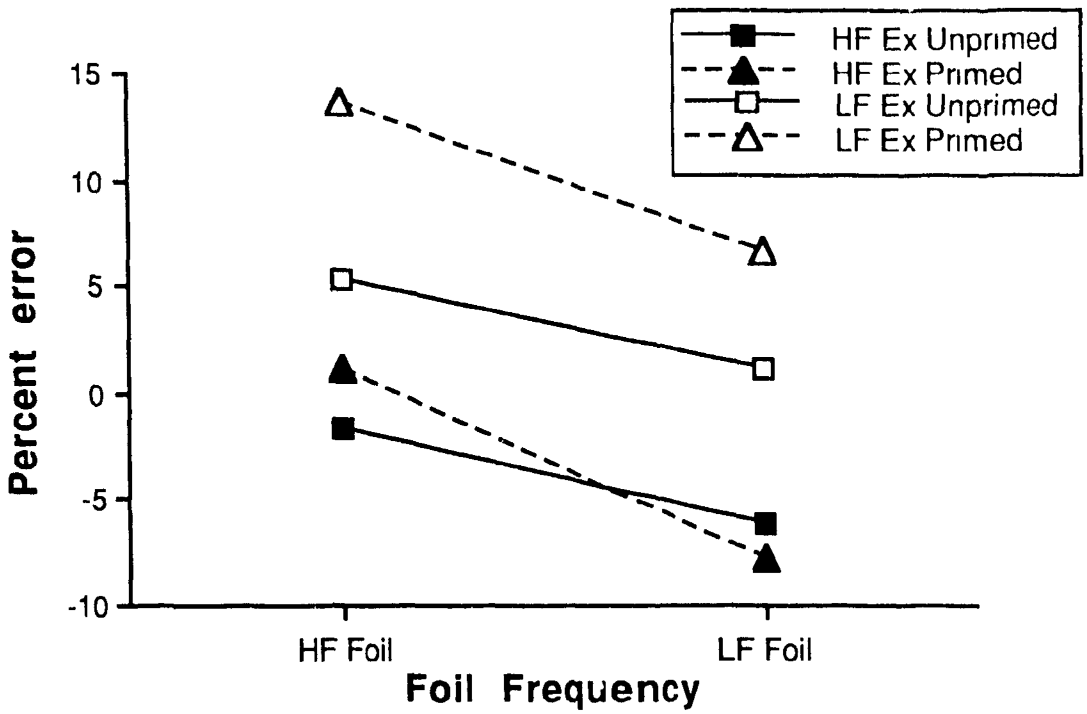
Results

Of the 640 trials in the experiment, only data from the 96 homophone exemplar trials were analyzed. Prior to analysis, the percent error data were arcsine transformed. A subject's response latency was only included in the analyses if the subject responded correctly to the exemplar in the primed and unprimed conditions. This criterion was met by 87.3% of responses. Another 4.5% of responses were correct but were discarded because the subjects made an error on the exemplar in the other prime condition. A total of 1.2% of response times were greater than 1500 ms and were replaced with times of 1500 ms. There were three factors in the analyses, prime condition (prime vs. unrelated control), frequency of the exemplar (high vs. low), and frequency of the foil (high vs. low). These were all treated as within factors in the subject analysis. In the item analysis, prime condition was treated as a within factor and the other two were treated as between factors

The mean percent errors for primed and unprimed exemplars are presented in Figure 8. There was a main effect of prime condition in the error data, $F(1,23) = 12.55$, $p < .01$ by subjects and $F(1,44) = 9.01$, $p < .01$ by items. Subjects made 6.6% errors on unprimed exemplars and 9.9% errors on primed exemplars. The main effect of exemplar frequency was significant by subjects, $F(1,23) = 11.33$, $p < .01$, but not by items, $F(1,44) = 2.26$, $p > .05$. Subjects made 5.8% errors on high frequency exemplars and 10.7% errors on low frequency exemplars. The main effect of foil frequency was significant by subjects, $F(1,23) = 39.01$, $p < .001$, but not by items, $F(1,44) = 1.72$, $p > .05$ Subjects made 11.8% errors on exemplars with high frequency foils and 4.8% errors on exemplars with low frequency foils.

There was a significant interaction between prime condition and exemplar frequency by subjects, $F(1,23) = 7.17$, $p < .05$, but not by items, $F(1,44) = 2.08$, $p > .05$. Simple main effects tests revealed a significant effect of priming for low frequency exemplars (5.7%), $F(1,23) = 24.37$, $p < .001$ by subjects and $F(1,44) = 9.87$, $p < .01$ by items, but no effect of priming for high frequency exemplars (0.9%), both F's < 1.2. The interaction between prime condition and foil frequency approached significance by subjects, $F(1,23) = 3.46$, $p < .08$, and was significant by items, $F(1,44) = 4.48$, $p < .05$. Simple main effects tests revealed a significant effect of priming for exemplars with

Figure 8: The mean percent errors for primed and unprimed homophone exemplars in Experiment 5.



HF Ex Unprimed
HF Ex Primed
LF Ex Unprimed
LF Ex Primed
15
10
5
0
-5
-10
Percent error
HF Foil
LF Foil
Foil Frequency

high frequency foils (4.7%), $F(1,23) = 12.16$, $p < .01$ by subjects and $F(1,44) = 13.09$, $p < .001$ by items, and an effect of priming for low frequency foils (1.8%) that was significant by subjects, $F(1,23) = 4.23$, $p = .05$, but not by items, $F < 1$. The triple interaction was not significant by subjects, $F(1,23) = 1.33$, $p > .05$, or by items, $F < 1$.

In the decision latency data, there was no main effect of prime condition, both F's < 1. The main effect of exemplar frequency (26 ms) was significant by subjects, $F(1,23) = 9.64$, $p < .01$, but not by items, $F < 1$. The main effect of foil frequency (17 ms) was also significant by subjects, $F(1,23) = 5.67$, $p < .05$ but not by items, $F(1,44) = 1.02$, $p > .05$. None of the interactions were significant.

Discussion

The results of Experiment 5 provide little evidence for the use of phonological information in the access of high frequency words. There was no effect of priming on either decision latencies or errors for these words even though the preliminary study demonstrated that the primes did indeed activate the foils. Had the meaning or lexical entry associated with the foil also been activated via the phonological representation of the exemplar, priming should have increased decision latencies and/or errors.

One could argue that the meanings of high frequency words were phonologically mediated but the spelling check did not err. However, although high frequency foils may not be likely to erroneously slip by the spelling check because their spellings are familiar, priming a low frequency foil should have led to it being chosen for the spelling check first on some occasions, and its less familiar spelling should have resulted in more errors. But, there were actually 1.1% fewer errors made on high frequency exemplars with low frequency foils when they were primed. Also, a delay would be expected when the foil was chosen for the spelling check first, and no effect of priming was observed in the latency data. These results suggest that access to the meanings of high frequency words is not phonologically mediated. In contrast, there was evidence for the use of phonological information in the access of the meanings of low frequency words, since priming the foil did produce more errors on homophone exemplars.

Further, the results of Experiment 5 suggest that a spelling check is not performed when "trick" foil trials (e.g. object-BREAK) are excluded from the experiment. Priming increased errors on exemplars with high frequency foils but had only a small effect on exemplars with low frequency foils. If a spelling check was being performed, the reverse pattern should have been observed. The spelling check should have been better able to avoid errors when the spelling of the nontarget member of the homophone pair was more familiar, as was found in Experiments 1-4 and in Van Orden's experiments. Subjects in Van Orden's experiments and in Experiments 1-4 probably were cautious and performed a spelling check because of the presence of homophone foil trials in the experiments. That is, they may have had enough information to respond without performing the spelling check, but because they were aware of the presence of trick trials in which the target sounded like a member of the category, they sought additional orthographic information before responding. Thus, the results of Experiment 5 imply that the spelling check is not an obligatory step in word recognition, as Van Orden (1987) argues, but rather it is a strategy used to avoid errors in a laboratory experiment.

The results of Experiment 5 support the dual route view that a spelling check is not a normal part of the word recognition process. For high frequency words, meaning is activated directly from orthography, with little effect of phonology. For low frequency words, the combination of activation from orthography and phonology is usually enough to prevent an error on a low frequency word since the other meaning of the homophone will receive activation from the phonological representation only, although decision latencies may be slowed by the presence of a second active semantic representation. However, when the foil of a homophone is primed, its activation level may exceed that of the target, particularly if the the foil is a common word. Priming of the foil has no effect on high frequency words because the meanings of these words are strongly activated from orthography prior to activation from their phonological representations.

Experiment 6:

The Effect of Homophony

Evidence from Experiments 1-5 suggests that the access of meanings is phonologically mediated for low frequency words and not for high frequency words. However, the same set of homophones pairs was used in all of the experiments (except for the few added in Experiment 4), and so the possibility exists that this result is specific to the categories and items used. In Experiment 6, a new category (VERB) and set of words was chosen in order to examine whether the finding of phonological effects only for low frequency words will replicate.

The design of the experiment was also different than in previous experiments. In Experiments 1-4 performance on homophone foils was compared to performance on spelling and homophone controls, and in Experiment 5, performance on exemplars when they were unprimed was compared to performance on exemplars when they were primed. In this experiment, performance on YES responses to exemplars (e.g. verb-MEET) was compared to performance on semantically similar nonhomophone words (e.g. verb-JOIN). If the access to meaning is phonologically mediated, subjects should produce longer decision latencies and/or more errors on homophones than on nonhomophone semantic controls. This design more closely approximates natural reading since it does not use homophone foils and there is no priming of the other member of the homophone pair. Observing YES decision latencies should allow the detection of small effects of homophony if they occur. The logic of the design is the same as that of the original homophony studies conducted by Rubenstein et al. (1971), and Davelaar et al. (1978), but the use of the semantic decision task rather than the lexical decision task ensures that meaning has been accessed, and avoids the problem of choosing the appropriate pseudoword distractors.

Homophone exemplars and semantic controls were matched on typicality of category membership, frequency, and length. Shoben (1982) has argued that these three factors must be controlled for when testing for an effect of another variable on YES trials in a semantic decision experiment. In addition, homophone exemplars and semantic controls were also matched on the frequency with which they are used as verbs. This was necessary because many verbs can also be used as nouns (e.g. to MEET, a track MEET). It was impossible to avoid these verbs altogether; however if homophone and control verbs are matched on their printed frequency and also on the frequency with which they are used as verbs, any differences that are observed in performance on these items cannot be attributed to ambiguity differences. No attempt was made to have semantic controls also be spelling controls as in Experiments 1-4. Spelling controls are not necessary as they were in those experiments because the trials of interest were not foil trials. With foils, subjects could falsely indicate that a foil was a member of the category because of a visual similarity to the exemplar. But subjects would be equally likely to falsely indicate that homophone or nonhomophone exemplars were not members because they looked like another nonmember word. Homophone exemplar targets would only be influenced by their visual similarity to the foil if the foil received activation from the phonological representation of the exemplar.

Method

Subjects. Twenty-five McGill University undergraduates were paid $3 each to participate in the study. An additional 25 subjects volunteered to fill out a typicality ratings questionnaire. All were native speakers of English.

Stimuli. There were 160 experimental words in the study. Half were homophones and half were semantically matched nonhomophones (see Appendix D). All experimental words were members of the category VERB (e.g. MEET, JOIN), and the other members of the homophone pairs were not usually used as verbs (e.g. MEAT). The homophones fell into four groups, high frequency exemplar/high frequency foil, high frequency exemplar/low frequency foil, low frequency exemplar/high frequency foil, and low frequency exemplar/low frequency foil. The semantically similar controls were matched as closely as possible to the homophones for printed word frequency, the frequency with which they appeared in the Francis and Kucera (1982) count as verbs, typicality, and length (see Table 3).

Typicality ratings for the 160 experimental words were obtained using a questionnaire based on those used by Rosch (1975). The 160 words were listed on two pages, and beside each word was a 7-point scale. Instructions on a separate page were a paraphrase of those used by Rosch (1975), and they asked subjects to indicate for each word how good an example of the category they felt the item was by circling a number on the scale. A 1 indicated a very good example and a 7 indicated a very poor example. The instructions included an illustration of what for me is a good (JUMP) and poor example (BELIEVE) of the category VERB. Subjects were encouraged to make their own judgments. These data were collected to ensure that the homophone exemplars and semantic matches were equally good exemplars of the category VERB, so that any differences could be attributed to differences in homophony. None of the differences between the mean typicality ratings of homophone exemplars and semantic controls for any of the four groups were significant (all t's < 1.5, p's $> .15$).

Two lists were created. Half of the homophone exemplars from each group appeared on each list. The semantic controls were placed on the lists such that they did not appear on the same list as their matched homophone exemplar. An additional 240 filler trials were included. Of these, 40 were exemplars of the category VERB and 200 were not verbs. The order of presentation of stimuli on each list was random. The lists thus had 200 trials each, half of which required a YES response and half required a NO response. A homophone was presented on 20% of trials. Another 16 words, half of which were verbs, were chosen to serve as practice stimuli.

Procedure. The subjects completed one experimental session lasting 30 minutes.

Table 3: Frequency, typicality, and length statistics for the words used in Experiment 6.

Group	Frequency			Typicality	Length
	Exemplar (word)	Foil (word)	Exemplar (verb)		
HF Exemplar/HF Foil	273.4	728.1[1]	261.6	2.9	4.0
Semantic Match	278.9	–	249.6	2.7	4.3
HF Exemplar/LF Foil	187.2	4.9	166.4	3.0	4.7
Semantic Match	181.3	–	168.4	2.8	4.6
LF Exemplar/HF Foil	6.4	266.2	5.3	3.5	4.6
Semantic Match	6.1	--	5.0	3.2	4.8
LF Exemplar/LF Foil	4.1	5.1	3.3	3.2	4.4
Semantic Match	4.1	–	3.2	2.9	4.8

Note: Word frequency was calculated using the Kucera and Francis (1967) norms and verb frequency was calculated using the norms of Francis and Kucera (1982). See text for a description of the typicality measure. HF = high frequency; LF = low frequency. [1]This figure is inflated by two items; the mean frequency of the other 18 items is 331.3.

They were told that they were to decide whether each word presented was a verb or not. They were then shown the practice words and were given feedback on each trial to ensure that they understood the task. The experimental lists were presented following the practice, half of the subjects saw list 1 first and half saw list 2 first. After a short break they were shown the remaining list. Words were presented on the screen one at a time and remained until the subject responded. The intertnal interval was 1.5 s. Other aspects of the procedure are the same as in the previous experiments.

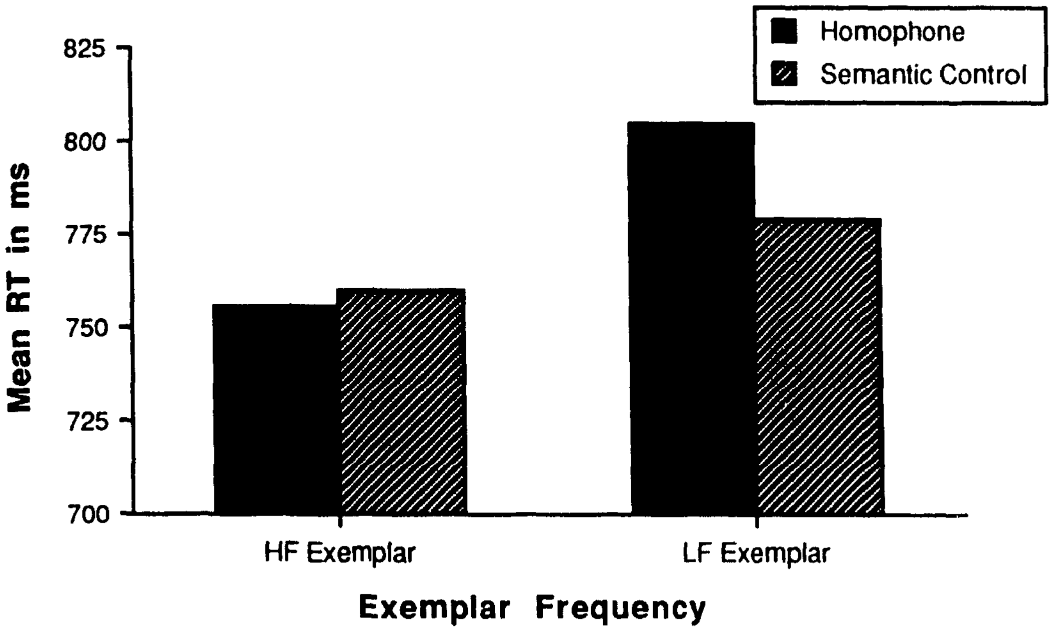
Results

Of the 400 trials in the experiment, only data from the 80 homophone exemplar and 80 semantic match YES trials were analyzed. Prior to analysis, the percent error data were arcsine transformed. A subject's response latency was only included in the analyses if the subject responded correctly to both the homophone and its semantic control. This criterion was met by 81.5% of responses. Another 8.0% of responses were correct but were discarded because the subjects made an error on another member of the stimulus pair. A total of 3.0% of response times were greater than 1500 ms and were replaced with times of 1500 ms. Almost a third of these were the response times of one subject. There were three factors in the analyses, frequency of the exemplar (high vs. low), frequency of the foil (high vs. low), and homophony (homophone vs. semantic control) The latency and error results are shown in Figures 9 and 10 respectively.

In the decision latency data, the main effect of homophony (11 ms) was not significant, $F(1,24) = 3.29$, $p > .05$ by subjects and $F(1,152) = 1.88$, $p > .05$. The main effect of exemplar frequency (34 ms) was significant, $F(1,24) = 9.78$, $p < .01$ by subjects and $F(1,152) = 6.83$, $p < .01$ by items. The main effect of foil frequency (4 ms) was not significant, both F's < 1.

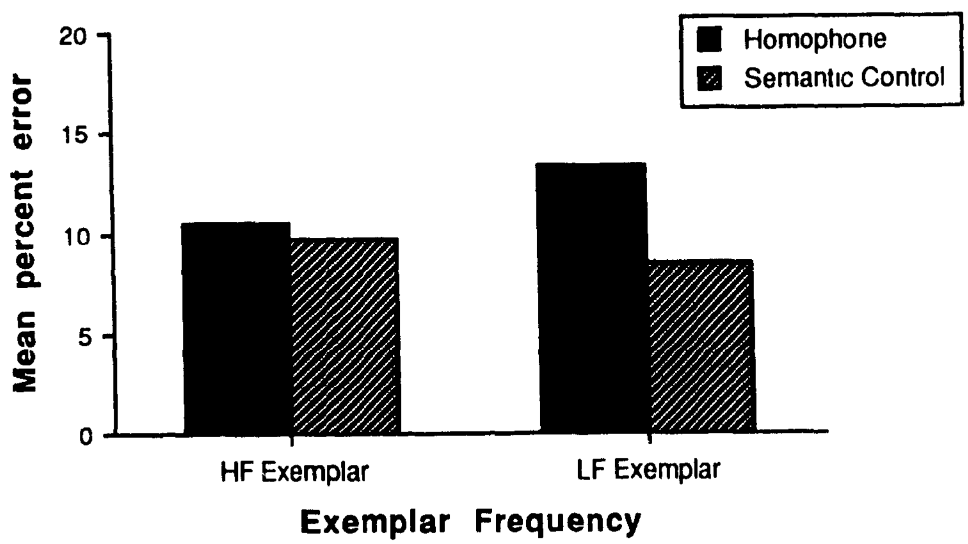
The interaction between homophony and exemplar frequency was significant by subjects, $F(1,24) = 4.46$, $p < .05$, and items, $F(1,152) = 3.77$, $p= .05$. The effect of

Figure 9: The mean decision latencies for homophone exemplars and semantic controls in Experiment 6.



Homophone
Semantic Control
825
800
775
750
725
700
Mean RT in ms
HF Exemplar
LF Exemplar
Exemplar Frequency

Figure 10: The mean percent errors for homophone exemplars and semantic controls in Experiment 6.


Homophone
Semantic Control
Mean percent error
20
15
10
5
0
HF Exemplar
LF Exemplar
Exemplar Frequency

homophony was -4 ms for high frequency words and 26 ms for low frequency words. The interaction between homophony and foil frequency was not significant by subjects, $F(1,24) = 1.97$, $p > .05$, or by items, $F(1,152) = 1.95$, $p > .05$. The triple interaction between homophony, exemplar frequency, and foil frequency was not significant, both F's < 1.

In the error data, the main effect of homophony (2.8%) approached significance by subjects, $F(1,24) = 3.80$, $p = .06$, but was not significant by items, $F(1,152) = 1.58$, $p > .05$. The main effect of exemplar frequency was not significant, $F(1,24) = 1.22$, $p > .05$ by subjects and $F < 1$ by items, nor was the main effect of foil frequency, both F's < 1.

The interaction between homophony and exemplar frequency approached significance by subjects, $F(1,24) = 3.52$, $p < .08$, but not by items, $F < 1$. The effect of homophony was 0.7% for high frequency words and 4.8% for low frequency words. The interaction between homophony and foil frequency was significant by subjects, $F(1,24) = 4.54$, $p < .05$, but not by items, $F < 1$. The effect of homophony was 1.1% when homophones had high frequency foils and 4.4% when they had low frequency foils. The triple interaction between homophony, exemplar frequency, and foil frequency was not significant by subjects, $F(1,24) = 3.08$, $p > .05$, or by items, $F < 1$.

A further analysis examined the correlation between the size of the homophone effect and the orthographic similarity of the homophone exemplar and its foil. Orthographic similarity was calculated using Weber's (1970) measure. If subjects use a spelling check, and a word is more likely to falsely pass if it is orthographically similar to the exemplar, then there should be a positive correlation between the size of the homophone effect in the error data and the similarity of the members of the homophone pair. When all items were included in the analysis of the error data, this relationship was not observed, $r = .01$. According to Van Orden (1987), the correlation would be expected particularly when homophone exemplars have low frequency foils because false candidates are more likely to pass the spelling check if their spellings are unfamiliar. However, when only items with low frequency foils were included, this relationship was extremely weak, $r = .11$. And, if a low frequency false candidate is only submitted to a spelling check if a higher frequency word has not been found to be an exemplar first, then there should be a correlation between the size of the homophony effect and spelling similarity when both members of a homophone pair are low in frequency. However, this relationship was again not observed when only the data from the low frequency exemplar/low frequency foil group were included, $r = .08$. In the latency data these correlations were slightly higher, $r = .19, .22$, and .15 in the three analyses respectively, but still weak, t's < 1.75.

Discussion

No evidence was found in Experiment 6 for the use of phonological information in the access of the meanings of high frequency words. On the other hand, subjects did show evidence that the access of the meanings of low frequency words is phonologically mediated since a significant effect of homophony for low frequency words was found in the decision latency data. The results of Experiment 6 do not provide support for the view that the meanings of all words are accessed phonologically (Van Orden, 1987, Van Orden et al., 1988), nor do they provide support for the view that a spelling check is performed on candidates (Becker, 1976, 1980; Davelaar et al., 1978; Paap et al., 1982; Rubenstein et al., 1971; Van Orden, 1987; Van Orden et al., 1988). If such a check had been performed on phonologically activated high frequency words, then there should have been an effect of homophony in the latency data for high frequency exemplars with high frequency foils, since on at least some occasions the exemplar would be submitted to the spelling check only after the foil had been checked. In fact, decision latencies were 16 ms faster for homophones than semantic controls in this group. Further, a strong correlation between spelling similarity and the size of the homophone effect should have been present in the data, and this was not observed.

The data instead support the dual route view that both orthographic and phonological information activate semantic representations, and that the meanings of high frequency words are highly activated via orthographic information before much or any activation from phonological representations contribute to semantic activation. For low frequency words, activation from the orthographic representation is slower and/or weaker, and so activation from phonology does have an opportunity to contribute to the activation of the semantic representation. While phonological activation may lead to inappropriate meanings being activated in the case of homophones, the combination of orthographic and phonological evidence for the word shown is usually enough to prevent an error, although it may take a little longer for the correct semantic representation to dominate when other possibilities are also activated. In Experiments 1-4 there was an effect of the frequency of the other member of the homophone pair, which indicated that subjects were performing a spelling check on low frequency words before responding. There was no effect of foil frequency in the present experiment. This suggests that the spelling check is a cautious strategy that subjects adopt when the experiment contains targets that sound like they require a YES response (object-BREAK) but actually require a NO response.

GENERAL DISCUSSION

The present studies have addressed the issue of whether phonological information plays a role in skilled word recognition. This issue is an important one, because it is central to the phonics - whole word method controversy in reading instruction. Further, resolution of this question may provide insight into the problems faced by deaf readers.

An indication of the importance of the issue is the large number of studies, reviewed earlier, that have been designed to determine whether skilled readers make use of phonological information in recognizing words The reason that so many studies have been conducted, is that it is difficult to find a way to determine whether phonological recoding is taking place, and to find a task that measures the time to access the meaning of a word that is not susceptible to task-specific strategies. Shortcomings of some of the methods that have been used in the past have become clear only after extensive research. A review of that research revealed that the use of homophones was the best way to examine phonological recoding in both high and low frequency words since homophones will necessarily activate two meanings if they are recoded phonologically. Another widely used method, comparing words that differ in spelling-sound consistency, has recently been cast into doubt as a valid method for examining phonological recoding for high frequency words by Seidenberg and McClelland's (1989) proposal that the translation process for a word becomes more efficient with practice. The absence of consistency effects for high frequency words can no longer be assumed to indicate that a visual code is being used The review of the literature also indicated that the semantic decision task was the best task to use to examine the access of the meanings of single words. Recent research has questioned whether the frequently used lexical decision task requires subjects to access the meanings of words. It appears to be susceptible to subject strategies with the result that different types of stimulus lists give different answers as to whether a phonological code is used to access meaning. Finally, it has become clear that performance on high frequency and low frequency words needs to be examined separately. The present studies have overcome some of the difficulties associated with the previous studies by examining performance on high and low frequency homophones in a semantic decision task. They also examined the use of subject strategies on that task.

Evidence for Phonological Mediation

The results of experiments 1-6 strongly suggest that access to the meanings of printed words is phonologically mediated only for low frequency words. There was no evidence for phonological mediation for high frequency words. Subjects were not more likely to make false positive errors on high frequency homophone foils than on spelling controls when priming from the category name was avoided. Nor was there any evidence that priming the other member of the homophone pair increased errors or decision latencies on high frequency homophone exemplars. And finally, subjects did not produce longer decision latencies or more errors on high frequency homophone exemplars than on nonhomophone semantic controls. The absence of effects of phonology for high frequency words suggests that their meanings are accessed directly on the basis of visual information.

In contrast, phonological effects were observed for low frequency words in each experiment. There were more false positive errors on low frequency homophone foils than on spelling controls, priming the other member of the homophone pair caused an increase in errors on low frequency homophone exemplars, and decision latencies for low frequency homophone exemplars were longer than for nonhomophonic semantic controls. Evidence that these phonological effects arise from processing prior to the access of meaning comes from the observation in Experiment 2 of similar false positive error rates for low frequency homophone and pseudohomophone foils. Further, there was evidence that the use of a phonological code is not an optional strategy. In Experiment 3, more errors were made on low frequency homophone foils than on spelling controls even when subjects were discouraged from using a phonological code by including a large proportion of homophones.

Using the lexical decision task, Andrews (1982), Waters et al. (1984), and Waters and Seidenberg (1985), also found evidence for phonological recoding only for low frequency words, although no phonological effects were observed by Seidenberg et al. (1984). The conclusion that meaning is phonologically mediated for low frequency but not high frequency words can be made more strongly on the basis of Experiments 1-6. One reason is that the task used here, the semantic decision task, requires access to meaning whereas the lexical decision task does not. The second reason is that evidence for phonological recoding only for low frequency words was observed in Experiment 6 which did not include pseudowords or foil trials. Thus, the results are more readily generalizable to normal reading than are the results from lexical decision experiments that are susceptible to special strategies because nonwords must be included. And finally, the present experiments used a homophony manipulation instead of a spelling-sound consistency manipulation.

Evidence for a Spelling Check

The results of the present experiments also suggest that a spelling check is not usually performed on low frequency words, but is a strategy that is used when the experiment contains foil trials, as in Experiments 1-4. In Experiments 1-3, subjects made fewer errors on low frequency homophone foils when they had high frequency exemplars than when they had low frequency exemplars. This is consistent with the view that a spelling check was performed, since subjects were better able to avoid errors when the spelling of the exemplar was familiar. In Experiment 4, subjects made more errors on low frequency homophone foils with similarly spelled exemplars than on foils with dissimilarly spelled exemplars. This is consistent with the hypothesis that exemplars are more li' ely to falsely slip by the spelling check if they are spelled like the target foil. These studies support the view that a spelling check is performed in experiments containing foil trials. In Experiment 5, however, no foil trials were included, and in that experiment, subjects made more errors on homophone exemplars when the other member of the homophone pair was primed and was high frequency than when the other member of the pair was primed and was low frequency. If a spelling check was being performed, subjects should have been better able to avoid errors when the spelling of the other member of the homophone pair was familiar. Further, in Experiment 6, the size of the homophony effect in the error data was not correlated with the similarity between the exemplar and the other member of the homophone pair. If a spelling check was being performed, subjects should have been more likely to make errors on homophone exemplars when the two were similarly spelled.

The failure to find evidence for the use of a spelling check when no foil trials were included suggests that activation from orthography is normally strong enough to allow readers to avoid making errors on low frequency homophones, although they will be slowed by the competition from the other member of the homophone pair. When foil trials are included, subjects are more cautious and perform a spelling check to avoid making false positive errors. This conclusion is not inconsistent with the spelling check literature. Van Orden (1987) noted that the primary evidence for verification models comes from lexical decision tasks that included pseudohomophone foils (Becker, 1976, 1980; Becker & Killion, 1977; but cf. Schvaneveldt & McDonald, 1981).

Implications for Theories of Word Recognition

The results of Experiments 1-6 do not support Van Orden's (1987) or Rubenstein et al.'s (1971) hypothesis that bottom up activation of candidates occurs exclusively via their phonological representations and that a spelling check is performed on the most highly activated candidates. They instead support the dual route view that both orthographic and phonological pathways to meaning exist. In the case of high frequency words, meaning is activated via the orthographic representation before processing has been completed along the phonological route. For low frequency words, orthographic activation of meaning is slower and/or weaker, and phonological information can contribute to semantic activation before any one has been sufficiently activated. This view can incorporate a spelling check to explain the finding of fewer errors on low frequency foils with high frequency exemplars in Experiments 1-4, but it is not normally part of the word recognition process. When two or more semantic representations are activated to similar degrees and the subjects are aware that phonological impostors are included in the experiment, they may check to see whether there is orthographic activation for the most highly activated representation before making a response.

Although these results are consistent with the broad dual route approach, they are not consistent with two proposals made by some dual route theorists. Coltheart (1978) argued that in skilled reading, processing along the visual route was nearly always completed prior to processing along the phonological route. The results of the present experiments suggest that this is true only for high frequency words. A common assumption of dual route models (e.g. Coltheart, 1978; Davelaar et al., 1978; Hawkins et al., 1976; McQuade, 1981) is that use of the phonological route is a strategy under control of the subject However, results of the present experiments suggest instead that it is the spelling check that is strategic. This view is consistent with the finding that phonological information appears to be activated quickly and automatically (Bakan & Alperson, 1967; Dalrymple-Alford, 1972; Humphreys et al., 1982; Perfetti et al., 1988, Tanenhaus et al., 1980).

Seidenberg and McClelland's (1989) connectionist dual route model can provide an account of the present results. As in other dual route models, there are two pathways to meaning, one directly from orthography, and the other from orthography to phonology to meaning. The model differs in that does not include representations of individual words; rather the spelling, pronunciation, and meaning of a word are represented by patterns of activation across simple processing units encoding each of these types of information. In their simulation, the connections between orthographic and phonological units are weighted, and it is these weights that encode the correspondences between spelling and sound. Theoretically, there are also weighted connections between orthographic and semantic units, and between phonological and semantic units that encode the correspondences between these representations. In the simulation, the weights start out as small random values, and then change via a back propagation learning algorithm during a training phase in which the model is exposed to a large number of words and their pronunciations. Gradually the weights come to encode facts about the consistency of spelling-sound correspondences in the training corpus, so that the correct phonological representation of a word is given when an orthographic pattern is presented. A similar process is assumed to apply for the weights on the connections between orthography and semantics and between phonology and semantics.

As in other dual route models, upon presentation of a stimulus, activation is assumed to spread from the orthographic units directly to the semantic units, and also from the orthographic units to phonological units and then to semantic units. However, whereas the assumptions of other models about which route would finish first are rather ad hoc, in the Seidenberg and McClelland model this is determined by the weights on the connections. These in turn depends on the amount of experience the model (or the subject) has with a given word. Further, in most of the other dual route models, either one or the other route determines the output, and it is not clearly specified what happens in the case of a conflict. In the Seidenberg and McClelland (1989) model, activation from orthographic and from phonological representations can both contribute to the activation of semantic representations, and the most highly activated representation determines the output. Whether or not phonological information influences (i.e. contributes to or competes with) the activation of the output representation depends on whether this semantic representation is sufficiently activated directly via orthography by the time activation from the phonological route arrives. In this view, activation from phonology can influence the semantic representation even before a clear pattern emerges on the phonological units, that is, a stimulus does not have to be fully recoded before it can begin to have an effect on the semantic representation. This makes it more likely that the two step phonological route will influence the representation activated by the one step orthographic route Finally, in this view, subjects can choose how long to wait before making a response, and can examine the orthographic representation corresponding to a semantic representation to see whether it is the same as the input orthographic representation, but they cannot prevent processing along the phonological route. This is consistent with the findings of Experiment 3.

The results of the present studies indicate that the orthographic to semantic route is a necessary part of the Seidenberg and McClelland (1989) theory, and must be included in the simulation in order to have a full model of the reading process. Further, the finding of phonological effects for low frequency but not high frequency words offers some information to connectionist modelers about the relative timing of activation along the two routes in skilled reading.

Previous Homophone Studies

This view outlined above can explain the results of the experiments by Davelaar et al. (1978, Expts. 3 & 4). They found an effect of homophony in a lexical decision experiment only for homophones that were the lower frequency member of the homophone pair Davelaar et al. (1978) then demonstrated that this homophone effect disappeared when half of the nonwords in the experiment were pseudohomophones. They suggested that the representations of both members of a homophone pair are activated via their phonological representations, and that a spelling check is conducted on the lower frequency member only after the spelling of the higher frequency member has been checked. Thus, there would be a delay for a lower frequency member of a homophone pair but not for a higher frequency member. Rubenstein et al. (1971) also proposed this explanation of the results of their homophone study. Davelaar et al. (1978) suggested that the homophony effect disappeared when pseudohomophones were included because subjects abandoned phonological recoding and used a visual code instead. As was pointed out in the introduction, it is unclear why the spelling check process that was postulated to explain the frequency difference would not serve to prevent errors in this condition. For both pseudohomophones and the lower frequency member of a homophone pair, the first spelling check should fail. For pseudohomophones, the subject could respond NO if there were no other activated words (as in the case of single pseudohomophones such as BRANE), and do a second spelling check if another alternative was available (as in the case of double pseudohomophones, such as GRONE, and word homophones). Thus, the spelling check procedure should still produce accurate performance when pseudohomophones are included.

The results of the studies above suggest another interpretation of Davelaar et al.'s (1978) experiment. They suggest that there was no effect for the higher frequency members of a homophone pair because the meanings of these words are highly activated prior to the influence of phonological information. The effect for low frequency members may have disappeared when pseudohomophones were included because subjects added a spelling check, and thus were able to catch these phonological impostors much as they did in the semantic decision experiments of Van Orden and Experiments 1-4 here. That is, instead of abandoning a phonological code in favor of a visual code, the change in strategy may have been to add a spelling check. Subjects could also have used a superficial orthographic familiarity strategy because meanings do not need to be accessed in order to respond YES in lexical decision. This would also account for the weak or nonexistent homophone effects in the Rubenstein et al. (1971), Coltheart et al. (1977), Dennis et al. (1985), and Barry (1981) studies because all included pseudohomophones among the nonwords. Thus, although Davelaar et al.'s (1978) explanation has difficulty accounting for both the influence of frequency and pseudohomophones on the homophone effect, the explanation just presented can easily account for both.

The results of the present experiments also suggest another interpretation of Davelaar et al.'s (1978, Expt. 2) priming experiment. They found facilitation for lower frequency homophones when they were preceded two trials earlier by the higher frequency member of the homophone pair, but found inhibition for higher frequency homophones when they were preceded by the lower frequency member of the homophone pair. According to their view, when the low frequency member is presented first, entries for both members are activated, and the entry for the high frequency member is chosen for the spelling check first. This check will fail and they proposed that inhibitory effects would reduce the availability of the entry for the high frequency member when it is presented two trials later When the higher frequency member is presented first, only its entry undergoes a spelling check so there are no inhibitory effects. The phonological representation of the higher frequency member will have activated the entry for the lower frequency member so that when it is presented two trials later, facilitation is observed.

The results of the present experiments suggest another interpretation of the inhibitory effect for high frequency homophones when they are preceded by the lower frequency member of the homophone pair. The lower frequency member will activate the phonological representation of the pair and this may still be active two trials later when the high frequency member is presented. The meaning of a high frequency member is normally sufficiently activated before information from the phonological route arrives, so that there are no effects of homophony. However, the computation along the phonological route may be much faster when it has been primed, so that the output from this route does contribute to semantic activation. This would slow responses on the high frequency members since subjects must await the outcome of the competition between the two representations of the homophone. When the higher frequency member is presented first, its meaning is activated prior to the arrival of the input from the phonological route, so phonological information does not affect decision latency. However, the phonological representation is activated, and it may facilitate responding on a subsequent low frequency member that depends on this information for recognition. That is, priming of the phonological representation of a homophone pair would slow processing on subsequent high frequency members that are normally processed before this information arrives, but speed processing on lower frequency members that need this information for recognition.

Word Frequency Effects

Although the present studies were designed to examine the use of phonological information in word recognition, the results from Experiment 6 can provide evidence in another debate in word recognition, and that is the debate about the locus of the effect of word frequency. Balota and Chumbly (1984) failed to observe an effect of word frequency on YES decision latencies for exemplars. They concluded that the large frequency effect they and others have found in lexical decision reflects a task-specific decision process that has little to do with the access of meaning. However, Monsell (1985) and Monsell et al. (1989) did find an effect of word frequency on YES semantic decision latencies, and they point to several problems with Balota and Chumbly's (1984) experiment that might explain the discrepancy. In Experiment 6, a significant effect of frequency was found on YES latencies, consistent with the results of Monsell (1985) and Monsell et al. (1989). This result offers further support to the widely held view that the process of accessing meanings is sensitive to word frequency. It is a central assumption of the Seidenberg and McClelland (1989) model.

Evaluation of the Semantic Decision Task

The semantic decision task seems to be a better measure of the access of the meanings of individual words than lexical decision or sentence verification. Several observations in the present studies have implications for the use of the semantic categorization task in the study of word recognition. The first is that the category names used must be broad in order to prevent priming of the target. Monsell et al. (1989) have also made this recommendation. Experiment 2 demonstrated that the use of specific category names in Experiment 1 and in Van Orden's (1987; Van Orden et al., 1988) experiments resulted in an overestimation of the size of the difference between homophone foils and spelling controls.

The second observation about the semantic decision task is that performance on words does not appear to be very sensitive to the proportion of homophones included in the experiment but that performance on pseudowords is somewhat sensitive to this manipulation. In Experiment 3, the size of the difference between homophone foils and spelling controls did not change from Experiment 2 for low frequency words with low frequency exemplars, but was smaller for pseudohomophone foils with low frequency exemplars. In addition, pseudowords were responded to faster than words in Experiment 3, whereas they had the longest response times in Experiment 2. It is possible that when the proportion of homophones was high, and thus decisions were difficult, subjects attempted to use an orthographic familiarity strategy. Such a strategy would have been useful for performance on pseudowords because the correct response is always NO, but would not have been useful for words, since the correct response depended on their meaning. This suggests that studies that involve making comparisons between pseudoword performance and word performance should keep the proportion of homophones low to reduce the likelihood that special strategies will be used on pseudowords.

A third observation is that subjects adopt a spelling check strategy when foil trials are included in the experiment. This suggests that foil trials should be avoided if the object is to study normal word recognition processes. However, the use of foils was essential in the present experiments because the comparison between erroneous YES responses on pseudohomophone foils and homophone foils determined whether the phonological effects arose prior to the access of meaning or following access. The use of exemplar targets avoids the use of foils but does not allow the observation of YES responses on pseudowords. A further complication that arises when YES responses on exemplars are examined is that, in most designs, a set of semantic control words needs to be used. These are not always easy to find, and care must be taken to ensure that they are equally good members of their category. A combination of strategies thus seems appropriate when phonological mediation of the access of meaning is studied.

Further Implications

The results of the present experiments have several other implications. They suggest that Adams (1990) is correct in saying that spelling-sound translation lessons are a valuable component of beginning reading instruction, since skilled readers continue to make use of phonological information to aid the recognition of low frequency words. The importance of phonological information for skilled reading in hearing persons suggests that a possible reason why the deaf often fail to become skilled readers is that they are unable to benefit from the clues to meaning that phonological information provides. And finally, the finding of a role for phonological information in skilled reading suggests that a reason that languages may have evolved to represent the phonemes of the language is that the phonological information offers a valuable source of information for the recognition of words.

REFERENCES

Adams, M. J. (1979). Models of word recognition. *Cognitive Psychology*, **11**, 133-176.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Allport, D. A. (1979). Word recognition in reading: A tutorial review. In P. A. Kolers, H. Bouma, & M. Wrolstad (Eds.), *Processing of visible language* (Vol. 1, pp. 227-257). New York: Plenum Press.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition*, **10**, 565-575.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Baddeley, A. D. & Lewis, V. J. (1981). Inner active processes in reading: The inner voice, the inner ear, and the inner eye. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 107-129). Hillsdale, NJ: Erlbaum.

Bakan, P., & Alperson, B. (1967). Pronounceability, attensity, and interference in the color-word test. *American Journal of Psychology*, **80**, 416-420.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word-frequency in the neglected decision stage *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, **24**, 89-106.

Banks, W. P., Oka, E., & Shugarman, S. (1981). Recoding of printed words to internal speech: Does recoding come before lexical access? In O. J. L. Tzeng, & H. Singer (Eds.), *Perception of print* (pp. 137-170). Hillsdale, NJ: Erlbaum.

Baron, J. (1973). Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology,* **25**, 241-246.

Baron, J. (1985). Back to basics. *The Behavioral and Brain Sciences,* **8**, 706.

Barry, C. (1981). Hemispheric asymmetry in lexical access and phonological encoding. *Neuropsychologia,* **19**, 473-478.

Bauer, D. W., & Stanovich, K. E. (1980). Lexical access and the spelling-to-sound regularity effect. *Memory and Cognition,* **8**, 424-432.

Beauvois, M. F., & Dérouesné, J. (1979). Phonological alexia: Three dissociations. *Journal of Neurology, Neurosurgery, and Psychiatry,* **42**, 1115-1124.

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* **2**, 556-566.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory and Cognition,* **8**, 493-512.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* **3**, 389-401.

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *The Quarterly Journal of Experimental Psychology,* **39A**, 467-478.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology,* **36**, 701-711.

Besner, D., & Davelaar, E. (1983). Suedohomofoan effects in visual word recognition: Evidence for phonological processing. *Canadian Journal of Psychology,* **37**, 300-305.

Besner, D., Davelaar, E., Alcott, D., & Parry, P. (1984). Wholistic reading of alphabetic print: Evidence from the FDM and the FBI. In L. Henderson (Ed.), *Perspectives from cognitive psychology, neuropsychology, and linguistics* (pp. 121-135). London: Erlbaum.

Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology*, **33A**, 415-437.

Besner, D., Dennis, I., & Davelaar, E. (1985). Reading without phonology? *The Quarterly Journal of Experimental Psychology*, **37A**, 477-491.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the connection between connectionism and data: Are a few words necessary? *Psychological Review*, **97**, 432-446.

Brown, G. D. A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language*, **26**, 1-23.

Bub, D., Cancelliere, A., & Kertesz, A. (1985). Whole-word and analytic translation of spelling to sound in a non-semantic reader. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 15-34). London: Erlbaum.

Butler, B. E., & Hains, S. (1979). Individual differences in word recognition latency. *Memory and Cognition*, **7**, 68-76.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, **25**, 85-95.

Chumbly, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory and Cognition*, **12**, 590-606.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, **12**, 335-359.

Colombo, L. (1986). Activation and inhibition with orhographically similar words. *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 226-234.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). New York: Academic Press.

Coltheart, M. (1980). Reading, phonological recoding, and deep dyslexia. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 197-226). London: Routledge & Kegan Paul.

Coltheart, M., Besner, D., Jonasson, J. T., & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology*, **31**, 489-507.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Patterson, K., & Marshall, J. C. (1980). *Deep dyslexia*. London: Routledge & Kegan Paul.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 387-397.

Conrad, R. (1977). The reading ability of deaf school-leavers. *British Journal of Educational Psychology*, **47**, 138-148.

Cosky, M. J. (1976). The role of letter recognition in word recognition. *Memory and Cognition*, **4**, 207-214.

Crowder, R. G. (1982). *The psychology of reading* New York: Oxford University Press.

Dalrymple-Alford, E. C. (1972). Sound similarity and color-word interference in the Stroop task. *Psychonomic Science*, **28**, 209-210.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. T. (1978). Phonological recoding and lexical access. *Memory and Cognition*, **6**, 391-402.

Davidson, B. J. (1986). Activation of semantic and phonological codes during reading. *Journal of Experimental Psychology Learning, Memory, and Cognition*, **12**, 201-207.

Dennis, I., Besner, D., & Davelaar, E (1985). Phonology in visual word recognition: Their is more two this than meats the I In D. Besner, T G. Waller, & E. M MacKinnon (Eds.), *Reading research Advances in theory and practice* (Vol. 5, pp. 167-197). Toronto. Academic Press.

Dérouesné, J., & Beauvois, M F. (1979) Phonological processing in reading: Data from alexia. *Journal of Neurology, Neurosurgery, and Psychiatry*, **42**, 1125-1132.

Dérouesné, J., & Beauvois, M. F (1985) The "phonemic" stage in the non-lexical reading process: Evidence from a case of phonological alexia In K E. Patterson, J C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia Neuropsychological and cognitive studies of phonological reading* (pp 399-457) London: Erlbaum

DiFrancesca, S. (1972). *Academic achievement results of a national testing program for hearing impaired students, United States, Spring 1971* Washington, DC Gallaudet College.

Eriksen, C. W., Pollack, M. D., & Montague, W. E. (1970). Implicit speech: Mechanism in perceptual encoding? *Journal of Experimental Psychology*, **84**, 502-507.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, **33A**, 325-350

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory and Cognition*, **5**, 335-339.

Forster, K. I. (1976). Accessing the mental lexicon. In E. C. T. Walker & R. J. Wales (Eds.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North-Holland.

Forster, K. I. (1988). *On knowing how many entries.* Paper presented at the Conference on Semantic Ambiguity, Adelphi University.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, **12**, 627-635.

Francis, W. N., & Kucera, H (1982). *Frequency analysis of English usage: Lexicon and grammar* Boston. Houghton-Mifflin.

Frederiksen, J R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 361-379.

Funnell, E. (1983) Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology*, **74**, 159-180.

Gernsbacher, M A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113**, 256-281.

Glushko, R J. (1979) The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 674-691.

Goodman, K (1972). Orthography in a theory of reading instruction. *Elementary English*, **49**, 1254-1261.

Goodman, K. (1973). The 13th easy way to make reading difficult: A reaction to Gleitman and Rozin. *Reading Research Quarterly*, **8**, 484-493.

Gordon, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior*, **22**, 24-44.

Gough, P. B. (1972). One second of reading. In J. P. Kavanagh & I. G. Mattingly (Eds.), *Language by eye and by ear* (pp. 331-358). Cambridge, MA: MIT Press

Green, D. W., & Shallice, T. (1976). Direct visual access in reading for meaning. *Memory and Cognition*, **4**, 753-758.

Hanson, V. L., & Fowler, C. A (1987). Phonological coding in word reading: Evidence from hearing and deaf readers. *Memory and Cognition*, **15**, 199-207.

Hawkins, H. L., Reicher, G. M., Rogers, M., & Peterson, L (1976) Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 380-385.

Henderson, L. (1981) Critical notice: Information processing approaches to acquired dyslexia. *Quarterly Journal of Experimental Psychology*, **33A**, 507-522.

Henderson, L. (1982). *Orthography and word recognition in reading* London: Academic Press

Henderson, L , Coltheart, M., & Woodhouse, D. (1973) Failure to find a syllabic effect in number-naming. *Memory and Cognition*, **1**, 304-306

Hildebrandt, N., Sokol, S., Dymkowski, T., & Ruzecki, V. (1990). *Implicit grapheme-to-phoneme conversion in deep dyslexia* Paper presented at TENNET, Montreal

Hillinger, M. L (1980). Priming effects with phonemically similar words The encoding-bias hypothesis reconsidered. *Memory and Cognition*, **8**, 115-123

Hillinger, M. L., & James, C. (1977). *Sow or sow? Phonemic coding and ambiguity decision.* Paper presented at the meeting of the Psychonomic Society, Washington, DC.

Hudson, P. T. W. & Bergman, M. W (1985). Lexical knowledge in word recognition Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language*, **24**, 46-58.

Huey, E. B. (1908/1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press

Humphreys, G. W., Evett, L. J., & Taylor, D E. T. (1982). Automatic phonological priming in visual word recognition. *Memory and Cognition,* **10**, 576-590.

Hung, D. L., & Tzeng, O. J. L. (1981). Orthographic variations and visual information processing. *Psychological Bulletin*, **90**, 377-414.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance,* **1**, 130-136.

Jared, D (1985). *The processing of multisyllabic words· Effects of phonological regularity, syllabic structure, and frequency.* Unpublished masters thesis, McGill University.

Jared, D., McRae, K., & Seidenberg, M. S. (in press). The basis of consistency effects in word naming. *Journal of Memory and Language.*

Jared, D., & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance.* **16**, 92-105.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology,* **13**, 278-305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior,* **14**, 534-537.

Johnson, N. F. (1975). On the function of letters in word identification: Some data and a preliminary model. *Journal of Verbal Learning and Verbal Behavior,* **14**, 17-29.

Kay, J., & Marcel, A (1981). One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology,* **33A**, 397-413.

Klapp, S. T. (1971). Implicit speech inferred from response latencies in same-different decisions. *Journal of Experimental Psychology,* **91**, 262-267.

Klapp, S. T. (1974). Syllable-dependent pronunciation latencies in number naming: A replication. *Journal of Experimental Psychology*, **102**, 1138-1140

Klapp. S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading, reconsidered. *Journal of Experimental Psychology*, **100**, 368-374

Kleiman, G. M. (1975). Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior*, **14**, 323-339.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English* Providence, RI: Brown University Press.

Laberge, D., & Samuels, J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, **6**, 293-323.

Lupker, S. J., & Williams, B. A. (1989). Rhyme priming of pictures and words: A lexical activation account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 1033-1046.

Martin, R. (1982). The pseudohomophone effect: The role of visual similarity in nonword decisions. *Quarterly Journal of Experimental Psychology*, **34A**, 395-409

Martin, R. C., & Jensen, C. R. (1988). Phonological priming in the lexical decision task: A failure to replicate. *Memory and Cognition*, **16**, 505-521.

Mason, M. (1978). From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. *Memory and Cognition*, **6**, 568-581.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 693-706

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **86**, 375-407.

McConkie, G. W., & Zola, D. (1981). Language constraints and the functional stimulus in reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 155-175) Hillsdale, NJ: Erlbaum.

McCusker, L X., Hillinger, M. L., & Bias, R. G. (1981). Phonological recoding and reading. *Psychological Bulletin,* **89**, 217-245.

McCutchen, D., & Perfetti, C. A. (1982). The visual tongue-twister effect: Phonological activation in silent reading. *Journal of Verbal Learning and Verbal Behavior,* **21**, 672-687.

McQuade, D V. (1981) Variable reliance on phonological information in visual word recognition. *Language and Speech,* **24**, 99-109.

Meyer, D. E., & Gutschera, K. (1975). *Orthographic vs. phonemic processing of printed words*. Paper presented at the meeting of the Psychonomic Society, Denver.

Meyer, D E , & Ruddy, M. G. (1973). *Lexical-memory retrieval based on graphemic and phonemic representations of printed words.* Paper presented at the meeting of the Psychonomic Society, St. Louis.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory and Cognition,* **2**, 390-321.

Monsell, S. (1985) Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2, pp. 147-195). London: Erlbaum.

Monsell, S., Doyle, M C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General,* **118**, 43-71.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, **76**, 165-178.

Morton, J., & Patterson, K. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 91-118). London: Routledge & Kegan Paul.

Myers, J. L. (1979). *Fundamentals of experimental design* Boston: Allyn & Bacon

Neely, J. H. (in press). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds), *Basic processes in reading: Visual word recognition.* Hillsdale, N. J.. Erlbaum

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 1003-1019.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects *Quarterly Journal of Experimental Psychology,* **17**, 273-281

Paap, K. R., McDonald, J. E., Schvaneveldt, R. W , & Noel, R. W. (1987) Frequency and pronounceability in visually presented naming and lexical decision tasks In M Coltheart (Ed.) *Attention and performance XII The psychology of reading* (pp 221-243). London: Erlbaum.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W (1982) An activation verification model for word and letter recognition: The word superiority effect. *Psychological Review,* **89**, 573-594

Parkin, A. J. (1982). Phonological recoding in lexical decision Effects of spelling-to-sound regularity depend on how regularity is defined. *Memory and Cognition,* **10**, 43-53.

Parkin, A. J., & Ellingham, R. (1983). Phonological recoding in lexical decision: The influence of pseudohomophones. *Language and Speech,* **26**, 81-90.

Parkin, A. J., McMullen, M., & Graystone, D. (1986). Spelling-to-sound regularity affects pronunciation latency but not lexical decision. *Psychological Research,* **48**, 87-92.

Parkin, A. J., & Underwood, G. (1983). Orthographic versus phonological irregularity in lexical decision *Memory and Cognition,* **11**, 351-355.

Patterson, K E. (1981). Neuropsychological approaches to the study of reading. *British Journal of Psychology,* **72**, 151-174.

Patterson, K. E. (1982). The relation between reading and phonological coding: Further neuropsychological observations. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 77-111) London: Academic Press.

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 335-359). London: Erlbaum.

Patterson, K., Seidenberg, M. S., & McClelland, J. L. (1989). Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neuroscience* (pp. 131-181). Oxford: Oxford University Press.

Perfetti, C. A., Bell, L C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language,* **27**, 59-70.

Richardson, J. T. E. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology,* **67**, 315-325.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General,* **104**, 192-233.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, **9**, 487-494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, **10**, 645-657.

Rubenstein, H., Richter, M. L., & Kay, E. J. (1975). Pronounceability and the visual recognition of nonsense words. *Journal of Verbal Learning and Verbal Behavior*, **14**, 651-657.

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review*, **81**, 99-118.

Schvaneveldt, R. W., & McDonald, J. E. (1981). Semantic context and the encoding of words: Evidence for two modes of stimulus analysis *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 673-687.

Schwartz, M. F., Saffran, E. M., & Marin, O. S M. (1980). Fractionating the reading process in dementia: Evidence for word specific print-to-sound associations In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 259-269). London: Routledge & Kegan Paul.

Seidenberg, M. S. (1985). The time course of information activation and utilization in visual word recognition. In D. Besner, T. G. Waller, & E. M. MacKinnon (Eds), *Reading research: Advances in theory and practice* (Vol. 5, pp 199-252) Toronto: Academic Press.

Seidenberg, M. S. (1988). Cognitive neuropsychology and language: The state of the art. *Cognitive Neuropsychology*, **5**, 403-426.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523-568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior,* **23**, 383-404.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. Pre- and postlexical loci of contextual effects on word recognition. *Memory and Cognition,* **12**, 315-328.

Shallice, T., & Warrington, E. K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 119-145). London: Routledge & Kegan Paul.

Shallice, T., Warrington, E. K., & McCarthy, R. (1983). Reading without semantics. *Quarterly Journal of Experimental Psychology,* **35A**, 111-138.

Shoben, E J. (1982). Semantic and lexical decisions. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 287-314). New York: Academic Press.

Shulman, H. G., Hornak, R., & Sanders, E. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory and Cognition,* **6**, 115-123.

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read.* New York: Holt, Rinehart, & Winston.

Snodgrass, J. G., & Jarvella, R. J. (1972). Some linguistic determinants of word classification times *Psychonomic Science,* **27**, 220-222.

Spoehr, K T., & Smith, E. E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology,* **5**, 71-89.

Stanners, R. F., Forbach, G. B., & Headley, D. B. (1971). Decision and search processes in word-nonword classification. *Journal of Experimental Psychology,* **90**, 45-50.

Stanovich, K. E., & Bauer, D. W. (1978). Experiments on the spelling-to-sound regularity effect in word recognition. *Memory and Cognition*, **6**, 410-415

Tanenhaus, M. K., Flanigan, H., & Seidenberg, M. S. (1980) Orthographic and phonological activation in auditory and visual word recognition. *Memory and Cognition*, **8**, 513-520.

Treiman, R., Freyd, J. J., & Baron, J. (1983). Phonological recoding and the use of spelling-sound rules in reading of sentences. *Journal of Verbal Learning and Verbal Behavior*, **22**, 682-700.

Treiman, R., & Hirsh-Pasek, K. (1983). Silent reading: Insights from second generation deaf readers. *Cognitive Psychology*, **15**, 39-65.

Trybus, R. J., & Karchmer, M. A. (1977). School achievement scores of hearing impaired children: National data on achievement status and growth patterns. *American Annals of the Deaf Directory of Programs and Services*, **122**, 62-69.

Tweedy, J. R., Lapinski, R. H., & Schvaneveldt, R. W. (1977). Semantic-context effects on word recognition: Influence of varying the proportion of items presented in appropriate context. *Memory and Cognition*, **5**, 84-89.

Underwood, G., & Thwaites, S. (1982). Automatic phonological coding of unattended printed words. *Memory and Cognition*, **10**, 434-442.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition*, **15**, 181-198.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology Learning, Memory, and Cognition*, **14**, 371-386.

Venezky, R. L. (1970). *The structure of English orthography* The Hague: Mouton.

Walker, J. H. (1973). Pronounceability effects on word-nonword encoding in categorization and recognition tasks. *Journal of Experimental Psychology,* **99**, 318-322.

Waters, G. S., Komoda, M. K., & Arbuckle, T. Y. (1985). The effects of concurrent tasks on reading: Implications for phonological recoding. *Journal of Memory and Language,* **24**, 27-45.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory and Cognition,* **13**, 557-572.

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adult's use of spelling-sound information in three reading tasks. *Memory and Cognition,* **12**, 293-305.

Weber, R. M. (1970). A linguistic analysis of first-grade reading errors. *Reading Research Quarterly,* **5**, 427-451.

Whaley, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior,* **17**, 143-154.

Wijk, A. (1966). *Rules of pronunciation for the English language.* Oxford: Oxford University Press.

APPENDIX A

Stimuli used in Experiments 1-3

Note: The homophone controls are not shown; they are simply the foil words paired with one of the other categories.

Exemplar	Foil	Spelling Control	Expt. 1 Category	Expt. 2 Category
		High Frequency Foil/High Frequency Exemplar		
principal	principle	municipal	school employee	living thing
son	sun	sin	male relative	living thing
bear	bare	beat	hibernating animal	living thing
male	mail	mile	biological category	living thing
patients	patience	patent	people found in hospitals	living thing
residents	residence	resistance	citizens	living thing
cent	sent	count	type of money	object
clothes	close	claims	stored in closets	object
meat	meet	mean	dinner food	object
plane	plain	play	air vehicle	object
presents	presence	preserve	typically at a birthday party	object
road	rode	round	used by travellers	object
		High Frequency Foil/Low Frequency Exemplar		
beech	beach	beer	type of tree	living thing
knight	night	knife	distinguished man	living thing
nun	none	non	religious person	living thing
seller	cellar	secret	store personnel	living thing
prophet	profit	project	Biblical person	living thing
boarder	border	broader	paying resident	living thing
sail	sale	soul	part of a boat	object
ladder	latter	labor	painter's equipment	object
pail	pale	pain	cleaning equipment	object
axe	acts	age	woodsman's tool	object
throne	thrown	throat	monarch's object	object
brake	break	brave	car part	object

Exemplar	Foil	Spelling Control	Expt. 1 Category	Expt. 2 Category
		Low Frequency Foil/High Frequency Exemplar		
guys	guise	guts	group of men	living thing
horse	hoarse	hose	farm animal	living thing
prince	prints	prance	royalty	living thing
rose	rows	robe	garden flower	living thing
guest	guessed	gust	people in hotels	living thing
boy	buoy	bog	young person	living thing
board	bored	boast	construction material	object
bread	bred	brood	baked item	object
gate	gait	gale	type of barrier	object
ring	wring	rinse	type of jewellery	object
ball	bawl	bail	child's toy	object
sign	sine	sigh	information medium	object
		Low Frequency Foil/Low Frequency Exemplar		
flea	flee	flex	parasite	living thing
fowl	foul	foil	edible birds	living thing
toad	towed	trod	amphibian	living thing
whale	wail	whack	sea mammal	living thing
baron	barren	bargain	nobleman	living thing
pigeon	pidgin	piston	city dwelling bird	living thing
sword	soared	sworn	weapon	object
pole	poll	poke	tent part	object
medal	meddle	medley	type of award	object
urn	earn	urea	container	object
lens	lends	lent	optical device	object
bridle	bridal	brittle	horseback riding equipment	object

Exemplar	Foil	Spelling Control	Expt. 1 Category	Expt. 2 Category
		Pseudohomophone Foil/High Frequency Exemplar		
writer	riter	writed	person in book industry	living thing
trees	treeze	treem	type of vegetation	living thing
daughter	dotter	daupher	female relative	living thing
soldiers	soljers	soltier	military personnel	living thing
dog	dawg	dag	domestic animal	living thing
chief	cheef	chiel	type of leader	living thing
boat	bote	boam	water vehicle	object
table	tabel	tadle	type of furniture	object
rock	rawk	roch	natural earth formation	object
phone	phoan	phand	means of communication	object
shoes	shews	shoss	footwear	object
scale	scail	scalm	measuring device	object
		Pseudohomophone Foil/Low Frequency Exemplar		
eagle	eagel	eaple	predatory bird	living thing
fox	focks	fow	canine	living thing
poppy	paupy	poggy	symbolic flower	living thing
worm	wirm	wurn	legless animal	living thing
beetle	beatel	beelet	crawling insect	living thing
cod	caud	col	fish	living thing
purse	perse	porse	fashion accessory	object
drawer	droar	drawen	part of furniture	object
spear	speer	spean	hunting equipment	object
nail	nale	noil	type of fastener	object
skate	scate	skote	hockey equipment	object
stove	stoave	stoze	household appliance	object

APPENDIX B

Words used in Experiment 4

Exemplar	Foil	Spelling Control	Category
	High Similarity Between Exemplar and Foil		
heroine	heroin	heroic	living thing
baron	barren	bargain	living thing
fowl	foul	foil	living thing
serf	surf	scarf	living thing
pigeon	pidgin	piston	living thing
toad	towed	toyed	living thing
boulder	bolder	balder	object
bridle	bridal	brittle	object
cord	chord	cod	object
lens	lends	leans	object
altar	alter	ajar	object
pearl	purl	peril	object
	Low Similarity Between Exemplar and Foil		
lynx	links	lanky	living thing
fairy	ferry	fancy	living thing
hawk	hock	haul	living thing
idol	idle	idiom	living thing
doe	dough	dot	living thing
whale	wail	warn	living thing
rack	wrack	wreck	object
medal	meddle	medley	object
sword	soared	seared	object
mast	massed	marred	object
urn	earn	urea	object
sleigh	slay	slam	object

APPENDIX C

Words used in Experiment 5

Homophone Exemplars	Related Primes	Unrelated Primes	Category
	High Frequency Exemplar/High Frequency Foil		
principal	ethic	bustle	living thing
son	star	magic	living thing
bear	exposed	launch	living thing
male	letters	echo	living thing
patients	serenity	phase	living thing
residents	dwelling	option	living thing
cent	shipped	bash	object
clothes	fasten	regret	object
meat	join	gamble	object
plane	obvious	aggravate	object
presents	existence	nimble	object
road	sat	health	object
	High Frequency Exemplars/Low Frequency Foils		
guys	semblance	parallel	living thing
horse	raspy	reason	living thing
prince	photographs	hoot	living thing
rose	columns	malice	living thing
guest	estimated	pardon	living thing
boy	beacon	liberty	living thing
board	restless	past	object
bread	born	jaunt	object
gate	gallop	bite	object
ring	twist	baffle	object
ball	cry	hint	object
sign	angle	antic	object

Homophone Exemplars	Related Primes	Unrelated Primes	Category
		Low Frequency Exemplar/High Frequency Foil	
beech	shore	babble	living thing
knight	evening	pinch	living thing
nun	few	league	living thing
seller	basement	riddle	living thing
prophet	revenue	moist	living thing
boarder	outline	myth	living thing
sail	discount	jolt	object
ladder	following	passion	object
pail	colorless	bungle	object
axe	pretends	eclipse	object
throne	tossed	focus	object
brake	shatter	bold	object
		Low Frequency Exemplar/Low Frequency Foil	
flea	run	advantage	living thing
fowl	polluted	glory	living thing
toad	pulled	anxious	living thing
whale	weep	density	living thing
baron	arid	accident	living thing
pigeon	slang	tune	living thing
sword	glided	blame	object
pole	opinion	lucky	object
medal	intrude	brief	object
urn	acquire	blink	object
lens	donates	bellow	object
bridle	wedding	amount	object

APPENDIX D

Words used in Experiment 6

High Frequency Words

Homophone Exemplars	Semantic Controls	Homophone Exemplars	Semantic Controls
(High Frequency	Foils)	(Low Frequency	Foils)
buy	shop	been	has
hear	speak	walk	run
knew	said	caught	shook
made	used	seem	appear
won	lost	rain	snow
led	fled	taught	fought
write	draw	sign	agree
roll	push	stayed	watched
see	look	shoot	kill
threw	struck	flew	slid
wait	seek	raise	lift
wore	hung	flow	fill
passed	reached	loan	grant
do	go	tied	fixed
seen	felt	shear	split
meet	join	pause	rush
rode	drove	cast	pitch
close	open	beat	hit
sent	kept	break	burst
sell	trade	thrown	swung

Low Frequency Words

Homophone Exemplars (High Frequency	Semantic Controls Foils)	Homophone Exemplars (Low Frequency	Semantic Controls Foils)
blew	stung	alter	amend
counsel	advise	creak	squeak
haul	lug	maul	flog
hire	quit	peer	squint
owe	lend	savor	crave
pour	stir	leak	drip
sew	rip	tow	drag
steal	raid	purl	knit
weigh	ponder	pray	preach
whine	plead	tease	taunt
err	botch	slay	stab
stares	glares	tacked	nailed
mourning	grieving	retch	gasp
lessen	loosen	bury	dig
pare	chop	heal	mend
cite	quote	earn	reap
bawl	yell	meddle	intrude
bred	sired	soar	glide
bored	amused	wail	moan
daze	stun	flee	sprint

APPENDIX E

Mean reaction times in milliseconds and mean percent errors obtained in Experiment 1

Group		Category-Target Relation					
Foil Freq.	Exemplar Freq.	Foil		Spelling Control		Homophone Control	
		RT	Errors	RT	Errors	RT	Errors
High	High	680	22.9	639	15.3	652	6.2
High	Low	666	15.3	631	6.3	631	2.8
Low	High	646	7.0	620	3.5	631	2 8
Low	Low	721	26.4	669	8.3	622	3 5
Non	High	670	7 6	683	4 9	633	2 1
Non	Low	717	18.1	672	2.8	613	1 4

APPENDIX F

Mean reaction times in milliseconds and mean percent errors obtained in Experiment 2

Group		**Category-Target Relation**					
Foil Freq.	**Exemplar Freq.**	**Foil**		**Spelling Control**		**Homophone Control**	
		RT	Errors	RT	Errors	RT	Errors
High	High	787	14.6	754	11.1	746	8.3
High	Low	777	11.1	764	13.9	750	8.3
Low	High	788	9.0	806	7.6	779	5.0
Low	Low	788	21.5	787	6.9	766	4.9
Non	High	782	11.8	800	6.2	803	7.6
Non	Low	836	17.4	809	6.9	756	4.9

APPENDIX G

Mean reaction times in milliseconds and mean percent errors obtained in Experiment 3

Group		**Category-Target Relation**					
Foil Freq.	**Exemplar Freq.**	**Foil**		**Spelling Control**		**Homophone Control**	
		RT	Errors	RT	Errors	RT	Errors
High	High	954	6.9	960	6.9	871	7.5
High	Low	982	11.8	951	9.0	893	10.0
Low	High	887	7.6	969	10.4	941	5.8
Low	Low	916	18.7	920	4.2	949	2.8
Non	High	873	2.1	854	1.4	864	2.1
Non	Low	905	10.4	834	2.1	860	1.4

APPENDIX H

Mean reaction times in milliseconds and mean percent errors obtained in Experiment 4

Similarity	Word Type			
	Foil		Spelling Control	
	RT	Errors	RT	Errors
High	933	25.0	913	6.8
Low	925	11.8	890	9.0

APPENDIX I

Mean reaction times in milliseconds and mean percent errors obtained in Experiment 5

Group		**Prime-Target Relation**			
Exemplar Freq.	**Foil Freq.**	**Unrelated**		**Related**	
		RT	Errors	RT	Errors
		Preliminary Study (Foil Target)			
High	High	756	9.7	708	6.0
High	Low	712	5.1	668	4.6
Low	High	769	7.4	721	4.6
Low	Low	785	19.9	740	11.6
		Main Experiment (Exemplar Target)			
High	High	703	6.6	707	9.4
High	Low	676	4.2	682	3.1
Low	High	721	12.2	715	18.8
Low	Low	719	3.5	709	8.3

APPENDIX J

Mean reaction times in milliseconds and mean percent errors obtained in Experiment 6

Group		Word Type			
Exemplar Freq.	Foil Freq.	Homophone		Spelling Control	
		RT	Errors	RT	Errors
High	High	742	6.8	758	9.0
High	Low	770	14.0	763	10.4
Low	High	803	14.6	789	10.2
Low	Low	807	12.0	769	6.8

APPENDIX K

Mean reaction times in milliseconds and mean percent errors for exemplar responses in Experiments 1-3

Group		**Experiment**					
Exemplar Freq.	**Foil Freq.**	**1**		**2**		**3**	
		RT	Errors	RT	Errors	RT	Errors
High	High	570	5.5	756	11.8	799	10.4
High	Low	600	8.3	697	7.6	818	7.6
High	Non	556	0	688	4.2	810	5.6
Low	High	605	25.7	781	22.9	835	18.8
Low	Low	571	4.9	734	5.5	844	6.9
Low	Non	640	11.1	721	7.6	802	8.3

Note: All correct reaction times on exemplars were included in the latency means. Only those latencies corresponding to false positive errors on homophone foils were included in the analyses in the text.