Journal of Experimental Psychology:
Learning, Memory, and Cognition
2001, Vol. 27, No. 1, 139–156

0278-7393/01/$5.00 DOI: 10.1037//0278-7393.27.1.139

Homophone Effects in Lexical Decision

Penny M. Pexman
University of Calgary

Stephen J. Lupker and Debra Jared
University of Western Ontario

The role of phonology in word recognition was investigated in 6 lexical-decision experiments involving homophones (e.g., *MAID–MADE*). The authors' goal was to determine whether homophone effects arise in the lexical-decision task and, if so, in what situations they arise, with a specific focus on the question of whether the presence of pseudohomophone foils (e.g., *BRANE*) causes homophone effects to be eliminated because of strategic deemphasis of phonological processing. All 6 experiments showed significant homophone effects, which were not eliminated by the presence of pseudohomophone foils. The authors propose that homophone effects in lexical decision are due to the nature of feedback from phonology to orthography.

One of the long-standing issues in research on visual word recognition concerns the role of phonology in the process of recognizing words. Although it is fairly clear that phonological representations of printed words are typically generated when people read silently, there is a major debate as to whether those representations are used to activate word meanings or whether they become available after meaning has been activated, perhaps for use in sentence comprehension. A related debate concerns whether readers have strategic control over their use of phonology. That is, can readers make use of this information when it is helpful and can they deemphasize it when it hinders performance? Both of these debates must be resolved to gain a full understanding of the word recognition process.

Much of the evidence that indicates phonological representations do play a role in silent reading for meaning comes from studies that have used homophones (for a review, see Jared, Levy, & Rayner, 1999). *Homophones* are pairs of words that sound the same but are spelled differently (e.g., *MAID–MADE*). The reason homophones are particularly useful is that the visual information of a homophone activates the meaning corresponding only to the presented word (e.g., *MAID*–"servant") whereas the phonological representation of a homophone activates the meanings corresponding to both members of the homophone pair (e.g., *MAID*—"servant" and "created"). If presenting one member of the homophone pair can be shown to result in the activation of the meanings of both members, then it can be concluded that the phonological representation played a role in activating word meanings.

Many of the studies that have investigated whether phonology plays a role in reading for meaning have used what is referred to as the "homophone error paradigm." In this paradigm, stimulus displays are created such that one member of the homophone pair is correct in a context and then that member is replaced by its homophone mate. If participants do not notice the substitution, then the inference is made that the phonological representation of the presented homophone was activated, and this in turn activated the meaning associated with the correct homophone.

Typically, performance on the incorrect homophone is compared with performance on a matched nonhomophone spelling control word that does not fit in the context. Investigations using this paradigm have involved a variety of tasks including sentence verification, semantic categorization, proofreading, and eye-movement monitoring (e.g., V. Coltheart, Avons, Masterson, & Laxon, 1991; Daneman & Reingold, 1993; Daneman, Reingold, & Davidson, 1995; Daneman & Stainton, 1991; Jared et al., 1999; Jared & Seidenberg, 1991; Van Orden, 1987; Van Orden, Pennington, & Stone, 1990). The typical result is that performance with homophones is substantially worse than with nonhomophones (in terms of errors detected), although this difference arises only for low-frequency words.

Similar results have been found in other paradigms as well. For example, Jared and Seidenberg (1991, Experiment 6) demonstrated a homophone disadvantage in a syntactic judgment task. Specifically, correct decision latencies were longer for low frequency homophones (e.g., *VERB–MEET*) than for matched controls (e.g., *VERB–JOIN*).

In most of the studies mentioned above, the investigators were attempting to demonstrate that the observed homophone effects arose because phonological representations were used to activate meanings (e.g., Daneman & Stainton, 1991; Jared et al., 1999;

Penny M. Pexman, Department of Psychology, University of Calgary, Calgary, Alberta, Canada; Stephen J. Lupker and Debra Jared, Department of Psychology, University of Western Ontario, London, Ontario, Canada.

The research presented here as Experiments 1 and 2 was part of a doctoral dissertation submitted by Penny M. Pexman at the University of Western Ontario. Portions of the research described in this article were presented at the 37th Annual Meeting of the Psychonomic Society, Chicago, November 1996, and also at the 7th Annual Meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Winnipeg, Manitoba, Canada, June 1997. This research was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada.

We thank Maggie Toplak and Aicha Rouibah for their assistance in selecting the stimuli for this research; Natalie Kacinik, Tamsen Taylor, Sandra Hessels, and Christine Barberic for testing participants; Derek Besner for providing the word stimuli used in Davelaar, Coltheart, Besner, and Jonasson (1978); and Sandy Pollatsek, Max Coltheart, and an anonymous reviewer for many helpful comments.

Correspondence concerning this article should be addressed to Penny M. Pexman, Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada. Electronic mail may be sent to pexman@ucalgary.ca.

Jared & Seidenberg, 1991; Van Orden et al., 1990). However, none of this research has ruled out the possibility that meanings were first activated by their spellings and that homophone effects arose subsequently, perhaps while readers were determining whether the word fit into the context. In the present research, we investigated homophone effects in a silent reading task involving shallower processing, specifically, the lexical-decision task. In the lexical-decision task, readers are not required to continue processing until they have understood the word in its context (e.g., whether it is a member of a designated category, whether it makes sense in the sentence, etc.). Thus, if homophone effects arise in lexical decision, that finding would be evidence that phonology makes a substantial contribution early in the word-recognition process.

HOMOPHONE EFFECTS IN LEXICAL DECISION

In what appears to be the first lexical-decision experiment to use homophones, Rubenstein, Lewis, and Rubenstein (1971) reported that response times (RTs) were longer for homophones than for nonhomophonic control words. This homophone effect arose only for low-frequency words. Rubenstein et al. interpreted their results by suggesting that under normal circumstances, readers generate a phonological code for each letter string that is presented. When the letter string is a homophone, this phonological code causes both lexical entries to be selected. The more frequent member of the homophone pair will usually be analyzed first. If the lower frequency member of the pair was actually presented, then that analysis, which includes a spelling check, will fail. A positive lexical-decision response will be made only when the lexical entry for the lower frequency member of the pair has been analyzed and has passed the spelling check. The delay caused by having to check an incorrect lexical entry first is what causes longer response times for low-frequency homophones.

The homophone effect reported by Rubenstein et al. (1971) had important implications because it suggested that (a) phonological codes were generated prelexically, that is, prior to the correct lexical entry being selected, and (b) those codes then drove the lexical selection process. The impact of the finding, however, was diminished when Clark (1973) pointed out that Rubenstein et al.'s effect was not significant in an analysis that treated both subjects and items as random factors. Thus, Clark suggested that Rubenstein et al.'s results may have been due to idiosyncrasies in some of their stimuli—not to the effects of homophony. Others (Cohen, 1976; Keppel, 1976; Smith, 1976; Wike & Church, 1976), however, pointed out that Clark's concern about item idiosyncrasies cannot actually be addressed by a statistical procedure but, instead, must be addressed by replication using different stimulus sets.

One attempt to replicate the findings of Rubenstein et al. (1971) was reported by M. Coltheart, Davelaar, Jonasson, and Besner (1977), who used a larger set of homophones. Unlike Rubenstein et al.'s stimuli, all of the homophones were substantially lower in frequency than their homophone mates (although not all were low-frequency words), which should have maximized the likelihood of observing a homophone effect according to Rubenstein et al.'s account. However, M. Coltheart et al. found no trace of a homophone effect and concluded that, at least in their experiment, lexical decisions for words were being made through direct access to the lexicon rather than through phonological mediation. Dennis, Besner, and Davelaar (1985, Experiment 4) also failed to find a homophone effect in a lexical-decision task using M. Coltheart et al.'s complete set of 39 homophones and matched control words. However, when Dennis et al. used a subset of 25 of the pairs in another study (Experiment 3), they did find a significant homophone effect in both the RT data and the error data (although the RT effect was not significant in an items analysis).

These results suggest that if homophone effects can be obtained in lexical-decision tasks, at the very least, they are somewhat fragile. However, if homophone effects cannot be obtained in lexical decision, even for low-frequency words, the implication would be that the homophone effects observed in tasks requiring activation of meaning are due to a relatively late influence of phonology. In the next section, we consider the possibility that homophone effects in lexical decision occur only in certain list contexts and not in others.

ARE HOMOPHONE EFFECTS SPORADIC BECAUSE THE USE OF PHONOLOGY IS OPTIONAL?

Davelaar, Coltheart, Besner, and Jonasson (1978) hypothesized that homophone effects occur in some experiments but not in others because using phonological codes during word recognition is optional for readers. To investigate this idea, they manipulated the nature of the pseudoword foils.[1] Their reasoning was that pseudohomophone foils, which sound like real words when pronounced (e.g., *GRONE*), should discourage the use of phonology because their phonological codes would produce activity in the lexical units of real words, making the word–nonword decision difficult. As a result, lexical units would tend to be selected instead on the basis of orthographic codes, which should eliminate homophone effects for words. In contrast, when foils are standard pronounceable pseudowords (e.g., *SLINT*), it should be more viable for readers to use prelexical phonology during lexical selection, and therefore a homophone effect should be observed. Their results supported this hypothesis. When the foils were pronounceable pseudowords, Davelaar et al. obtained a significant homophone effect for homophones that were the lower frequency member of the homophone pair but not for homophones that were the higher frequency member. There were no homophone effects for either type of homophone when the foils were pseudohomophones. Davelaar et al. concluded that using prelexical phonological codes in the process of lexical selection is an option for readers and that in lexical-decision tasks homophone effects arise unless readers choose not to use such a strategy. The implication is that a homophone effect was not obtained in the M. Coltheart et al. (1977) study because participants adopted the strategy of selecting a lexical entry on the basis of orthographic representations.

There are several reasons to question whether the Davelaar et al. (1978) study provides strong evidence for the optional use of phonology in the lexical-decision task and, hence, in the word recognition process itself. The first concern is that different homophones and control words were used in their pseudoword and pseudohomophone foil conditions, and thus it is possible that a homophone effect did not appear with the pseudohomophone foils because of the particular word stimuli in that condition. A potentially important difference is that the homophones in the

[1] For present purposes, the term *pseudowords* is used to refer to nonwords that are pronounceable but non-pseudohomophonic. Examples of these nonwords are *CLANE, BRAX,* and *SHART.*

pseudohomophone foil condition were somewhat more frequent than the homophones in the pseudoword foil condition. In the pseudohomophone foil condition, 30% of homophones had a frequency of 40 or more per million (Kučera & Francis, 1967), compared with only 3% (one word) in the pseudoword foil condition; 57% had a frequency of 15 per million or less in the pseudohomophone foil condition, compared with 83% in the pseudoword foil condition. It is possible, then, that no effect of homophony was observed in the pseudohomophone foil condition because homophone effects occur only for low-frequency words and many of the stimuli in that condition were reasonably frequent.

A second concern is that Davelaar et al.'s (1978) participants responded more quickly when the foils were pseudohomophones than when they were pseudowords. Most studies in the literature have reported that participants responded more slowly when foils were pseudohomophones (e.g., Berent, 1997; Ferrand & Grainger, 1996; Gibbs & Van Orden, 1998; James, 1975; Parkin & Ellingham, 1983; Stone & Van Orden, 1993; although not Andrews, 1982; Pugh, Rexer, & Katz, 1994). These discrepant results may be due to differences in the word-likeness of the pseudohomophones. If Davelaar et al.'s (1978) pseudohomophones contained unusual orthographic patterns (the foils were not included in the Appendix), then participants may have been able to make their lexical decisions primarily on the basis of visual information rather than on the basis of stored lexical information. If so, then the results from the pseudohomophone condition tell us little about the information readers use to recognize words.

A third concern with Davelaar et al.'s (1978) study was that the pseudohomophone foil condition immediately followed the pseudoword foil condition for all participants in the study. Therefore, the participants would always have been more practiced in the pseudohomophone foil condition than the pseudoword foil condition, possibly attenuating homophone effects.

These three concerns suggest that we should be cautious about accepting Davelaar et al.'s (1978) conclusion that inconsistent findings in lexical-decision studies of homophone effects indicate that readers can alter their reliance on phonology. In the next section, we consider other evidence for strategic control of phonology in the lexical decision task.

OTHER STUDIES INVESTIGATING STRATEGIC CONTROL OF PHONOLOGICAL PROCESSING IN LEXICAL DECISION

Other studies have investigated whether readers can deemphasize phonological processing in the lexical-decision task using indicators besides changes in the size of the homophone effect. One group of studies observed changes in the size of the difference between pseudohomophones and matched pseudoword controls as a function of either the proportion of pseudohomophones in the study (McQuade, 1981), or the presence of homophones among the words (Dennis et al., 1985; Underwood, Roberts, & Thomason, 1988). Unfortunately, findings from nonword responses may not be relevant for determining whether readers can strategically control word-recognition processes (Gibbs & Van Orden, 1998). One reason is that rejection latencies for nonwords are typically much longer than acceptance latencies for words, suggesting that additional processing is occurring (M. Coltheart et al., 1977).

Another group of studies examined whether the size of phonological priming effects is influenced by the type of foils included in the study (Berent, 1997; Ferrand & Grainger, 1992, 1996). In these studies, the primes were either phonologically related pseudohomophones of the target word (e.g., *SUNE–SOON*) or were pseudowords phonologically unrelated to the target word (e.g., *BARP–SOON*). A phonological priming effect is observed when lexical decisions to target words are faster when they are preceded by phonologically related pseudohomophone primes than by phonologically unrelated pseudoword primes. If participants restrict the availability of phonology in the presence of pseudohomophone foils, then phonological priming effects should be smaller with pseudohomophone foils than with pseudoword foils. However, Berent and Ferrand and Grainger found similar-sized phonological priming effects with the two types of foils, suggesting that participants did not deemphasize phonological processing in pseudohomophone foil conditions.

Other researchers have focused on word responses using spelling–sound consistency or regularity effects as an indicator of phonological processing. Spelling–sound consistency effects involve slower RTs for words that are pronounced differently than their body neighbors (e.g., *HAVE* is pronounced differently than *SAVE, WAVE, CAVE,* etc.), compared with RTs for words that are pronounced like their body neighbors (e.g., *LINE* is pronounced like *WINE, MINE,* etc.). Regularity effects involve slower RTs for words that do not follow rules of spelling–sound translation (e.g., *PINT*), compared with RTs for words that do follow those rules (e.g., *NAME*). The results of five of these studies suggest that readers cannot alter their reliance on phonology in situations where it would be to their advantage to do so. In two of these studies, there was no change in the size of these effects when the foils included pseudohomophones compared with legal pseudowords (Andrews, 1982; Parkin & Ellingham, 1983), whereas in two others these effects were larger, not smaller, with pseudohomophone foils (Berent, 1997; Gibbs & Van Orden, 1998). Finally, Waters and Seidenberg (1985) found that including words with unusual spelling–sound correspondences (e.g., *TONGUE*), which should have caused participants to reduce their reliance on phonology, actually produced a low-frequency regularity effect for the same words that did not produce a regularity effect when the strange words were absent. This set of results suggests that instead of reducing reliance on phonology, pseudohomophones (word-like nonwords) and words with unusual spelling–sound correspondences (nonword-like words) may actually increase the role of phonology in a lexical-decision task.

In contrast, Pugh et al. (1994) claimed, on the basis of their results, that readers can reduce their reliance on phonology in the lexical-decision task. In one experiment, their participants showed a significant effect of number of unfriendly word–body neighbors on RTs for low-frequency words when foils were pseudowords but not when foils were pseudohomophones. Unfriendly word-body neighbors are words that have the same body as a target word but are pronounced differently (e.g., *PINT*'s unfriendly word–body neighbors include *MINT* and *HINT*). However, these findings must be interpreted cautiously. Aside from contradicting the findings above that indicate that readers cannot reduce their reliance on phonology, they are incompatible with numerous lexical-decision studies that have failed to find consistency or regularity effects for low-frequency words when the foils are legal pseudowords and no strange words are included (e.g., Berent, 1997; Gibbs & Van Orden, 1998; Jared, McRae, & Seidenberg, 1990; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Waters & Seidenberg, 1985).

In Pugh et al.'s final experiment in which participants were shown two stimuli and made a positive response only if both were words, Pugh et al. found inhibition for pairs of words that were orthographically similar but phonologically dissimilar (e.g., *COUCH–TOUCH*), compared with unrelated controls when the foils were pseudowords. When the foils were pseudohomophones, however, these word pairs produced facilitation. They interpreted the inhibition in the pseudoword foil condition as evidence for the involvement of phonology in word recognition and the facilitation in the pseudohomophone foil condition as an indication that participants had reduced their reliance on phonology.

Gibbs and Van Orden (1998) suggested that Pugh et al.'s (1994) apparent evidence that participants can reduce their reliance on phonology in the presence of pseudohomophone foils was obtained because many of Pugh et al.'s pseudohomophones had strange spellings and were less word-like than their pseudowords. Hence, participants in the pseudohomophone foil condition may have been able to respond superficially on the basis of orthographic familiarity. Indeed, Shulman, Hornak, and Sanders (1978) demonstrated that a facilitation effect is observed for *COUCH–TOUCH* pairs when orthographically illegal foils (random letter strings) are used. Further, when Pexman, Lupker, Jared, Toplak, and Rouibah (1996) attempted a replication of Pugh et al.'s study with pseudohomophones that were more word-like, *COUCH–TOUCH* pairs produced inhibition rather than facilitation, indicating that phonology had not been deemphasized. It is also worth noting that even when using the exact same stimuli as Pugh et al. (i.e., same word pairs and foils), Pexman et al. failed to obtain a facilitation effect for *COUCH–TOUCH* pairs in the pseudohomophone foil condition. Thus, until a replication of Pugh et al.'s facilitation effect is produced, their result should be interpreted with caution.

One final study that is relevant to this discussion used a semantic categorization task instead of a lexical-decision task. Jared and Seidenberg (1991, Experiment 3) found that increasing the proportion of homophones in the study did not change the rate of errors to low-frequency homophone foils (e.g., *LIVING THING–*ROWS) compared with spelling controls. If participants had deemphasized phonological processing when a higher proportion of homophones was used, the result should have been a decrease in the difference in error rates between homophone foils and spelling controls.

THE PRESENT RESEARCH

The results discussed above provide only minimal evidence that readers can reduce their reliance on phonological information when performing a lexical-decision task. Determining whether readers can exert this kind of strategic control is one of the main goals of the present investigation. The first step, however, is to determine whether homophone effects can actually be obtained in a lexical-decision task. Our analysis of the literature revealed several important considerations for the design of the present studies. One was the frequency of the homophones. If findings from studies using tasks that require activation of meaning generalize to lexical decision, then homophone effects in lexical decision should be strongest for low-frequency words. Thus, in all of the present experiments except Experiment 6, the lower frequency member of the homophone pair was always a low-frequency word.

The second consideration was the word-likeness of the foil stimuli. If foils are not sufficiently word-like orthographically, then participants may be able to make their decisions on the basis of a superficial analysis of the orthography of the stimuli before they have determined whether the stimulus is a word. In that case, the results would tell us little about the processes involved in word recognition. This factor was investigated in Experiments 1 and 2.

The third consideration was that our pseudohomophone stimuli did indeed function as pseudohomophones for our participants. To ensure this, we pretested them in both a naming experiment and a pseudohomophone identification task. Only those pseudohomophones that were named quickly and accurately and that were identified as "sounding like words" were included in the main studies.

The fourth consideration was the comparability of the word stimuli in the pseudoword and pseudohomophone foil conditions. To determine whether participants are able to deemphasize phonology in some situations, it is extremely important that both the word and foil stimuli be carefully matched in pseudoword and pseudohomophone conditions. In our studies, the word stimuli were counterbalanced across participants so that each word appeared with pseudoword foils for half of the participants and with pseudohomophone foils for the other half. This avoids the problem with the Davelaar et al. (1978) study in which a different set of homophones and controls were used in each foil condition. As well, in the present studies, the pseudowords and the pseudohomophones were matched for word body (e.g., *CADE* and *RADE*). This ensured that the foils were not more word-like in one condition than in the other.

The fifth consideration was the manner in which the two foil conditions were presented. We wanted to avoid the problem of the Davelaar et al. (1978) study in which the pseudohomophone foil condition always followed the pseudoword foil condition, in case changes in performance in the pseudohomophone foil condition were a result of practice or fatigue rather than a change in emphasis on phonology. Thus, in Experiments 1 and 2 we used a between-subjects design. However, in Experiment 3 we switched to a within-subjects design to determine whether such a design was necessary to observe strategic deemphasis of phonology.

The sixth consideration was the frequency composition of the experimental lists. Davelaar et al. (1978) obtained evidence for strategic deemphasis of phonology in a study in which all words were low in frequency. Our initial studies mixed high- and low-frequency words, but in Experiments 4 and 5 we included only low-frequency words to determine whether participants can be made to deemphasize phonology only when many words on the list cannot be recognized quickly on the basis of their spellings.

In Experiment 5 we examined whether Davelaar et al.'s (1978) findings were specific to the homophone and control words used in their study. Unable in our first four experiments to replicate their finding that homophone effects disappear with pseudohomophone foils, we explored the possibility that the homophone and control word pairs used in their *GRONE* condition would not produce a homophone effect even with pseudoword foils.

Finally, to determine precisely which types of homophones generate homophone effects, in Experiment 6 we included the two types of homophones that were not examined in the previous experiments. These were low-frequency homophones with low-frequency mates (*low–low* pairs) and high-frequency homophones with higher frequency homophone mates (*high–higher* pairs). A summary of the main characteristics of each experiment is presented in Table 1.

Experiment 1

Method

Participants

The participants for all of the studies reported in this article were undergraduate students at the University of Western Ontario who received partial course credit in an introductory psychology course or payment for their participation. No participants took part in more than one of the experiments. All participants considered English to be their native language and had normal or corrected-to-normal vision. Sixty-four students participated in Experiment 1; however the data from 2 participants were discarded because of very high error rates. Both participants had more than 30% errors for foil responses.

Stimuli

A set of 160 English words was used in Experiments 1 through 3 (see Appendix). The set consisted of 40 pairs of homophones and 80 matched nonhomophone control words. One member of each homophone pair was a low-frequency word (frequency less than 32 per million; Kučera & Francis, 1967) and the other was a high-frequency word (frequency greater than 40 per million). Both homophones in a pair had the same initial letter. Homophones and nonhomophones were matched as closely as possible for frequency, length, first letter, and neighborhood size using Coltheart's *N* (M. Coltheart et al., 1977).[2] Homophones and nonhomophones could not be matched for word body (e.g., homophone *HAIR* and nonhomophone *LAIR*) because it would have been impossible to do so and still maintain the required control of word frequency. Mean frequency and neighborhood size for each word type are presented in Table 2.

Two lists of word stimuli were created such that in each list there were 20 low-frequency homophones, 20 corresponding low-frequency control words, 20 high-frequency homophones, and 20 corresponding high-frequency control words. Each list contained only one member of a homophone pair. The two lists were matched as closely as possible for mean frequency and neighborhood size in an effort to prevent list effects. In addition, word body repetitions were kept to a minimum within each list.

The foil stimuli consisted of 60 pseudohomophones and 60 pronounceable pseudowords matched for word bodies (e.g., *RADE* and *CADE*). This matching minimized the orthographic differences between the two types of foils. To ensure that the foils had spelling patterns that were typical of English, they contained word bodies that occurred in at least two English words. As well, the pseudohomophone foils were constructed so as to maximize the likelihood that their phonological representations could be generated quickly and accurately by participants. To ensure that this was the case, a rigorous stimulus development process was used.

We began by choosing monosyllabic word bodies that occurred in at least two English words. Only consistent word bodies were selected, thus, they had only one correct pronunciation. For each word body, various first letters were added to make viable pseudowords and pseudohomophones. All pseudohomophones sounded like a real word that had a frequency of greater than 10 per million so that participants would be almost certain to know the base word. The initial set consisted of 286 pseudowords and pseudohomophones.

Pilot Study 1. A naming study was run with the 286 pseudowords and pseudohomophones to determine whether any of these letter strings had inordinately long naming latencies, which might indicate that it was difficult to generate a phonological code for that letter string. There were 33 participants. The naming latencies were quite uniform for the 286 stimuli (pseudowords: M = 626.2, SD = 46.2; pseudohomophones: M = 603.9, SD = 39.5). Only 6 stimuli (all pseudowords) were removed from the list because of long naming latencies. The list of foil stimuli was limited further by including no more than two pseudowords or pseudohomophones with the same word body. This additional restriction pared the list to 192 letter strings. Mean naming latencies were recalculated for these 192 stimuli (pseudowords: M = 615.1, SD = 31.2; pseudohomophones: M = 598.3, SD = 38.2).

Pilot Study 2. The 192 pseudowords and pseudohomophones were printed on paper in random order. Twenty-three participants were asked to circle the letter strings that sounded like real words when pronounced. We selected pseudohomophones that were circled by at least 20 of the 23 participants and pseudowords that were circled by no more than 3 of the 23 participants. The final set consisted of 60 pseudoword–pseudohomophone pairs that were matched for word bodies.

Two experimental lists were created by adding the 60 pseudowords to each of the two word lists, and two more were created by adding the 60 pseudohomophones to each of the word lists.

Procedure

In each experiment an IBM PC-clone computer was used to control the presentation of the stimuli and time responses. Stimuli were presented one at a time in the center of the computer monitor in uppercase 12-point font. Participants were asked to indicate whether each letter string was a word or nonword by pressing either the left button (labeled *NONWORD*) or the right button (labeled *WORD*) on a response box. This arrangement was reversed for left-handed participants. Participants were told to make their decisions as quickly and as accurately as possible.

At the beginning of each trial there was a 50-ms, 400-Hz beep signal, and then a fixation point appeared in the center of the screen. The fixation point was displayed for 500 ms, followed by a 500-ms blank screen and then the presentation of the target. The target remained on the screen until the participant responded.

Participants completed 16 practice trials and were given verbal feedback if they responded incorrectly to any of the practice stimuli. They then received one of the four experimental lists; 16 participants were given each list (although recall that the data from 2 participants were not analyzed). The stimuli within a list were presented in a different random order for each participant.

Results

In this and all subsequent experiments, the RTs and response error data were examined with analyses in which subjects (F_s) and items (F_i) were separately treated as random factors. Planned

[2] Most homophone studies have used the homophone error paradigm rather than the lexical-decision task. Recall that in the homophone error paradigm one member of the homophone pair is chosen to be correct and then a context is created to fit that homophone. The homophone is then removed and the other member of the homophone is put in its place. In this paradigm it is important that the control words be "spelling controls," that is, words that are as similar in spelling to the correct homophone as the homophone foils are. If homophone foils are then detected less frequently than control words, we can rule out the possibility that this was simply due to the fact that homophone foils are similar in spelling to the correct homophone. In lexical decision, the situation is quite different. Stimuli are presented individually and all the homophones (and control words) that are presented are "correct" because they are all real words. Thus, the most appropriate control words for homophones in a lexical-decision task are words that have the same frequency as the homophones. Note that we also matched our homophones and controls on initial letter, length, and orthographic neighborhood size. Numerous lexical-decision studies have shown that orthographic neighborhood size has a significant impact on RTs, particularly for low-frequency words (see Andrews, 1997, for a review).

Table 1
Summary of Variables Manipulated in Experiments 1 Through 6

Experiment	Frequency of homophones	Word-likeness of foils	Design	Source of stimuli
Experiment 1	low–high high–low	word-like	between-subjects	chosen for Experiment 1
Experiment 2	low–high high–low	less word-like	between-subjects	same as Experiment 1
Experiment 3	low–high high–low	word-like	within-subjects	same as Experiment 1
Experiment 4	low–high	word-like	within-subjects	only low–high from Experiment 1
Experiment 5	low–high	word-like	within-subjects	Davelaar et al. (1978)
Experiment 6	low–low low–high high–higher	word-like	between-subjects	chosen for Experiment 6

Note. low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*); high–low = high-frequency homophones with low-frequency homophone mates (e.g., *MADE–MAID*); low–low = low-frequency homophones with low-frequency homophone mates (e.g., *BAIL–BALE*); high–higher = high-frequency homophones with higher frequency homophone mates (e.g., *FOUR–FOR*).

comparisons were conducted in each foil condition to determine whether each homophone effect was significant. This involved comparing RTs and errors for each type of homophone with RTs and errors for the corresponding control words. These planned comparisons were one-tailed tests, and significant comparisons are indicated by an asterisk in the tables of means. In all analyses, a value of $p < .05$ was used to determine statistical significance. In all experiments, a trial was excluded from the analyses of RTs if an incorrect response was made or if the RT for the trial was more than two standard deviations from the participant's mean RT for that condition. In this experiment, response errors occurred on 5.8% of trials, and data from 3.8% of trials were removed because of extremely long or short RTs. Mean RTs and response error percentages for Experiment 1 are presented in Table 3.

Word Responses

Homophones had significantly longer RTs than matched control words, $F_s(1, 58) = 16.08$, $MSE = 3{,}851.25$; $F_i(1, 76) = 6.01$, $MSE = 14{,}367.24$, and produced more errors, $F_s(1, 58) = 20.26$, $MSE = 26.76$, and $F_i(1, 76) = 2.91$, $p = .08$, $MSE = 200.03$. The homophone effect was larger for low-frequency words than for high-frequency words, although the interaction of homophony and frequency was significant only by subjects in both the RT, $F_s(1, 58) = 9.99$, $MSE = 3{,}225.93$; $F_i(1, 76) = 3.23$, $p = .08$, $MSE = 14{,}367.24$, and error data, $F_s(1, 58) = 6.67$, $MSE = 33.37$; $F_i < 1.5$. Contrary to Davelaar et al.'s (1978) results, the homophone effect was larger, not smaller, with pseudohomophone foils than with pseudoword foils. The interaction of homophony and foil condition was only marginally significant in the RT data, $F_s(1, 58) = 2.48$, $p = .12$, $MSE = 3{,}851.25$; $F_i(1, 76) = 3.19$, $p = .08$, $MSE = 3{,}429.74$, but was significant in the error data, $F_s(1, 58) = 10.23$, $MSE = 26.76$; $F_i(1, 76) = 5.62$, $MSE = 60.56$. The three-way interaction of foil condition, frequency, and homophony was not significant in the RT data, $F_s < 1.5$; $F_i < 1.5$, indicating that the pattern of a larger homophone effect for low-frequency words was present in both foil conditions. However, the three-way interaction was significant in the error data, $F_s(1, 58) = 5.58$, $MSE = 33.37$; $F_i(1, 76) = 3.50$, $MSE = 60.56$; the difference in error rates for low-frequency homophones and controls was much larger in the pseudohomophone foil condition than in the pseudoword foil condition.

RTs for words were significantly longer with pseudohomophone foils than with pseudoword foils, $F_s(1, 58) = 4.66$, $MSE = 59{,}765.70$; $F_i(1, 76) = 27.22$, $MSE = 3{,}807.48$, and significantly more errors were made with pseudohomophone foils, $F_s(1, 58) = 6.86$, $MSE = 55.88$; $F_i(1, 76) = 9.46$, $MSE = 56.39$. There was a significant main effect of frequency in the RT data, $F_s(1, 58) = 170.57$, $MSE = 8{,}073.06$; $F_i(1, 76) = 154.16$, $MSE = 14{,}241.19$, and error data, $F_s(1, 58) = 151.29$, $MSE = 34.45$; $F_i(1, 76) =$

Table 2
Mean Frequencies and Mean Neighborhood Sizes for Word Stimuli

	Target homophone		Homophone mate
Stimulus type	*M* frequency	*M* neighborhood size	*M* frequency
Experiments 1–4			
High–low homophone	160	9	8
High-frequency control	160	8	—
Low–high homophone	8	8	160
Low-frequency control	7	7	—
Experiment 6			
High–higher homophone	80	9	1026
High-frequency control	80	8	—
Low–high homophone	8	9	241
Low-frequency control	9	8	—
Low–low homophone	9	8	9
Low-frequency control	9	7	—

Note. High–low = high-frequency homophones with low-frequency homophone mates (e.g., *MADE–MAID*); low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*); High–higher = high-frequency homophones with higher frequency homophone mates (e.g., *FOUR–FOR*); Low–low = low-frequency homophones with low-frequency homophone mates (e.g., *BAIL–BALE*). Dashes indicate control words that did not have homophone mates.

Table 3
Mean Lexical-Decision RTs (and Error Percentages) and Homophone Effects for Experiments 1 and 2

	Pseudoword foil condition			Pseudohomophone foil condition		
		Homophone effect			Homophone effect	
Stimulus type	RT	RT	Errors	RT	RT	Errors
Experiment 1 (word-like foils)						
High–low homophone	630 (1.71)			683 (2.50)		
High-frequency control	634 (1.09)	−4	0.61	660 (1.16)	23*	1.34*
Low–high homophone	774 (8.75)			874 (16.49)		
Low-frequency control	732 (7.81)	42*	0.94	813 (8.33)	61*	8.16*
Foil	824 (4.22)			941 (6.72)		
Experiment 2 (less word-like foils)						
High–low homophones	558 (1.09)			644 (1.25)		
High-frequency control	567 (0.62)	−9	0.47	638 (1.25)	6	0.00
Low–high homophones	655 (3.91)			785 (10.32)		
Low-frequency controls	641 (4.68)	14	−0.77	746 (8.12)	39*	2.20
Foil	709 (2.50)			811 (3.64)		

Note. Response time (RT) is given in milliseconds. High–low = high-frequency homophones with low-frequency homophone mates (e.g., *MADE–MAID*); Low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*).
* $p < .05$.

44.38, $MSE = 165.54$, as well as an interaction of frequency with foil condition in both the RT, $F_s(1, 58) = 5.22$, $MSE = 8{,}073.06$; $F_i(1, 76) = 13.22$, $MSE = 3{,}807.48$, and error data, $F_s(1, 58) = 7.24$, $MSE = 34.45$; $F_i(1, 76) = 5.39$, $MSE = 56.39$, indicating that the frequency effect was larger with pseudohomophone foils than with pseudoword foils (replicating Stone & Van Orden, 1993).

Foil Responses

Participants took significantly longer to reject pseudohomophone foils than pseudoword foils, $F_s(1, 58) = 5.72$, $MSE = 36{,}978.53$; $F_i(1, 118) = 3.78$, $MSE = 5{,}012.34$, and they made more errors on pseudohomophone foils, $F_s(1, 58) = 3.89$, $MSE = 26.15$; $F_i(1, 118) = 3.34$, $p = .07$, $MSE = 55.67$.

Discussion

A significant homophone effect was found for low-frequency words, indicating that phonology played a role in the recognition of these words. More important, the effect was not eliminated with pseudohomophone foils as in Davelaar et al. (1978). With pseudohomophone foils, RTs were slower and the homophone effect was larger and also appeared for high-frequency homophones. Thus, the present results argue strongly against Davelaar et al.'s claim that participants can strategically deemphasize phonology in the lexical-decision task.

Experiment 2

In Experiment 2 we investigated whether the word-likeness of the foils has an impact on the size of homophone effects. If foils do not resemble English words, then participants may be able to make their decisions on the basis of a superficial orthographic analysis. If so, then phonological effects should be smaller. Such a finding would provide a possible explanation for some of the failures to find homophone effects in lexical decision. It would also provide an account of the conflicting results concerning the influence of pseudohomophone foils if in some studies the pseudohomophones were less word-like than the pseudowords.

Experiment 2 was a replication of Experiment 1 but with less word-like foils. As before, pseudoword and pseudohomophone foils were matched for word body.

Method

Participants

There were 64 participants in this study.

Stimuli

The two word lists from Experiment 1 were used, but the foils differed. The foil stimuli here contained bodies that did not occur in any real English words, although they were pronounceable. These less word-like pseudowords and pseudohomophones are the kind that generally have been used in previous studies, although the authors of those studies were not intentionally choosing bodies that were not real. We took potential bodies from the following studies: Lukatela and Turvey (1991), Martin (1982), McCann and Besner (1987), McQuade (1983), Pring (1981), Pugh et al. (1994), Stone and Van Orden (1993), and Taft (1982). By adding first letters to these bodies we created a list of 172 pseudowords and pseudohomophones.

Pilot Study 3. The 172 pseudowords and pseudohomophones were printed on paper in random order. Twenty-three participants were asked to circle the letter strings in the list that sounded like real words. Because the letter strings in this list were all spelled with nonreal bodies, they did not look very word-like. As a result, participants were much less likely to identify the pseudohomophones correctly than were the participants who evaluated the list in Pilot Study 2. To select the same number of stimuli that had been used in Experiment 1, the criteria had to be less stringent. Pseudohomophones were selected if at least 15 of the 23 participants circled them, and pseudowords were selected if no more than 6 of the 23 participants circled them. The list of foil stimuli was limited further by including no more than two pseudowords or pseudohomophones with the same body. The final set consisted of 60 pseudoword-pseudohomophone pairs that were matched for word bodies (see the Appendix).

Two experimental lists were created by adding the 60 pseudowords to each of the two word lists, and two more were created by adding the 60 pseudohomophones to each of the word lists.

Procedure

The procedure was the same as in Experiment 1.

Results

Response errors occurred on 3.5% of trials, and data from 3.9% of trials were removed because of extremely long or short RTs. Mean RTs and response error percentages for Experiment 2 are presented in Table 3.

Word Responses

As in Experiment 1, homophones had significantly longer RTs than matched control words, although here that effect was significant only by subjects, $F_s(1, 60) = 4.86$, $MSE = 2,296.38$; $F_i < 1.5$, and the homophone effect was not significant in the error data ($F_s < 1.5$; $F_i < 1.5$). The homophone effect was larger for low-frequency words, although the interaction of homophony and frequency was significant only in the RT data and only by subjects, $F_s(1, 60) = 4.17$, $MSE = 3,834.49$; $F_i(1, 76) = 1.66$, $p = .20$, $MSE = 6,417.29$. The homophone effect was larger with pseudohomophone foils although, again, only by subjects and only in the RT data, $F_s(1, 60) = 3.89$, $MSE = 2,296.38$; $F_i(1, 76) = 1.92$, $p = .16$, $MSE = 3,416.36$. The three-way interaction of foil condition, frequency, and homophony was not significant in the RT data ($F_s < 1.5$; $F_i < 1.5$) but was significant by subjects in the error data, $F_s(1, 60) = 3.98$, $MSE = 12.12$; $F_i < 1.5$: The difference in error rates between low-frequency homophones and control words was somewhat larger in the pseudohomophone condition.

RTs for words were significantly longer with pseudohomophone foils, $F_s(1, 60) = 6.67$, $MSE = 93,319.02$; $F_i(1, 76) = 261.00$, $MSE = 3,358.16$, and significantly more errors were made with pseudohomophone foils, $F_s(1, 60) = 11.31$, $MSE = 43.86$; $F_i(1, 76) = 18.04$, $MSE = 38.21$. There was a significant main effect of frequency in the RT data, $F_s(1, 60) = 74.90$, $MSE = 9,528.60$; $F_i(1, 76) = 94.89$, $MSE = 10,776.48$, and the error data, $F_s(1, 60) = 91.35$, $MSE = 25.24$; $F_i(1, 76) = 31.50$, $MSE = 101.68$, as well as an interaction of foil condition and frequency that approached significance by subjects in the RT data, $F_s(1, 60) = 2.89$, $p = .09$, $MSE = 9,528.06$; $F_i(1, 76) = 14.78$, $MSE = 3,358.16$, and was significant in the error data, $F_s(1, 60) = 14.31$, $MSE = 25.24$; $F_i(1, 76) = 12.60$, $MSE = 38.21$, indicating that the frequency effect was larger with pseudohomophone foils.

Foil Responses

Participants took significantly longer to reject pseudohomophone foils, $F_s(1, 60) = 3.99$, $MSE = 42,144.35$; $F_i(1, 118) = 71.63$, $MSE = 9,125.96$, and made more errors on pseudohomophone foils, $F_s(1, 60) = 3.16$, $p = .07$, $MSE = 10.47$; $F_i(1, 118) = 3.05$, $p = .07$, $MSE = 14.35$.

Discussion

Homophone effects were smaller in Experiment 2 with less word-like foils than they were with the more word-like foils in Experiment 1, and they were not significant even for low-frequency words in the pseudoword foil condition. This, along with the generally faster RTs here than in Experiment 1, suggests that participants made decisions earlier in processing, perhaps based on a very cursory analysis of the letter strings on some occasions. As in Experiment 1, there was no evidence that participants strategically deemphasized phonology in the presence of pseudohomophones because again the homophone effect was larger, not smaller, with pseudohomophone foils.

Experiment 3

Strategic deemphasis of phonology may not have been obtained in Experiments 1 and 2 because participants were only presented with one foil condition, whereas Davelaar et al.'s (1978) participants were all presented first with the pseudoword foil condition and then the pseudohomophone foil condition. Perhaps strategic deemphasis of phonology is observed only when participants have already developed a task strategy that they discover to cause problems when the foil type is changed. To test this idea, Experiment 3 used a within-subject manipulation of foil type. Half of the participants were presented with the foil conditions in the same order as in the Davelaar et al. experiments, and half were presented with them in the reverse order.

Method

Participants

There were 40 participants in this study.

Stimuli

For Experiment 3, each of the lists of word stimuli from Experiment 1 was divided in two. Ten of each type of word appeared in each half list. Forty pseudowords from Experiment 1 were included to be used with one half list and the 40 matched pseudohomophones were included to be used with the other half list. Forty filler stimuli were also used; 20 were words, 10 were pseudowords, and 10 were pseudohomophones.

Procedure

The procedure was the same as in Experiment 1 except for the order of presentation of the experimental stimuli. Half of the participants were presented with, first, the 40 words from one half of a list and the 40 pseudowords, in random order. Next, they were presented with 10 filler words and 10 pseudowords, in random order. Then they were presented with 10 filler words and 10 pseudohomophones, in random order. Finally, they were presented with 40 words from the other half of a list and the 40 pseudohomophones, in random order. For the other half of the participants, the order of presentation of the foil types was reversed. Presentation of the half word lists was counterbalanced such that each was presented equally often with pseudoword and pseudohomophone foils, and each was presented equally often in the first and second parts of the experiment.

Results

Response errors occurred on 6.9% of trials, and data from 2.4% of trials were removed because of extremely long or short RTs. Mean RTs and response error percentages for Experiment 3 are presented in Table 4.

Word Responses

Homophones had significantly longer RTs than matched control words, $F_s(1, 32) = 18.89$, $MSE = 12{,}200.69$; $F_i(1, 72) = 8.32$, $MSE = 45{,}629.06$, and produced more errors, $F_s(1, 32) = 14.24$, $MSE = 28.44$; $F_i(1, 72) = 1.54$, $p = .21$, $MSE = 467.91$, although the homophone effect in the error data was not significant by items. The homophone effect was larger for low-frequency words than for high-frequency words, and hence there was a significant interaction of homophony and frequency in the RT data, $F_s(1, 32) = 26.29$, $MSE = 7{,}166.87$; $F_i(1, 72) = 5.91$, $MSE = 45{,}629.07$, although not in the error data, $F_s(1, 32) = 2.63$, $p = .11$, $MSE = 38.44$; $F_i < 1.5$. The homophone effect was again larger with pseudohomophone foils such that the interaction of homophony and foil condition was significant by subjects in the RT data, $F_s(1, 32) = 6.85$, $MSE = 7{,}932.97$; $F_i(1, 72) = 2.57$, $p = .11$, $MSE = 19{,}175.26$, but not in the error data ($F_s < 1.5$; $F_i < 1.5$). The three-way interaction of foil condition, frequency, and homophony was not significant in the RT data ($F_s < 1.5$; $F_i < 1.5$) or in the error data ($F_s < 1.5$; $F_i < 1.5$), suggesting a larger homophone effect for low-frequency words in both foil conditions.

RTs for words were again longer with pseudohomophone foils, $F_s(1, 32) = 5.40$, $MSE = 230{,}499.77$; $F_i(1, 72) = 14.83$, $MSE = 13{,}482.93$, but the effect of foil condition was not significant in the error data ($F_s < 1.5$; $F_i < 1.5$). There was a significant main effect of frequency in the RT data, $F_s(1, 32) = 67.27$, $MSE = 28{,}315.12$; $F_i(1, 72) = 85.82$, $MSE = 45{,}284.39$, and in the error data, $F_s(1, 32) = 57.47$, $MSE = 146.25$; $F_i(1, 72) = 33.08$, $MSE = 494.10$, but no interaction of frequency with foil condition in either the RT data, $F_s < 1.5$; $F_i(1, 72) = 1.67$, $p = .21$, $MSE = 13{,}482.93$, or the error data ($F_s < 1.5$; $F_i < 1.5$). RTs were slower when participants were presented with the pseudoword foil condition first, although this effect only approached significance by subjects, $F_s(1, 32) = 2.96$, $p = .09$, $MSE = 253{,}559.16$; $F_i(1, 72) = 70.87$, $MSE = 20{,}102.51$, and was not significant in the error data ($F_s < 1.5$; $F_i < 1.5$). The frequency effect was also larger when the pseudoword foil condition was presented first, and hence there was an interaction of frequency and order condition in the RT data, although the effect was not significant by subjects, $F_s(1, 32) = 2.27$, $p = .14$, $MSE = 28{,}315.12$; $F_i(1, 72) = 7.22$, $MSE = 20{,}102.51$, and a significant interaction in both analyses in the error data, $F_s(1, 32) = 4.28$, $MSE = 57.19$; $F_i(1, 72) = 5.52$, $MSE = 132.69$.

Foil Responses

Participants took significantly longer to reject pseudohomophone foils, $F_s(1, 38) = 22.15$, $MSE = 15{,}083.00$; $F_i(1, 78) = 42.33$, $MSE = 14{,}992.56$, and made more errors on pseudohomophone foils, $F_s(1, 38) = 21.52$, $MSE = 28.76$; $F_i(1, 78) = 7.39$, $MSE = 167.38$. Participants also took longer to reject foils when the pseudoword foil condition was presented first, $F_s(1, 38) = 4.50$, $MSE = 112{,}455.55$; $F_i(1, 78) = 222.19$, $MSE = 4{,}505.29$, but there was no such effect in the error data. The difference in RTs for the pseudoword and pseudohomophone foil conditions was larger when the pseudoword foil condition was presented first, and hence there was an interaction of foil condition and order condition in the RT data only, $F_s(1, 38) = 4.29$, $MSE = 15{,}083.00$; $F_i(1, 78) = 18.87$, $MSE = 4{,}505.29$.

Discussion

Once again, a significant effect of homophony was found for low-frequency words, indicating that phonology played a role

Table 4
Mean Lexical-Decision RTs (and Error Percentages) and Homophone Effects for Experiment 3

	Pseudoword foil condition			Pseudohomophone foil condition		
		Homophone effect			Homophone effect	
Stimulus type	RT	RT	Errors	RT	RT	Errors
	Pseudoword foil condition presented first					
High–low homophone	723 (1.00)			803 (1.50)		
High-frequency control	726 (0.50)	−3	0.50	782 (1.00)	21	0.50
Low–high homophone	948 (11.00)			1020 (15.00)		
Low-frequency control	870 (10.50)	78*	0.50	904 (11.50)	116*	3.50
Foil	924 (4.00)			1088 (8.50)		
	Pseudohomophone condition presented first					
High–low homophone	698 (5.00)			696 (2.00)		
High-frequency control	710 (2.50)	−12	2.50	673 (1.00)	23	1.00
Low–high homophone	836 (15.00)			886 (14.00)		
Low-frequency control	766 (11.00)	70*	4.00	768 (8.50)	118*	5.50*
Foil	819 (4.12)			887 (10.75)		

Note. Response time (RT) is given in milliseconds. High–low = high-frequency homophones with low-frequency homophone mates (e.g., *MADE–MAID*); Low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*).
* $p < .05$.

in recognition of these words. Again, the effect was not eliminated with pseudohomophone foils despite the fact that we used a within-subject manipulation of foil type, as in Davelaar et al. (1978). Although half of our participants received the pseudohomophone foil condition first, even those who received the pseudoword foil condition first, as did Davelaar et al.'s participants, produced a larger, not smaller, homophone effect with pseudohomophone foils for low-frequency words. Thus, we again find no evidence that participants can strategically deemphasize phonology in the lexical-decision task.

Experiment 4

In Experiment 4 we tested the possibility that participants can deemphasize phonology only when all words in the list are low in frequency. In Experiments 1 through 3, our stimulus lists included both high- and low-frequency words. In contrast, Davelaar et al.'s (1978) evidence for strategic deemphasis of phonology with low-frequency words came from a study in which word frequency was blocked. That is, there were no high-frequency words in the lists containing low-frequency words. Perhaps participants can only deemphasize phonology when many words in the list cannot be recognized quickly on the basis of their spellings. In Experiment 4 participants received only low-frequency words and were presented with (a) the list that included pseudoword foils first and (b) the list that included pseudohomophone foils second, because this closely replicates Davelaar et al.'s procedure.

Method

Participants

There were 28 participants in this study.

Stimuli

The stimuli for Experiment 4 were the same as those used in Experiment 3, except that the high-frequency homophones and high-frequency control words were not included. There were two experimental lists instead of four. One half of the low-frequency homophones and controls appeared in one list mixed with pseudowords and in the other mixed with pseudohomophones, and vice versa for the other half of the low-frequency homophones and controls.

Procedure

The procedure was the same as in Experiment 3 except that all participants were given a list with pseudoword foils first and pseudohomophone foils second, as was done in Davelaar et al. (1978).

Results

Response errors occurred on 5.2% of trials, and data from 1.6% of trials were removed because of extremely long or short RTs. Mean response times and response error rates for Experiment 4 are presented in Table 5.

Word Responses

Homophones had longer RTs than matched controls, $F_s(1, 26) = 13.58$, $MSE = 3428.53$; $F_i(1, 38) = 2.37$, $p = .14$, $MSE = 25{,}151.19$, and produced more errors, $F_s(1, 26) = 10.51$, $MSE = 14.35$; $F_i < 1.5$, although the homophone effects in both the RT and error data were not significant by items. The homophone effect was again larger, not smaller, with pseudohomophone foils such that the interaction of homophony and foil condition was marginally significant by subjects in the RT data, $F_s(1, 26) = 3.46$, $p = .07$, $MSE = 1{,}325.22$; $F_i < 1.5$, but not in the error data ($F_s < 1.5$; $F_i < 1.5$).

RTs for words were again longer with pseudohomophone foils, $F_s(1, 26) = 6.99$, $MSE = 11{,}575.49$; $F_i(1, 38) = 4.35$, $MSE = 4{,}437.65$, and the effect of foil condition was also significant by subjects in the error data, $F_s(1, 26) = 11.02$, $MSE = 32.42$; $F_i < 1.5$.

Foil Responses

Participants took significantly longer to reject pseudohomophone foils, $F_s(1, 27) = 5.02$, $MSE = 6{,}851.39$; $F_i(1, 78) = 9.33$, $MSE = 8{,}219.47$, and made more errors on pseudohomophone foils, $F_s(1, 27) = 15.19$, $MSE = 16.32$; $F_i(1, 78) = 4.42$, $MSE = 79.70$.

Discussion

Once again, we found no evidence that participants can strategically deemphasize phonology in the lexical-decision task, even when all words are low in frequency. Because our design in Experiment 4 was the same as in Davelaar et al. (1978), it seems most likely that our inability to replicate their findings is due to differences in the stimulus items used.

Experiment 5

In Experiments 1 through 4 we used our own set of word and foil stimuli, because Davelaar et al. (1978) included in their appendix only the homophones that they used, not the control words

Table 5
Mean Lexical-Decision RTs (and Error Percentages) and Homophone Effects for Experiment 4

	Pseudoword foil condition			Pseudohomophone foil condition		
		Homophone effect			Homophone effect	
Stimulus type	RT	RT	Errors	RT	RT	Errors
Low–high homophone	803 (4.29)			865 (8.93)		
Low-frequency control	775 (3.04)	28*	1.25	818 (5.54)	47*	3.39*
Foil	887 (2.86)			940 (7.06)		

Note. Response time (RT) is given in milliseconds. Low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*).
* $p < .05$.

or the foils. After four unsuccessful attempts to replicate their findings, we decided it was necessary to attempt a replication using their stimuli to try to resolve the discrepancy between their findings and ours. Derek Besner kindly provided us with the complete list of words; however the foils were no longer available. In Experiment 5 we examined whether the sets of homophones and control words that Davelaar et al. used in their two foil conditions could both produce a homophone effect with pseudoword foils. It is possible that the words that they used in their pseudoword foil condition would produce a homophone effect with any reasonable set of foils, whereas the words used in their pseudohomophone foil condition would not produce a homophone effect with either pseudohomophone or pseudoword foils, possibly because the word were reasonably high in frequency. Experiment 5, then, is a replication of Experiment 4 with Davelaar et al.'s (1978, Experiment 3) word stimuli.

Method

Participants

There were 32 participants in this study.

Stimuli

The critical word stimuli for Experiment 5 were the same as those used by Davelaar et al. (1978, Experiment 3). These included 60 homophones (all the lower frequency members of homophone pairs) and 60 control words. The foils and most of the fillers were the same as those used in Experiment 4. Two of the filler words used in Experiment 4 were replaced because they were among Davelaar et al.'s control words. There were two sets of word stimuli. Set A words were the homophone and control words that Davelaar et al. used in their *SLINT* (i.e., pseudoword) foil condition, and Set B were the homophone and control words used in their *GRONE* (i.e., pseudohomophone) foil condition. As in Experiment 4 there were two stimulus lists. The first list was exactly like Davelaar et al.'s: Set A words appeared with pseudoword foils in the first part of the list, and then after the fillers, Set B words appeared with pseudohomophone foils. The second list had the Set B words mixed with pseudoword foils in the first part, and then after the fillers, the Set A words appeared with the pseudohomophone foils. The fillers were arranged as in Experiment 4.

Procedure

The procedure was the same as in Experiment 4—the pseudoword foil condition was always presented before the pseudohomophone foil condition, as it was in Davelaar et al. (1978).

Results

Response errors occurred on 5.6% of trials, and data from 1.3% of trials were removed due to extremely long or short RTs. Mean RTs and response error percentages for Experiment 5 are presented in Table 6.

Word Responses

Overall, homophones had significantly longer RTs than control words, $F_s(1, 30) = 3.87$, $MSE = 5{,}137.55$; $F_i < 1.5$, and produced more errors, $F_s(1, 30) = 87.81$, $MSE = 7.85$; $F_i(1, 57) = 4.75$, $MSE = 292.34$, although the homophone effect in the RT data was not significant by items. This homophone effect was, however, qualified by a significant interaction with word set in the RT data, $F_s(1, 30) = 6.43$, $MSE = 2{,}556.97$; $F_i < 1.5$, such that the homophone effect was only present for Set A stimuli regardless of the nature of the foils. The three-way interaction of homophony, word set, and foil condition was not significant in the RT data ($F_s < 1.5$; $F_i < 1.5$), but was significant in the error data, $F_s(1, 30) = 9.15$, $MSE = 12.92$; $F_i(1, 57) = 5.41$, $MSE = 34.49$. The nature of this interaction was that the homophone effect in the error data increased more in the pseudohomophone condition for Set B words. In the error data only, the homophone effect was larger with pseudohomophone foils and hence there was a marginally significant interaction of homophony with foil condition by

Table 6
Mean Lexical-Decision RTs (and Error Percentages) and Homophone Effects for Experiment 5 (Davelaar et al., 1978, Stimuli)

	Pseudoword foil condition			Pseudohomophone foil condition		
		Homophone effect			Homophone effect	
Stimulus type	RT	RT	Errors	RT	RT	Errors
			Set A[a]			
Low–high homophone	896 (6.64)			932 (12.00)		
Low-frequency control	847 (3.10)	49*	3.54*	945 (4.29)	−13	7.71*
Foil	1023 (3.36)			1168 (9.86)		
			Set B[b]			
Low–high homophone	847 (5.00)			932 (8.16)		
Low-frequency control	839 (3.04)	8	1.96	895 (2.71)	37*	5.45*
Foil	1056 (4.70)			1168 (9.78)		

Note. Response time (RT) is given in milliseconds. Low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*).
[a] This list presented with pseudowords (same presentation as in Davelaar et al.). [b] This list presented with pseudowords (opposite to presentation in Davelaar et al.).
* $p < .05$.

subjects, $F_s(1, 30) = 3.87, p = .06, MSE = 7.85$; $F_i(1, 57) = 1.54$, $p = .22$, $MSE = 34.49$.

RTs for words were again longer with pseudohomophone foils, $F_s(1, 30) = 12.49$, $MSE = 11,867.42$; $F_i(1, 57) = 26.84$, $MSE = 11,823.06$, and the effect of foil condition was also significant in the error data, $F_s(1, 30) = 16.92$, $MSE = 10.68$; $F_i(1, 57) = 7.66$, $MSE = 39.71$.

Foil Responses

Participants took significantly longer to reject pseudohomophone foils than pseudoword foils, $F_s(1, 30) = 26.44$, $MSE = 9,700.61$; $F_i(1, 78) = 35.34$, $MSE = 23,347.48$, and made more errors on pseudohomophone foils, $F_s(1, 30) = 10.09$, $MSE = 44.77$; $F_i(1, 78) = 7.79$, $MSE = 215.91$.

Discussion

The results of Experiment 5 show that a homophone effect was generated only from Davelaar et al.'s (1978) Set A words (the set that showed a homophone effect in the original article). A homophone effect was not found with Set B words in the RT data in either the pseudoword or the pseudohomophone foil conditions. One surprising aspect of these data was the fact that, even for Set A words, the homophone effect in the RT data was not larger in the pseudohomophone foil condition. This is in contrast with the pattern observed in each of our previous experiments. This difference may be due to the nature of the homophones and control words in Set A. Among the nonhomophonic control words for Set A, there were two repeated words, one high-frequency word misclassified as a low-frequency word, and, perhaps most important, three homophones. The fact that there are homophones among the control words raises the possibility that the homophone effect observed for Set A is not purely a homophone effect. The observed effect could, instead, just be a function of poor matching between homophones and control words, and if so, there is no obvious reason why such an effect would be larger with pseudohomophone foils.

Experiment 6

The results of Experiments 1 through 5 showed that homophone effects do exist in lexical-decision tasks, that the effects are larger with more orthographically word-like pseudowords, and that the phonological processing that produces homophone effects cannot be strategically deemphasized in the context of pseudohomophones. We next turn our attention to determining whether all types of homophones show the same pattern as those investigated in the previous experiments. Homophone effects were found primarily for low-frequency words that were homophonic with high-frequency words. However, Rubenstein et al.'s (1971) account also predicts that all homophones should produce a homophone effect as long as they have a higher frequency mate. No study that we are aware of has examined such homophones. In Experiment 6 we investigated whether homophone effects arise for low–low and high–higher homophones. As such, this experiment provides a more complete evaluation of homophone effects than previous studies.

Method

Participants

There were 64 participants in this study. However, 1 participant in the pseudohomophone foil condition had extraordinarily long RTs ($M = 2341$ ms). That participant was excluded, and 1 participant in the pseudohomophone foil condition who was given the other list was randomly selected to be excluded from the analyses.

Stimuli

The words in the study were 60 homophones and 60 matched nonhomophone control words (see Appendix). There were 20 low-frequency homophones with high-frequency mates (low–high homophones), which were a subset of the low-frequency homophones used in Experiments 1 through 3, 20 high-frequency homophones with higher frequency mates (high–higher homophones), and 20 low-frequency homophones with low frequency mates (low–low homophones). The high–higher homophones were selected such that the target word had a frequency of more than 32 per million and that the target word's homophonic mate, which was not presented, had a higher frequency. The low–low homophones were selected such that both the target word and its homophonic mate had frequencies of less than 32 per million. Homophones and nonhomophones were matched as closely as possible for frequency, length, first letter, and neighborhood size using Coltheart's N (M. Coltheart et al., 1977). Mean frequencies and neighborhood sizes for each word type are presented in Table 2.

The foil stimuli were the word-like pseudoword and pseudohomophone foils used in Experiment 1. To equate the number of words and foils on the list, the word stimuli were divided into two lists, so that each participant was presented with 60 word stimuli and the 60 foil stimuli. In each word list there were 10 of each of these types of word stimuli. Each list was presented equally often with pseudoword and pseudohomophone foils.

Procedure

The procedure was the same as in Experiment 1.

Results

Response errors occurred on 6.2% of trials, and data from 3.1% of trials were removed because of extremely long or short RTs. Mean RTs and response error percentages for Experiment 6 are presented in Table 7.

Word Responses

The RTs and error rates for each type of homophones were analyzed separately. The high–higher homophones had significantly longer RTs than matched control words, $F_s(1, 58) = 4.42$, $MSE = 6,671.64$; $F_i(1, 18) = 4.47$, $MSE = 13,743.65$, but did not produce more errors ($F_s < 1.5$; $F_i < 1.5$). This main effect, however, was qualified by a significant interaction of foil condition and homophony in the RT data, $F_s(1, 58) = 3.73$, $p = .07$, $MSE = 6,671.64$; $F_i(1, 18) = 3.07$, $p = .09$, $MSE = 8,611.64$, but not in the error data, $F_s < 1.5$; $F_i(1, 18) = 3.92$, $p = .06$, $MSE = 12.46$. The nature of this interaction was that there was a homophone effect only in the pseudohomophone foil condition. RTs for these words were also significantly longer with pseudohomophone foils, $F_s(1, 58) = 11.20$, $MSE = 75,954.45$; $F_i(1, 18) = 69.12$, $MSE = 8,924.79$, but there were not significantly more errors ($F_s < 1.5$; $F_i < 1.5$).

Table 7
Mean Lexical-Decision RTs (and Error Percentages) and Homophone Effects for Experiment 6

	Pseudoword foil condition			Pseudohomophone foil condition		
		Homophone effect			Homophone effect	
Stimulus type	RT	RT	Errors	RT	RT	Errors
High–higher homophone	621 (5.94)			821 (6.00)		
High-frequency control	618 (3.44)	3	2.50	757 (5.33)	64*	0.67
Low–high homophone	702 (7.81)			918 (10.00)		
Low-frequency control	665 (6.88)	37*	0.93	848 (9.66)	70*	0.44
Low–low homophone	685 (5.63)			863 (14.33)		
Low-frequency control	687 (3.13)	−2	2.50	844 (5.67)	19	8.66*
Foil	887 (2.86)			940 (7.06)		

Note. Response time (RT) is given in milliseconds. High–higher = high-frequency homophones with higher frequency homophone mates (e.g., *FOUR–FOR*); Low–high = low-frequency homophones with high-frequency homophone mates (e.g., *MAID–MADE*); Low–low = low-frequency homophones with low-frequency homophone mates (e.g., *BAIL–BALE*).
* $p < .05$.

The low–high homophones had significantly longer RTs than matched control words, $F_s(1, 58) = 9.67$, $MSE = 8{,}699.33$; $F_i(1, 18) = 3.87$, $p = .07$, $MSE = 16{,}473.91$, but the homophones did not produce significantly more errors, $F_s < 1.5$; $F_i < 1.5$. The homophone effect was larger with pseudohomophone foils, but the interaction of homophony and foil condition was not significant in the RT data, $F_s < 1.5$; $F_i < 1.5$, or in the error data, $F_s < 1.5$; $F_i < 1.5$. Words had longer RTs with pseudohomophone foils, $F_s(1, 58) = 15.20$, $MSE = 82{,}286.90$; $F_i(1, 18) = 103.96$, $MSE = 8{,}223.64$, but there was no parallel effect in the error rates, $F_s < 1.5$; $F_i(1, 18) = 2.98$, $p = .10$, $MSE = 41.62$.

The low–low homophones did not have significantly longer RTs than matched control words, $F_s < 1.5$; $F_i(1, 18) = 2.64$, $p = .12$, $MSE = 14{,}643.32$, but did produce significantly more errors, $F_s(1, 58) = 14.21$, $MSE = 67.96$; $F_i(1, 18) = 4.75$, $MSE = 131.38$. The homophone effect was slightly larger with pseudohomophone foils; the interaction of homophony and foil condition was not significant in the RT data ($F_s < 1.5$; $F_i < 1.5$) but was significant in the error data, $F_s(1, 58) = 4.33$, $MSE = 67.96$; $F_i(1, 18) = 3.40$, $p = .07$, $MSE = 55.92$. Words had longer RTs with pseudohomophone foils, $F_s(1, 58) = 9.33$, $MSE = 104{,}450.24$; $F_i(1, 18) = 32.84$, $MSE = 17{,}785.15$, and error rates were higher with pseudohomophones, $F_s(1, 58) = 10.70$, $MSE = 91.55$; $F_i(1, 18) = 13.54$, $MSE = 46.75$.

Foil Responses

Participants took significantly longer to reject pseudohomophone foils, $F_s(1, 58) = 10.87$, $MSE = 60{,}685.16$; $F_i(1, 118) = 124.94$, $MSE = 21{,}613.37$, and they made more errors on pseudohomophone foils, $F_s(1, 58) = 3.37$, $p = .07$, $MSE = 32.95$; $F_i(1, 118) = 3.40$, $p = .07$, $MSE = 126.40$.

Discussion

The low–high homophones produced a significant effect of homophony in the pseudoword foil condition, just as they did in the previous experiments with word-like foils. The high–higher homophones, however, did not produce a significant effect of homophony in the pseudoword foil condition. This finding is contrary to the predictions of Rubenstein et al. (1971) because that account predicts homophone effects for homophones of any frequency as long as the homophone mate is higher in frequency. The absence of a homophone effect for high–higher homophones is analogous to the absence of an effect for *high–low* homophones in most conditions in Experiments 1 and 2, suggesting that in those conditions, lexical decisions for high-frequency homophones are unaffected by their homophonic mate regardless of its frequency. The low–low homophones also did not produce a homophone effect. Thus it appears that the only type of homophone that produces an effect of homophony in a lexical-decision task with word-like pseudoword foils is low–high homophones.

Once again, there was no evidence that participants strategically deemphasized phonology in the presence of pseudohomophones because the homophone effect was larger, not smaller, with pseudohomophone foils for each of the three types of homophones. As was found with high–low homophones in Experiment 1, a significant homophone effect appeared for high–higher homophones with pseudohomophone foils that was not present with pseudoword foils. A significant homophone effect also appeared, although only in the error data, for the low–low homophone pairs with pseudohomophone foils, an effect that was not present with pseudoword foils.

GENERAL DISCUSSION

These experiments revealed three factors that influence whether a homophone effect will be observed in the lexical-decision task. One factor is the frequency of the homophone and of its mate. Homophone effects are most likely to be observed for low-frequency words with high-frequency homophone mates. No other type of homophone produced a significant effect of homophony with pseudoword foils. A second factor is the orthographic nature of the foils. Homophone effects are more likely to be observed with foils that are orthographically similar to English words than with foils that do not resemble English words. A third factor is the phonological nature of the foils. Homophone effects in these studies were larger, not smaller, with foils that sounded like real words than with foils that did not. Thus, these experiments produced no evidence that participants' strategic changes in emphasis on phonology are responsible for the sporadic nature of homophone effects in previous research. We first discuss whether these three factors can account for the conflicting findings in the lexical-decision literature on homophone effects, and then we discuss what these findings tell us about the processes involved in word recognition.

Relation to Past Literature

Our results using pseudoword foils nicely replicate Rubenstein et al.'s (1971) results. Those authors observed an effect of homophony for the lower frequency members of homophone pairs but not for the higher frequency members. Their effect may have been weaker than ours in that it was not significant in Clark's (1973) analysis in which both subjects and items were treated as random factors. This may have been due to the orthographic characteristics of Rubenstein et al.'s foils. Not all foils were provided in the article, but of the 56 that were, 19 used word bodies that do not occur in any real English words (e.g., *SLIC*). Thus, at least one third of Rubenstein et al.'s foils were not very word-like. As demonstrated in the present Experiment 2, homophone effects are smaller when foils are not word-like.

There may have been several reasons why Coltheart et al. (1977) did not observe a homophone effect. Although all of their words were the lower frequency member of the homophone pair, 6 of their 39 words had a frequency greater than 40 per million. In Experiment 6, we found an effect of homophony for high–higher words only when all foils were pseudohomophones. In addition, 7 of their words had a homophone mate with a frequency less than 32. In Experiment 6, low–low words only showed a homophone effect in the error data and only when all the foils were pseudohomophones.

The second possible reason that Coltheart et al. (1977) did not observe a homophone effect is that the word–nonword discriminations in their experiment were probably somewhat easier than in the present experiments. For example, 27 of their 78 foils had a word body that does not occur in English, and as we have shown, homophone effects are smaller when foils are less word-like. In addition, there was considerable repetition of letter patterns in their stimuli. That is, half of their foils were pseudowords and half were pseudohomophones and these were matched so that a particular pseudohomophone (e.g., *ILE*) only differed from a particular pseudoword (e.g., *IFE*) by one letter. Because the pseudohomophones and pseudowords were presented in the same block of trials, Coltheart et al.'s participants saw the same body among the foils as many as six times. Both of these aspects of Coltheart et al.'s experiment probably made word–nonword discriminations relatively easy which may explain why they failed to obtain a significant homophone effect, even with pseudohomophones as foils. The same comments apply to Dennis et al.'s (1985, Experiment 4) replication with the same stimuli in which there was no effect of homophony. Note also that Dennis et al. (Experiment 3) did find a homophone effect with a subset of Coltheart et al.'s homophone and control pairs and a new set of foils, half of which were pseudohomophones. Thus, at least some of Coltheart et al.'s word pairs produce a homophone effect in some foil environments.

With respect to Davelaar et al.'s (1978) study, we did replicate their finding of a homophone effect for lower frequency homophones that were presented with pseudoword foils, although only when the pseudowords were very word-like. Where our results differed, however, is that our homophone effect typically was larger with pseudohomophone foils, whereas theirs was reported to have disappeared. The present Experiment 3 showed that this difference was not a consequence of their presenting the pseudohomophone foil condition to all participants after the pseudoword foil condition, because our homophone effect was larger with pseudohomophone foils when we used a within-subjects manipulation regardless of which foil condition was presented first. Experiment 4 demonstrated that the difference between our results and Davelaar et al.'s was not due to their inclusion of only low-frequency words. The reason for the discrepancy was revealed in Experiment 5, where we demonstrated that the word set Davelaar et al. used in their pseudohomophone foil condition did not produce a homophone effect even with pseudoword foils. Thus, there was no "disappearing" homophone effect. This word set may not have produced a homophone effect because, as we have noted above, there were more high-frequency homophones in the set used in the pseudohomophone foil condition, and such homophones are less likely to produce a homophone effect.

With respect to our finding that readers do not have strategic control of their use of phonology, our results are consistent with word-decision data in most other lexical-decision studies except Pugh et al. (1994, Experiment 3). We consider that study to be weak evidence for strategic control of phonology, however, because those results do not seem to be easily replicable (L. Katz, personal communication, April 28, 1999; Pexman et al., 1996).

Implications for Word Recognition

An effect of homophony was reliably found for low-frequency homophones with high-frequency mates when they appeared among word-like pseudoword foils. This provides evidence that phonology plays a role in the recognition of low-frequency words. No evidence was obtained indicating that phonology played a role in the recognition of high-frequency words when foils were pseudowords.

Why does the activation of phonology produce slowed recognition of low-frequency homophones? One possible explanation is that the phonological representation of a homophone activates two meanings, and these semantic representations compete with one another and slow recognition. This is unlikely to be the cause of homophone effects in lexical decision, however, because research involving polysemous words (i.e., words with multiple meanings, e.g., *BANK*) has shown faster lexical decision RTs for polysemous words than for control words (Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas, Ferraro, & Simpson, 1988; Pexman & Lupker, 1999; Rubenstein, Garfield, & Millikan, 1970). Thus, activation of more than one semantic representation seems to actually speed lexical-decision responses. As such, it seems unlikely that the delay produced in responding to homophones could be accounted for in terms of semantic competition.

A second, more likely explanation of the homophone effect is that through feedback, the phonological representation of a homophone activates two orthographic representations. As a result, competition is created at the orthographic level and it will take longer for the system to settle on the pattern of orthographic activation appropriate to the presented word (Stone, Vanhoy, &

Van Orden, 1997; Ziegler, Montant, & Jacobs, 1997). Under the assumption that lexical decisions are primarily based on activation within the orthographic level, this extra settling time leads directly to longer latencies in lexical-decision tasks. We envision that this feedback mechanism could easily be incorporated in a fully-interactive, parallel distributed processing (PDP)–type of framework (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). In fact, feedback connections from phonological units to orthographic units through hidden units were included in Seidenberg and McClelland's (1989) theoretical framework, although these were not implemented in their simulation. Feedback connections are also present in the dual-route cascaded model (M. Coltheart, Curtis, Atkins, & Haller, 1993). However, in that model, homophone effects in lexical decision would be accounted for by feedback from the phonological output lexicon to the visual word detectors (or visual lexicon).

Initial evidence of feedback effects were presented by Stone et al. (1997), who claimed to have demonstrated that words with phonological bodies that can be spelled more than one way (e.g., "eep" can be spelled *EAP* in *HEAP* or *EEP* in *SHEEP;* Stone et al. referred to these as "feedback inconsistent" words) produce slower correct *YES* responses than words with bodies that can be spelled only one way (e.g., *OBE* in *PROBE;* Stone et al. referred to these as "feedback consistent" words). This claim, however, has been challenged by Peereman, Content, and Bonin (1998), who were unable to replicate Stone et al.'s findings when feedback-inconsistent and feedback-consistent words were matched for familiarity.

Even if Peereman et al. (1998) were correct in arguing that there is little impact of phonological–orthographic feedback when considering sublexical units like word bodies, the situation may be somewhat different when that feedback activates lexical units (or collections of sublexical units that together form the orthographic code for a word). These units (or patterns of units) may have much more ability to compete once activated because they are such familiar, unitary patterns. One implication is that this ability to compete would be a direct function of the familiarity or frequency of the word. Thus, the largest homophone effect would be expected for low-frequency words with high-frequency mates because the orthographic patterns for low-frequency words would initially be less activated than those for high-frequency words, and the feedback from phonology for a high-frequency mate would produce relatively high levels of activation in the high-frequency word's orthographic representation. Presentation of a low-frequency homophone with a high-frequency mate would, therefore, result in the greatest competition at the orthographic level.

The impact of pseudohomophone list context can also be explained within this framework. What pseudohomophone foils do is essentially make the lexical-decision process more difficult than with pseudoword foils. Specifically, unlike pseudowords, pseudohomophones activate the phonological representation for a word. This activation then feeds back to the orthographic level, activating the orthographic representation corresponding to that word. Thus, to avoid responding positively when a pseudohomophone is presented in a lexical-decision task, participants have to set a more strict activation criterion for discriminating between words and nonwords. The result is longer RTs for both positive and negative responses. A secondary result of the more strict criterion setting is that there is more time for feedback activation to accumulate and create competition at the orthographic level. The result is larger homophone effects for the standard low–high homophones and emerging homophone effects for high-frequency homophones and for low-frequency homophones with low-frequency homophone mates whenever pseudohomophones are used.

Summary

These experiments have shown that homophone effects do occur for low-frequency words in the lexical-decision task, indicating that phonology plays a role in the recognition of these words. This finding provides a nice parallel to the results from a number of other tasks, tasks that required the activation of meaning. The size of the homophone effect was shown to be influenced by the frequency of a homophone and of its mate, and by the orthographic and phonological characteristics of the foils. However, no evidence was obtained that homophone effects could be diminished by encouraging participants to strategically reduce their reliance on phonology, suggesting that readers are unable to do so. In general, existing data from the lexical-decision literature now overwhelmingly support the view that readers have little, if any, strategic control over the activation of phonology by a visually presented word.

References

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

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review, 4,* 439–461.

Berent, I. (1997). Phonological priming in the lexical decision task: Regularity effects are not necessary evidence for assembly. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1727–1742.

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.

Cohen, J. (1976). Discussion of Wike and Church's comments. *Journal of Verbal Learning and Verbal Behavior, 15,* 257–266.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589–608.

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

Coltheart, V., Avons, S. E., Masterson, J., & Laxon, V. J. (1991). The role of assembled phonology in reading comprehension. *Memory & Cognition, 19,* 387–400.

Daneman, M., & Reingold, E. M. (1993). What eye fixations tell us about phonological recoding during reading. *Canadian Journal of Experimental Psychology, 47,* 153–178.

Daneman, M., Reingold, E. M., & Davidson, M. (1995). Time course of phonological activation during reading: Evidence from eye fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 884–898.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 618–632.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. T. (1978). Phonological recoding and lexical access. *Memory & Cognition, 8,* 195–209.

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. Waller, & E. M. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 167–197). San Diego, CA: Academic Press.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45A,* 353–372.

Ferrand, L., & Grainger, J. (1996). List context effects on masked phonological priming in the lexical decision task. *Psychonomic Bulletin and Review, 3,* 515–519.

Gibbs, P., & Van Orden, G. C. (1998). Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1162–1187.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1331–1356.

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., Levy, B. A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General, 128,* 219–264.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358–394.

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.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 601–609.

Keppel, G. (1976). Discussion of Wike and Church's comments. *Journal of Verbal Learning and Behavior, 15,* 257–266.

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

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 951–966.

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

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 14–24.

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

McQuade, D. V. (1983). Phonological recoding becomes automatic with stimulus repetition. *Language and Speech, 26,* 393–405.

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

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback inconsistency in visual word recognition. *Journal of Memory and Language, 39,* 151–174.

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects? *Canadian Journal of Experimental Psychology, 53,* 323–334.

Pexman, P. M., Lupker, S. J., Jared, D., Toplak, M., & Rouibah, A. (1996, November). *Strategically controlling phonology in response to pseudohomophones?* Paper presented at the 37th Annual Meeting of the Psychonomic Society, Chicago.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Pring, L. (1981). Phonological codes and functional spelling units: Reality and implications. *Perception & Psychophysics, 30,* 573–578.

Pugh, K. R., Rexer, K., & Katz, L. (1994). Evidence of flexible coding in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 807–825.

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.

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.

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

Smith, J. E. (1976). Discussion of Wike and Church's comments. *Journal of Verbal Learning and Verbal Behavior, 15,* 257–266.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337–359.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 744–774.

Taft, M. (1982). An alternative to grapheme–phoneme conversion rules? *Memory & Cognition, 10,* 465–474.

Underwood, G., Roberts, M., & Thomason, H. (1988). Strategical invariance in lexical access: The reappearance of the pseudohomophone effect. *Canadian Journal of Psychology, 42,* 24–34.

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

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488–522.

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

Wike, E. L., & Church, J. D. (1976). Comments on Clark's "The language-as-fixed-effects fallacy." *Journal of Verbal Learning and Verbal Behavior, 15,* 249–255.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37,* 533–554.

Appendix

Stimuli Used in Experiments

Experiments 1–4 Word Stimuli

High-frequency homophone	Control word	Low-frequency homophone	Control word	High-frequency homophone	Control word	Low-frequency homophone	Control word
aid	ask	aide	arch	mind	miss	mined	mused
break	broad	brake	bleed	none	note	nun	nap
beat	boat	beet	bait	pain	post	pane	pest
blue	bill	blew	boil	pair	pick	pear	perk
board	black	bored	baked	pale	page	pail	peep
chance	choice	chants	champs	read	rest	reed	rail
course	church	coarse	cheese	real	rate	reel	rude
dear	draw	deer	deed	rose	rock	rows	rags
die	dog	dye	den	sale	soil	sail	silk
fair	fast	fare	fake	seem	soon	seam	seep
feet	five	feat	flip	size	step	sighs	skids
grown	grass	groan	graze	soul	safe	sole	seal
hair	hard	hare	hack	thrown	twelve	throne	thrill
hall	hope	haul	hoop	wine	wide	whine	whack
higher	having	hire	heap	allowed	applied	aloud	aloof
horse	heart	hoarse	hearse	council	concern	counsel	costume
least	large	leased	loomed	latter	larger	ladder	locate
loan	lake	lone	lace	manner	middle	manor	manic
made	must	maid	mess	minor	model	miner	molar
main	mass	mane	maze	morning	million	mourning	mounting

Experiment 6 Word Stimuli

High-frequency homophone	High-frequency mate (not presented)	Control word	Low-frequency homophone	High-frequency mate (not presented)	Control word
four	for	felt	blew	blue	boil
hear	here	hell	bored	board	baked
male	mail	mile	brake	break	bleed
meat	meet	meal	coarse	course	cheese
piece	peace	price	deer	dear	deed
plain	plane	prime	feat	feet	flip
rode	road	rare	hare	hair	hack
roll	role	risk	haul	hall	hoop
scene	seen	spent	hire	higher	heap
sea	see	sir	ladder	latter	locate
site	sight	shop	leased	least	loomed
sum	some	sin	maid	made	mess
sun	son	sat	mane	main	maze
threw	through	throw	mined	mind	mused
vary	very	vital	mourning	morning	mounting
wore	war	wage	reed	read	rail
weak	week	wise	reel	real	rude
wear	where	wash	seam	seem	seep
weather	whether	winter	sighs	size	skids
wood	would	worse	sole	soul	seal

Low-frequency homophone	Low-frequency mate (not presented)	Control word	Low-frequency homophone	Low-frequency mate (not presented)	Control word
bail	bale	bang	heel	heal	hint
bridal	bridle	beware	peal	peel	perk
chord	cord	chore	peer	pier	peck
creak	creek	croak	pray	prey	pint
dough	doe	ditch	stair	stare	stump
fairy	ferry	freak	stake	steak	stall
fur	fir	fix	tail	tale	tent
flair	flare	frail	tee	tea	tar
foul	fowl	fuss	warn	worn	weep
hay	hey	ham	yoke	yolk	yawn

(Appendix continues)

Experiments 1, 3, 4, 5, and 6 Foil Stimuli

Pseudoword	Pseudohomophone	Pseudoword	Pseudohomophone	Pseudoword	Pseudohomophone	Pseudoword	Pseudohomophone
nace	chace	treal	wheal	cleep	leep	froar	scoar
cade	rade	feap	sleap	zeer	heer	foast	koast
naff	laff	meap	keap	geet	heet	loast	goast
spail	scail	ched	ded	reet	neet	shong	rong
blain	sain	ped	hed	sern	lern	klor	bor
gair	shair	gree	plee	ferse	kerse	prore	rore
jair	squair	weech	teech	fet	swet	bort	cort
nait	hait	deef	greef	shet	thret	broze	noze
swait	mait	meef	cheef	pife	nife	shung	yung
lale	jale	jeek	bleek	thipe	tipe	turge	murge
drale	trale	fleek	speek	jite	tite	turl	gurl
clane	brane	geel	deel	noak	joak	blurse	vurse
tane	rane	greel	meel	loak	smoak	murt	shurt
brate	wate	neem	creem	broal	boal	furt	durt
brax	trax	feen	cleen	troar	floar	rutch	tutch

Experiment 2 Foil Stimuli

Pseudoword	Pseudohomophone	Pseudoword	Pseudohomophone	Pseudoword	Pseudohomophone	Pseudoword	Pseudohomophone
slahr	stahr	varck	darck	derl	gerl	trownd	fownd
mahrn	bahrn	darr	pharr	creth	breth	howt	dowt
blaie	plaie	dawlt	fawlt	troab	gloab	boyn	coyn
taige	raige	dawp	stawp	woald	soald	roynt	joynt
haik	caik	vawx	bawx	broald	goald	thufe	prufe
paik	waik	yawx	fawx	goarn	hoarn	lufe	rufe
yayk	tayk	blayl	mayl	foarn	boarn	cunn	runn
vayk	rayk	layvy	nayvy	choart	spoart	sunth	munth
baip	taip	thayze	fayze	hoart	shoart	murld	wurld
draive	braive	gedd	hedd	soize	noize	bruve	pruve
claive	graive	chedd	redd	tolph	golph	huve	muve
taize	daize	bleez	pleez	tooce	jooce	byne	myne
kaks	taks	deez	peez	fooce	looce	hyne	nyne
balce	falce	lerce	perce	loond	woond	byve	dyve
nande	sande	berce	nerce	boope	soope	chyze	pryze

Received October 29, 1998
Revision received May 24, 2000
Accepted July 14, 2000 ■