A ROWS is a ROSE:
Spelling, sound, and reading

GUY C. VAN ORDEN
University of California at San Diego, La Jolla, California

Skilled readers generally are assumed to make little or no use of words' phonological features in visual word identification. Contrary to this assumption, college students' performance in the present reading experiments showed large effects of stimulus word phonology. In Experiments 1 and 2, these subjects produced larger false positive error rates in a semantic categorization task when they responded to stimulus foils that were homophonic to category exemplars (e.g., ROWS for the category A FLOWER) than when they responded to spelling control foils. Additionally, in Experiment 2, this homophony effect was found under brief-exposure pattern-masking conditions, a result consistent with the possibility that phonology is an early source of constraint in word identification. Subjects did, however, correctly reject most homophone foils in Experiments 1 and 2. Experiment 3 investigated the source of this ability. The results of Experiment 3 suggest that subjects detected homophone impostors, such as ROWS, by verifying target foil spellings against their knowledge of the correct spellings of category exemplars, such as ROSE.

The possibility that the sound (or pronunciation) of words influences the identification of printed words has been repeatedly both affirmed and denied throughout the history of the study of the psychology of reading. This possibility, usually termed "phonological mediation," assumes that there are phonological sources of activation that precede word identification. Hence, word identification in reading is thought to be analogous to a process that transforms spelling to sound and then maps sound to meaning (e.g., see Gough, 1972; Rubenstein, Lewis, & Rubenstein, 1971; Spoehr & Smith, 1973). An alternative "direct-access" view makes little, if any, use of phonology in word identification. Rather, the orthographic representation of a printed word activates its lexical entry directly (e.g., see Aaronson & Ferres, 1983; Baron, 1973; Becker, 1976, 1980; Bower, 1970; Goodman, 1969; Kleiman, 1975; Kolers, 1970; Paap, Newsome, McDonald, & Schvaneveldt, 1982; Smith, 1971).

For "dual-access" theories that include both phonological mediation and direct access (e.g., Coltheart, 1978; Coltheart, Davelaar, Jonasson, & Besner, 1977; McClelland & Rumelhart, 1981; Meyer & Roddy, 1973; Norris & Brown, 1985; Rumelhart & McClelland, 1982; Seiden-

berg, 1985; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Waters, Seidenberg, & Bruck, 1984), the exact nature of the collaboration between these two sources of activation becomes a critical issue. All dual-access theories assume both orthographic and phonological sources of lexical activation analogous to direct access and phonological mediation. These theories differ, however, in the presumed relative time course of the two sources. Some theories assume that phonological activation arises more slowly than orthographic activation and that, consequently, the influence of this phonological activation is generally delayed and/or usually absent in word identification (Allport, 1977; Coltheart, 1978; McCusker, Hillinger, & Bias, 1981; Seidenberg, 1985; Seidenberg et al., 1984; Waters et al., 1984). In other theories, lexical nodes are activated by both sources simultaneously, and both will influence every instance of word identification (McClelland & Rumelhart, 1981; Perfetti, 1985; Van Orden, 1984).

A third issue concerns whether word identification includes some verification process (as is assumed in the theories of Becker, 1976, 1980; Grossberg & Stone, 1986; Paap et al., 1982; Rubenstein et al., 1971; and Schvaneveldt & McDonald, 1981) that is essentially a spelling check wherein the spelling representation associated with a candidate lexical entry is verified against the spelling representation of the word being read. Verification theories complicate the discussion of word identification by extending the usual meaning of the term "lexical access" and by amending the previously presumed all-inclusive categories of pre- and postlexical processes to include "intralexical" processes.

In its more usual, specific sense, lexical access refers to the action of prelexical processes that retrieve information contained at the lexical entry for use by postlexical processes downstream from word identification (e.g., processes of test integration, pronunciation, categoriza-

tion, etc.). In this specific sense, lexical access is the end to which word identification proceeds (e.g., see the discussions of lexical access in Becker, 1976; Carpenter & Just, 1981; Coltheart, 1978; Crowder, 1982; Van Orden, 1984) and sometimes is equated explicitly with word identification (e.g., see Danks & Hill, 1981).

In contrast, lexical access in the most general sense occurs whenever lexical information is retrieved from long-term memory, even for use *within* the process of word identification. For example, in verification models, information from memory (i.e., a representation of a word's spelling) is retrieved for use in the verification procedure, a subprocess of word identification. Hence, in verification models, lexical access in the general sense can occur repeatedly before word identification runs to completion. This important difference between current verification and nonverification theories motivates both the conception of intralexical processes (processes such as verification that use information stored in the lexicon toward the eventual selection of a lexical entry) and the respective general and specific meanings of the term lexical access.

The results of the experiments that follow are pertinent to each of these three issues: (1) Does phonological mediation occur in skilled reading? (2) What is the relative time course of orthographic and phonological influences upon word identification? (3) Does word identification include a verification process? I will begin with a brief review of the evidence that motivates the phonological mediation hypothesis and introduce the other issues in the course of later discussions.

The best current evidence for the phonological mediation hypothesis comes from experiments using the lexical decision task, in which subjects judge whether a letter string is a word (the reference language is usually English). Subjects in this task take more time to correctly reject pseudohomophone foils, such as DYME, than to reject control foils (Coltheart et al., 1977; Rubenstein et al., 1971). One explanation of this pseudohomophone effect is that DYME activates the phonological representation /DIM/, which in turn activates the lexical entry for the word DIME. Activation of the lexical entry for DIME makes it more difficult for subjects to classify DYME as a nonword. In less specific process terms, some representation of the phonological code for DYME increases its similarity to DIME and thus makes DYME harder to classify as a nonword.

There are at least two problems with this interpretation. Coltheart et al. (1977) questioned whether this effect bears upon how real words are identified (see also Henderson, 1982; McCusker et al., 1981). They note that the pseudohomophone effect is observed on *no* trials, which are generally slower than *yes* trials in lexical decision. Thus, the nonword DYME's effect in the lexical decision task may arise after the time has elapsed that is usually required for word identification in normal reading. Consequently, the pseudohomophone effect may not bear on the issue of whether phonological mediation occurs in normal reading.

Another problem is that perceived stimulus familiarity, evaluated independently of word identification, may interfere with (i.e., delay) a lexicality judgment (cf. Besner, 1984). By this account, a pseudohomophone is harder to classify as a nonword because its familiar phonological representation increases its perceived familiarity, without actually activating any particular lexical entry. It is even possible that all lexical decisions are based solely upon perceived stimulus familiarity (Besner, Davelaar, Alcott, & Parry, 1984). Thus, once again, DYME's ability to cause interference in the lexical decision task may tell us nothing about how we normally read DIME.

A result analogous to the pseudohomophone effect comes from experiments using a categorization task, in which subjects judge whether a target word is a category exemplar (e.g., Is PAIR an exemplar of the category FRUIT?). Meyer and Gutschera (1975) and Meyer and Ruddy (1973) reported that subjects in this task take more time to correctly reject PAIR for the category FRUIT than to reject other nonexemplars such as TAIL. Unfortunately, the authors failed to include spelling controls, which confounded homophony with spelling similarity. Another problem is that this effect, like the pseudohomophone effect, is an effect upon the time it takes to respond "no," and "no" responses in the categorization task are generally slower than "yes" responses. Thus, just as in the lexical decision task, a variable that delays "no" responses may reveal nothing of the identification process that precedes a "yes" response.

In order to determine whether phonological representations actually mediate word identification, we would like to observe an effect upon "yes" responses. Thus, a categorization task that both includes spelling controls and affects "yes" responses should directly address the issue of phonological mediation—unlike the pseudohomophone effect in the lexical decision task or effects upon "no" latencies in earlier categorization tasks. In the experiments reported here, subjects were presented with a category name (e.g., A FLOWER) followed by a target word (e.g., TULIP, ROWS, or ROBS). The phonological mediation hypothesis predicts that false positive categorization errors should occur as a function of the phonological similarity of a stimulus foil to some category exemplar. Accordingly, when the target word is a homophone foil such as ROWS, its phonological representation should strongly activate the lexical entry of the category exemplar ROSE prior to word identification. If so, the homophone foil ROWS will tend to be misidentified as the flower ROSE.

For this version of the categorization task, subjects are required first to judge whether a target word is an exemplar of the preceding category (by saying "yes" or "no") and then to name (identify) the target. The categorization judgment requires subjects to evaluate the meaning that is activated by a target word, but that judgment alone may not indicate which lexical entry is selected in word identification. For example, if the phonological features of stimuli such as ROWS merely accomplish lexical access

(in the general sense), the effect of phonological activation might be merely to increase the availability of all homophonic words' meanings, even though a category exemplar meaning is not eventually selected by word identification. In other words, phonology's errant activation of meanings could merely increase the subjects' bias toward responding "yes" in the categorization task, without actually causing false word identification. (Evidence for this yes-bias artifact hypothesis would be found if subjects produced responses such as "Yes, I mean, no, ROWS" to a foil such as ROWS that could indicate a mismatch between the outcomes of categorization and identification.) Therefore, it is only because subjects are also required to name the target stimulus that the source of false positive responses can be localized within the process of word identification. Word identification presumably is required before the identity and meaning of a particular word can be reported. Thus, if a subject responds /YES, ROZ/ to the stimulus ROWS (for the category A FLOWER), that subject has *categorized* (indicating the activation of the meaning of the flower ROSE) and *identified* (indicating the selection of a lexical entry corresponding to ROSE) the stimulus ROWS as the flower ROSE.

There are several advantages of this categorization task over the lexical decision task (the task used in most tests of the phonological mediation hypothesis). The categorization judgment, because it is a semantic judgment, cannot be based solely upon perceived stimulus familiarity (Besner et al., 1984). (Experiment 3 provided an additional test of this unlikely possibility.) Consequently, use of the categorization task precludes the possibility that performance circumvents lexical access (in the general sense) by employing stimulus familiarity. Additionally, this categorization task avoids the other problem noted above of extending inferences drawn from effects upon generally slower "no" latencies because observations are made upon the rate of false positives.

The categorization task also allows the use of word foils. If we are to extend inferences drawn from these data to reading in general, then we are best informed by observations of reading performance on word stimuli. Nonword stimuli (or unfamiliar word stimuli) could require a greater degree of phonological coding than is normally required in the skilled reading of familiar words (McCusker et al., 1981; Seidenberg, 1985; Seidenberg et al., 1984; Waters et al., 1984). (I will return to this issue later.)

Because most words that sound alike are also spelled alike, it is necessary to ensure that "sound-alike" words are not being miscategorized because of spelling similarity. One way to accomplish this is to yoke each homophone foil such as ROWS with a spelling control foil such as ROBS. A word such as ROBS is as orthographically similar to ROSE as the homophone ROWS by almost any reasonable metric (the metric used here is described in Appendix A). Thus, if ROWS is mistaken for ROSE because they are similarly spelled, then there should be no difference between the false positive error rate for homo-

phone foils (ROWS) and the false positive error rate for spelling controls (ROBS). However, if the sound of ROWS causes it to be categorized as A FLOWER, then subjects should make more false positive categorization errors to target foils such as ROWS than to spelling controls such as ROBS.

Dual-access theories, which allow both orthographic and phonological influences in word identification, would predict categorization errors both when foils sound like category exemplars and when they are spelled similarly to category exemplars. Therefore, it should be possible to see an effect of the orthographic similarity of foils independent of the effect of phonological similarity. This prediction was tested using homophone foils that varied in the degree to which they were spelled similar to their sound-alike category exemplars. Target foil MEET (for the category A TYPE OF FOOD) is very similar in spelling to MEAT. Target foil ROWS is less similar in spelling to ROSE. If an orthographic representation is used directly in word identification, then the likelihood that this process will be misled by orthographically similar foils should be a function of the number of orthographic characteristics that they share with a category exemplar. Thus, MEET is more likely to be mistaken for MEAT than ROWS is to be mistaken for ROSE.

It would be tempting to compare the error rates of similarly spelled (MELT) and less similarly spelled (ROBS) spelling control foils, possibly to provide an additional test for effects of spelling similarity. This comparison, however, is contaminated by the lack of control over phonological similarity between these two stimulus types. Although these two sets of stimuli were carefully controlled for spelling similarity to their respective, corresponding category exemplars, it was impossible also to control relative phonological similarity to the same respective, corresponding category exemplars.

EXPERIMENT 1

Experiment 1 included two factors. The manipulation of homophony (homophone foils such as MEET and ROWS vs. spelling control foils such as MELT and ROBS) was derived from the phonological mediation hypothesis. It tested for an effect of the phonological characteristics of homophone foils while controlling for the fact that these foils were orthographically similar to category exemplars. The manipulation of orthographic similarity (MEET, spelled similarly to MEAT, vs. ROWS, spelled less similarly to ROSE) tested for an effect of spelling similarity independent of the effect of homophony.

Method

Subjects. The subjects, 5 women and 5 men, all undergraduates at the University of California at San Diego, received credit for a course requirement for their participation. All were native English speakers and had normal or corrected-to-normal vision.

Procedure. The subjects were seated before a Gerbrands B1128 Harvard Model T-3A tachistoscope with which the stimuli were presented. Each trial began with the presentation of a category name

directly above a fixation point ("+"). The subjects' instructions were to read the category name (silently) and then to look at the fixation point. The fixation point indicated where a forthcoming target word would appear (the fixation point was always in the same position and always indicated the exact center of the expected target word). The target word was subsequently replaced by a pattern mask.

Upon presentation of the target word, the subjects were to respond orally "yes" or "no," indicating "yes, it is a category exemplar" or "no, it is not," and then to name the target word. Accuracy was stressed, but the subjects were also instructed to respond quickly. No accuracy or response time (RT) feedback was provided.

Each subject received 50 practice trials followed immediately by 122 experimental trials. Of the experimental trials, only 16.4% were trials in which the target words were homophone foils. This percentage was intentionally low to avoid inducing processing strategies by the presence of a high percentage of phonologically ambiguous stimuli (see, e.g., Davelaar, Coltheart, Besner, & Jonasson, 1978; Hawkins, Reicher, Rogers, & Peterson, 1976; McQuade, 1981).

The order of presentation for all trials in the practice task and the filler trials in the experimental task was the same for all subjects. However, for each subject, the key trials containing homophone or spelling control foils were assigned randomly to predetermined presentation slots separated by one, two, or three filler trials.

The entire experiment took approximately 50 min per subject. **Viewing conditions.** All stimuli were typed in capital letters on 4×6 in. white index cards in the IBM Selectric Orator typeface. The illumination controls for all three channels of the tachistoscope were set at their highest levels and held constant across subjects. Four-letter words subtended approximately .6° of visual arc horizontally and .3° vertically.

The timing of stimulus presentation was as follows: The category name and "+" remained visible for 1,500 msec and then were replaced by a target word. The target word was visible for 500 msec before it was replaced by a pattern mask. This stimulus-onset asynchrony (SOA) of 500 msec was chosen to allow the subjects adequate time to see the target word. The trailing pattern mask was included to match the conditions of Experiment 2 as closely as possible (excluding the SOA between targets and mask) and thus to promote comparisons between these two data sets.

Stimuli. Each subject saw 172 target words, 50 in the practice trials and 122 in the experimental trials. The subjects saw each target word only once. The key targets (see Appendix B) were 20 homophone foils (10 similarly spelled, e.g., MEET, and 10 less similarly spelled, e.g., ROWS) and 20 yoked spelling controls (e.g., MELT and ROBS). The 50 practice targets and 82 filler experimental targets comprised equal numbers of category exemplar targets and nonexemplar foils (nonexemplar foils were not chosen to be systematically similar in any respect to category exemplars).

Degree of spelling similarity was calculated using an estimate of orthographic similarity (OS) (see Appendix A) that ranges from 0 to 1, where 1 indicates an identical match. This estimate was used to construct two groups of homophone foils (foils such as MEET vs. foils such as ROWS) that differed in their degree of spelling similarity to corresponding category exemplars (MEAT and ROSE). It was also used to ensure that each spelling control (ROBS) matched its yoked homophone foil (ROWS) in spelling similarity to their corresponding category exemplar (ROSE). Mean OS for homophone foils in the similarly spelled condition (MEET) was .74, and mean OS for their yoked spelling control foils (MELT) was also .74. Mean OS for homophone foils in the less similarly spelled condition (ROWS) was .45, and mean OS for their yoked spelling control foils (ROBS) was .46.

Because the likelihood of false positives to a homophone foil such as ROWS is a function of the word frequency (Kučera & Francis,

1967) of its corresponding category exemplar (see Experiment 3), it was important to control this factor between the similarly spelled homophone foils (MEET) and the less similarly spelled homophone foils (ROWS). The mean Kučera and Francis frequency count of corresponding category exemplars was 26.1 for items such as MEAT and 31.7 for items such as ROWS. The difference between these means did not approach statistical significance [$t(18) = -.39$].

A final control took into account the length of stimulus foils. For each homophone (MEET) in the similarly spelled condition, there was a corresponding homophone (ROWS) of the same length in the less similarly spelled condition. Consequently, the mean lengths (4.5 letters) of the two groups were equal.

The pattern mask was constructed by overlapping alphabetic and nonalphabetic characters typed by the same element that was used to type other stimuli. To further ensure that the mask would not be seen as containing letters, it was inverted and then shown through the "mirror-reverse" channel of the tachistoscope. The mask subtended approximately 1.8° of visual arc horizontally and 1° vertically.

Results

The key trials were those that contained targets that were either homophone foils (MEET and ROWS) or yoked spelling controls (MELT and ROBS). The results in Table 1 are shown as the percentage of key trials in each condition that resulted in a false positive error.[1] By definition, a false positive categorization error to a foil such as ROWS occurred only if the subject's response was of the type /YES, ROZ/. Any other type of response was excluded from the error analysis. As noted previously, because this special definition of false positive was used, we can infer that the process of word identification is the source of false positive responses. In each comparison, separate statistics considering subjects and then items as random factors were computed. The percentage of false positives was the only dependent measure in Experiments 1 and 2. (I was unable to measure RTs because I was unable to obtain a voice key.)

The mean percentage of false positives for homophone foils such as MEET and ROWS (18.5%) was significantly greater than that for spelling controls such as MELT and ROBS (3%) [$t(9) = 3.90$, $p < .05$, for subjects, and $t(19) = 3.64$, $p < .05$, for items]. This effect of homophony provides strong support for the phonological mediation hypothesis, but is inconsistent with reading models that do not include a phonological source of influence upon word identification. Also, the percentage of false positives to similarly spelled homophone foils such as MEET (29%) was significantly greater than that to less similarly spelled homophone foils such as ROWS (8%) [$t(9) = 3.71$, $p < .05$, for subjects, and $t(9) = 2.27$,

Table 1
Percentage of False Positive Responses to Homophone and Spelling Control Foils in Experiment 1

	Similarly Spelled Foils		Less Similarly Spelled Foils		
	Mean	SE	Mean	SE	Mean
Homophones	29	6.9	8	2.9	18.5
Spelling Controls	5	3.1	1	1.0	3

$p < .05$, for items]. Thus, the results of Experiment 1 support reading models in which both orthographic and phonological representations influence the process of word identification.

Discussion

Homophones such as MEET and ROWS caused false positive errors on 18.5% of trials, whereas spelling controls such as MELT and ROBS caused such errors on only 3%. The error rate of 18.5% ($SE = 4.4\%$) to homophone foils is very close to the error rates of 16% and 14% that were reported for the categorization experiments of Meyer and Gutschera (1975). (Meyer & Ruddy, 1973, did not report the exact error rate for responses to homophones in their categorization experiment.) This effect of homophony upon "yes" responses cannot be dismissed as a latecomer to the process of word identification (e.g., an effect that arises only in generally slow "no" latencies—see also Experiment 2) and cannot be attributed to orthographic similarity (a possible confound in the experiments of Meyer and his colleagues). Rather, it is consistent with the phonological mediation hypothesis, in which a phonological representation of printed words mediates access to their meaning. This finding is evidence against any model that lacks a phonological source of influence in word identification.

Dual-access models were introduced to accommodate both phonological and orthographic effects in simple reading tasks. Thus, the results of Experiment 1 are generally consistent with all of these models, but these models do not all include assumptions pertaining to other factors that affect reading performance, such as pattern masking. Experiment 2 examined the process of word identification under pattern-masking conditions to test a prediction derived from the masking assumptions of one particular dual-access model, the interactive activation model of McClelland and Rumelhart (1981; Rumelhart & McClelland, 1982).

EXPERIMENT 2

The interactive activation model includes assumptions pertaining to the effects of pattern masking that are not present in most other dual-access models. These masking assumptions originally appeared in Johnston and McClelland's (1980) explanation of the word-superiority effect (Reicher, 1969; Wheeler, 1970), in which a pattern mask, composed of letter features, "overwrites" lower, letter-level (orthographic) representations but spares representations at the higher, word level (lexical representations). In other words, the word-superiority effect arises because word-level representations are less vulnerable to the effects of pattern masking than are letter-level representations.

Although Johnston and McClelland (1980) argued that word-level representations underlie the word-superiority effect, Johnston (1981, p. 78) hinted that the spared representation responsible for the word-superiority effect

could be phonological (see also Henderson & Chard, 1980). I do not necessarily assume that phonological coding underlies the word-superiority effect; however, the following experimental hypothesis does presume that phonological representations are relatively invulnerable to the effects of pattern masking. (Empirical support for this assumption comes from results reported in Hawkins et al., 1976, and Underwood and Thwaites, 1982, in which homophonic stimuli interfered with subjects' performance in tasks that included pattern masking.)

If phonological representations are relatively invulnerable to the effects of pattern masking, then any effect of phonology that is observed without pattern masking should also be observed under pattern-masking conditions. Specifically, the effect of homophony in Experiment 1 (more false positives to homophone foils such as ROWS than to spelling controls such as ROBS) should also be observed under pattern-masking conditions. However, the effect of orthographic similarity in Experiment 1 (more false positives to similarly spelled homophone foils such as MEET than to less similarly spelled homophone foils such as ROWS) should be reduced because orthographic representations are relatively vulnerable to the effects of pattern masking. These were the predictions tested in Experiment 2.

Method

Subjects. As in Experiment 1, the subjects, 5 women and 5 men, all undergraduates at the University of California at San Diego, received credit for a course requirement for their participation. All were native English speakers and had normal or corrected-to-normal vision.

Other than this new group of subjects and a change in the viewing conditions, the method in Experiment 2 was identical to that of Experiment 1.

Viewing conditions. The viewing conditions in Experiment 2 were identical to those in Experiment 1 except for the timing of stimulus presentation. In Experiment 2, SOA between the target word and the pattern mask was established separately for each subject. Because pilot work had shown that subjects varied widely in the critical SOA at which their categorization performance was affected by pattern masking, the following procedure was used to set SOA by individual performance.

This procedure was a method of descending limits. Practice began with a 150-msec SOA, but after several warm-up trials the SOA was decreased by increments of 10 msec. The SOA eventually became so brief that the subjects could not identify every target. The subjects had been advised of this possibility and were instructed not to guess, but to respond "I didn't see anything" if they could not identify the target (although this instruction was also included in Experiment 1, the response was used on less than .5% of trials). The critical SOA was that at which the subjects could still report a large percentage of the practice target words that were exemplars of their preceding category but could no longer report any practice target words that were not exemplars of their preceding category. Because the priming of category exemplars made them easy to identify (relative to identification of nonexemplars), this proved to be a straightforward method for setting critical SOA for the experimental trials.

In the experimental trials, if the subjects responded "I didn't see anything" for 10 consecutive trials, the SOA between target words and the pattern mask was lengthened by 10 msec (these trials were not replaced). Alternatively, if the subjects correctly reported three consecutive nonexemplar target words, the critical SOA was short-

ened by 10 msec. Thus, the critical SOA was set near threshold in the practice trials and was maintained throughout the experimental trials.

Results

As in Experiment 1, the key trials were those in which the target words were either homophones (MEET and ROWS) or yoked spelling controls (MELT and ROBS). The independent variables were homophony (MEET and ROWS vs. MELT and ROBS), orthographic similarity (MEET vs. ROWS), and, in a comparison of the data of Experiments 1 and 2, viewing conditions (long SOA of Experiment 1 vs. brief SOA of Experiment 2). The results in Table 2 are shown as the mean percentage of false positives in each condition. Again, as in Experiment 1, the only response to the examples ROWS or ROBS that met the special definition of false positive was /YES, ROZ/.

Orthographic and phonological similarity. As in Experiment 1, the planned comparison between the mean percentage of false positives to homophone foils (43%) and that to spelling controls (17.5%) revealed an effect of homophony [$t(9) = 6.41, p < .05$, for subjects, and $t(19) = 5.52, p < .05$, for items]. (The rate of unsystematic false alarms to filler nonexemplar foils was only 5.1%.) Unlike in Experiment 1, however, the percentage of false positives to similarly spelled homophone foils such as MEET (40%) was not significantly different from that to less similarly spelled homophone foils such as ROWS (46%) [$t(9) = -.90, p = .40$, for subjects, and $t(9) = -.77, p = .46$, for items].

Viewing conditions. Table 3 shows the change in relative error rates to similarly spelled homophones and less similarly spelled homophones between Experiments 1 and 2. As one can see, pattern masking eliminated the effect of orthographic similarity (the percentage of false positives to MEET was greater than that to ROWS in Experiment 1 but not in Experiment 2). The statistical significance of this interaction between orthographic similarity and viewing conditions was tested by comparing the data from the similarly spelled (MEET) and less similarly spelled (ROWS) homophone foils in the long SOA conditions of Experiment 1 versus the corresponding data set for the brief SOA (pattern-masking) conditions of Experiment 2. This analysis verified the apparent interaction between viewing conditions and orthographic similarity [$F(1,18) = 9.47, p < .05$, for subjects, and $F(1,9) = 14.88, p < .05$, for items]. (The data from spelling control trials were excluded from this analysis because, as noted previously, the spelling con-

Table 2
Percentage of False Positive Responses to Homophone and Spelling Control Foils in Experiment 2

	Similarly Spelled Foils		Less Similarly Spelled Foils		
	Mean	SE	Mean	SE	Mean
Homophones	40	4.2	46	6.4	43
Spelling Controls	22	5.1	13	1.5	17.5

Table 3
Percentage of False Positive Responses to Similarly Spelled and Less Similarly Spelled Homophone Foils in Experiments 1 and 2

	Experiment 1		Experiment 2	
	Mean	SE	Mean	SE
Similarly Spelled Homophone Foils	29	6.9	40	4.2
Less Similarly Spelled Homophone Foils	8	2.9	46	6.4

trol stimuli confound orthographic and phonological similarity to an unknown degree.[2]

Discussion

In Experiment 1, similarly spelled homophones (MEET) caused more false positive errors than less similarly spelled homophones (ROWS), but under the pattern-masking conditions of Experiment 2 this effect of orthographic similarity was not observed. In contrast, effects of homophony were observed in both experiments.

At least two points relevant to dual-access models follow: (1) Some orthographically determined process of word identification is highly vulnerable to the effects of pattern masking. This point is addressed by Johnston and McClelland's (1980) and McClelland and Rumelhart's (1981) assumptions about the effect of pattern masking in which lower, letter-level (orthographic) representations are especially vulnerable to masking. (2) Phonological mediation occurs soon enough to exert its influence under the very brief exposure conditions of pattern masking. As noted earlier, it is commonly assumed that the influence of phonological codes upon word identification is delayed relative to the influence of orthographic codes. If this were true, exposure conditions that require quick word identification should favor the effects of fast-acting orthographic codes. The pattern-masking conditions of Experiment 2 provided a situation in which word identification was best served by its most rapidly available sources of activation. Under those conditions, the effects of orthographic similarity disappeared, but the effect of homophony remained relatively unperturbed.

Further evidence against the delayed phonology hypothesis comes from Van Orden, Johnson, and Hale (in press). They assumed that if phonological activation is generally delayed relative to orthographic activation, then false positive "yes" RTs to foils such as ROWS should be generally slower than correct "yes" RTs to yoked control category exemplars such as TULIP (for the category A FLOWER). Contrary to this assumption, they found only a small difference, concentrated in their slow RT tails, between these two RT distributions.

It is important to note here that there is no direct evidence that supports the delayed phonology hypothesis. Rather, this hypothesis is inferred from the fact that most previously observed phonology effects were delays in "no" latencies, and "no" latencies are generally longer than "yes" latencies (as noted earlier), but the present phonological effects upon "yes" responding undermine

the basis for this inference. Furthermore, within this logic, the failure to find effects of orthographic similarity under masking conditions in Experiment 2, coupled with the Van Orden et al. (in press) finding of nearly identical distributions of false positive and correct "yes" latencies, is hypothetically consistent with the possibility that some orthographically determined subprocess of word identification is delayed relative to phonological activation. The following discussion explores this possibility.

The results of Experiments 1 and 2 indicate that representations derived from both the sound (and/or pronunciation) and the spelling of words are used in word identification. These results do not, however, provide exclusive support for dual-access models. The following outline of a verification model illustrates this point. This model is a variant of the general scheme proposed by Rubenstein et al. (1971), Becker (1976, 1980), Schvaneveldt and McDonald (1981), and Paap et al. (1982).

The presentation of a target word is assumed to activate many candidate lexical entries exclusively via input from a phonological representation. (A possible mechanism of phonological coding is described in the General Discussion.) Before an activated lexical entry can be selected, it must pass a verification test, which is essentially a spelling check. The orthographic representation of the most active lexical entry is retrieved from memory and compared with the orthographic representation of the stimulus word. If a match occurs, the lexical entry is selected; otherwise, the process is repeated using the next most active candidate entry.

In this model, bottom-up activation comes from the phonological representation, whereas the verification process operates upon the orthographic representation. The model easily explains the results of Experiment 1 because a false candidate is more likely to be available to the verification procedure if it is phonologically similar to the stimulus word (thus the high error rate to homophone foils). In turn, a false candidate is more likely to slip by the verification procedure if it is orthographically similar to the stimulus word (thus the highest error rate to similarly spelled homophone foils in Expeirment 1). Pattern masking interrupts processing before the verification procedure occurs (Paap et al., 1982), and, consequently, effects that result from verification are not found under masking conditions. Thus, the effect or orthographic similarity disappears under the masking conditions of Experiment 2. This verification hypothesis can be differentiated from dual-process theory (Coltheart, 1978), possibly the most influential dual-access theory, by the way in which readers are assumed to disambiguate homophones. The dual-process view is illustrated (and supported) by an experiment reported by Doctor and Coltheart (1980).

Doctor and Coltheart (1980) used a sentence verification task in a cross-sectional study to track the development of the direct route of lexical access. Consistent with dual-process theory, they found that children's false positive error rates to sentence foils that sound meaningful

(e.g., "She blue up the balloon") decreased as a function of the children's age (and, presumably, their degree of reading experience). The mean false positive error rate for beginning readers (children 6 years of age) was 70.8%, but the most experienced readers (children 10 years of age) had a corresponding error rate of only 20.8%. If we assume (as does dual-process theory) that older children (experienced readers) bypass phonology in favor of direct access for the bulk of their reading vocabulary, we can see why older children are less likely to make false positive errors to "sound-okay" foils. This bypass hypothesis makes the general prediction that experimental effects attributable to phonological mediation should not be found when the experimental stimuli are familiar words—words that subjects have read many times before and for which phonology is bypassed.

Doctor and Coltheart's (1980) evidence for the bypass hypothesis—that older children are less likely to be seduced by homophonic impostors in a sentence verification task—is also accommodated by the verification hypothesis. In the verification account, word identification always includes phonological mediation, but the mechanism of verification develops to avoid phonological confusions. The efficacy of verification, however, depends upon the acquisition of complete knowledge of a word's spelling, and the completeness of a reader's knowledge about a word's spelling would be a function of his/her overall reading experience. If we assume that the older children in Doctor and Coltheart's study had relatively better knowledge of words' spellings, we can attribute their lower false positive error rates to more efficacious verification.

A categorization task with homophonic foils could discriminate between this verification account versus the bypass account of readers' ability to disambiguate correctly most homophonic words. For example, consider again the homophonic foil ROWS for the category A FLOWER. The bypass hypothesis suggests that when the stimulus ROWS is a familiar word, a direct association exists between its spelling and its entry in the lexicon. This direct association will bypass phonological coding and circumvent the influence of the ambiguous phonology of ROWS. Thus, in the categorization task, if stimulus foil ROWS is a high-frequency word, direct access will bypass the possibility of miscategorizing ROWS as A FLOWER.

Alternatively, the verification hypothesis is not concerned with the familiarity of stimulus ROWS. Rather, it is familiarity with ROSE, the corresponding category exemplar for ROWS, and the consequent knowledge of the spelling of ROSE, that will allow readers to avoid a false positive categorization error to the stimulus foil ROWS. If exemplar ROSE is a high-frequency word, readers are more likely to have complete knowledge of its spelling and are thus more likely to detect the stimulus impostor ROWS. Consequently, the likelihood that ROWS will be miscategorized as A FLOWER is reduced when ROSE is a high-frequency word.

EXPERIMENT 3

These are the alternative predictions tested in Experiment 3. The bypass hypothesis predicts that the percentage of false positive categorization responses to homophonic foils should be lowest when actual stimulus foils such as ROWS are high-frequency words. The verification hypothesis predicts that the percentage of false positive responses to stimulus foils such as ROWS should be lowest when corresponding sound-alike category exemplars such as ROSE are high-frequency words. In other words, the bypass hypothesis predicts an effect of the frequency of the stimulus homophone foils that appear in the experiment, but the verification hypothesis predicts an effect of the frequency of corresponding category exemplars that are never presented.

Method

Subjects. The new group of subjects in Experiment 3, 5 women and 5 men, were also undergraduates at the University of California at San Diego and received credit for a course requirement for their participation. All were native English speakers and had normal or corrected-to-normal vision.

Procedure. The subjects were seated facing a Tectronix 601 storage scope upon which stimuli were presented. Each trial began with the appearance of the word READY. The subjects proceeded by pressing a READY button with the left hand. As the READY button was pressed, a category name appeared. The category name remained visible for 1,500 msec, after which it was replaced by a ".", which served as a fixation point. The "." remained on the screen for 500 msec, after which it was replaced by a target word (the "." always appeared at the center of the forthcoming target word). The target word remained visible until the subject responded. The subjects indicated that a target was an exemplar of its preceding category by pressing a "yes" key with the right index finger and that a target was *not* an exemplar by pressing a "no" key with the right middle finger. Neither accuracy nor RT feedback was provided. Immediately following a subject's response, the READY signal appeared again, to signal the beginning of the next trial. The subjects were instructed to be accurate, but were also instructed to respond quickly.

The special definition of false positive categorization error that was used in Experiments 1 and 2 was not used in Experiment 3. That special definition was used to ensure that incorrect "yes" responses to homophone foils did not result merely from a "yes" bias caused by errant meaning activation. As I noted in the introduction to Experiment 1, mere "yes" bias might be indicated by responses of the type "Yes, I mean, no, ROWS" (hereafter "y→n") to a stimulus such as ROWS. In Experiment 1, which like Experiment 3 used a categorization task with adequate exposure time, the y→n response was observed on 5.5% of homophone foil trials and 2% of spelling control trials [$F(1,9) = 2.19, p < .18$, for subjects, and $F(1,19) = 2.26, p < .15$, for items]. Thus, there is some indication of a "yes" bias caused by homophony. (This y→n response was found on less than .3% of filler trials.) However, there is equally compelling evidence that homophony causes a "no" bias. "No, I mean, yes, ROSE" type responses (hereafter "n→y") to foils such as ROWS resulted on 2% of homophone foil trials and on 0% of spelling control trials [$F(1,9) = 3.27, p < .11$, for subjects, and $F(1,19) = 4.75, p < .05$, for items]. (This n→y response was also found on less than .3% of filler trials.) Thus, a bias hypothesis must assume that homophone foils induce both positive and negative bias.

An alternative possibility is that these types of responses reflect successive, inconsistent outcomes of an iterative verification process,

possibly like the resonant process proposed by Grossberg and Stone (1986). Further support for this possibility comes from the fact that y→n and n→y responses almost never occur (.25% of key trials for both of these types of mismatched responses) under the masking conditions of Experiment 2. This would be consistent with the previous assumption that pattern masking interferes with effects that are due to verification. However, in any case, given this failure to support exclusively a yes-bias artifact hypothesis, it is reasonable to dispense with the special definition of false positive that was necessary in Experiments 1 and 2. Notice also that requiring merely a "yes" response before the attribution of false positive is consistent with Doctor and Coltheart's (1980) assumptions concerning false positives in their sentence verification task.

Each subject received 40 practice trials and 160 experimental trials. Stimuli in the practice trials were presented in the same order for all subjects; in the experimental trials, however, each subject saw a different randomly generated ordering of stimuli. The entire experiment took approximately 25 min.

Viewing conditions. All stimuli were presented on the Tectronix 601 storage scope controlled by a PDP-12/30 computer. The experiment was conducted in a dark room for maximum stimulus visibility. The focus and intensity of the storage scope were adjusted for the clearest image (as judged by the experimenter) and were held constant across subjects.

Letters were presented with a 4×6 dot matrix. Four-letter words subtended approximately .4° of visual arc with .794 mm between letters. Each letter subtended approximately .09° vertically. (Using stimuli that subtend larger visual angles, Greg Stone and I have since replicated this experiment.)

Stimuli. Each subject saw 200 target words. The subjects saw each target word only once. None of the category names used in the practice trials reappeared in the experimental trials. The key targets were the 20 homophone foils and their 20 yoked controls (see Appendix C).

The homophone foils were chosen such that they could be sorted into four frequency groups of five items per group either by their own frequency counts (the frequency count of ROWS) or by the frequency counts of their corresponding category exemplars (the frequency count of ROSE). This allowed the alternative experimental hypotheses to be tested upon the same data. Because this plan would be compromised if the frequency counts of the stimulus foils (ROWS) were correlated with the frequency counts of their corresponding category exemplars (ROSE), significant correlation was avoided [$r = -.17, t(18) = -.71$]. The four frequency groups were: (1) a very low-frequency group (Group VL), with frequency counts less than 1 per million (words that do not appear in Kučera & Francis, 1967); (2) a low-frequency group (Group L), with frequency counts of 1-9 per million; (3) a high-frequency group (Group H), with frequency counts of 10-99 per million; and (4) a very high-frequency group (Group VH), with frequency counts of greater than 99 per million. These groups were the same for both arrangements of the stimuli—whether items were grouped using their own frequency counts (ROWS) or that of their corresponding category exemplar (ROSE).

Each homophone foil was yoked to a control word, but these control foils were not spelling controls as in Experiments 1 and 2. The need for homophone pairs of specific frequency counts severely limited the pool of candidate homophones, and, consequently, spelling controls were unavailable for most of these candidate homophones. However, because the logic of this experiment concerned comparisons between frequency groups of homophone foils, but not between homophone foils and spelling controls, it was necessary only that spelling similarity between the homophone foils and their corresponding category exemplars be approximately equal across the respective frequency groups (see below). But, control items were still included to establish a baseline for the rate of false positives to items other than homophones. For this purpose, yoked controls were matched to homophone foils along dimensions other

than spelling similarity. These yoked pairs were identical in number of letters and in part of speech, and as close as possible in frequency count. (Since many words can function in more than one part of speech, "part of speech" was estimated by the part of speech of each word's first entry in *The American Heritage Dictionary*; Morris, 1976.)

To control for spelling similarity between homophonic foils and their corresponding category exemplars, a source of false positives in Experiment 1, both arrangements of the frequency groups were matched closely in spelling similarity using OS (see Appendix A). To test for between-group differences in spelling similarity, one-way analyses of variance were computed for each of the arrangements of the stimuli, with OS as the dependent variable and items as the random factor. No significant differences were found when the frequency groups were assigned by stimulus homophone (ROWS) frequency counts [$F(3,16) = .70$] or by corresponding category exemplar (ROSE) frequency counts [$F(3,16) = .26$]. Mean OS for the VL, L, H, and VH frequency groups were .70, .65, .64, and .64 when groups were arranged according to the frequency count of ROWS, and .66, .65, .68, and .64 when groups were arranged according to the frequency count of ROSE.

Results

The key trials were those in which the target words were either homophone foils (e.g., ROWS) or yoked controls. The primary dependent measure was the percentage of these key trials that resulted in a false positive categorization response (e.g., "yes" indicating ROWS is A FLOWER), but RTs were also measured. The independent variables were frequency group (VL, L, H, and VH) and homophony (homophones vs. controls).

Two analyses of variance were performed upon the error data: one in which homophone foils were assigned to frequency groups by their own (ROWS's) frequency count and a second, alternative arrangement of the data in which homophone foils were assigned to frequency groups by their corresponding category exemplar's (ROSE's) frequency count. The first analysis tested the prediction of the bypass hypothesis that false positives should be least likely when stimulus homophone foils such as ROWS are very high in frequency. The second analysis tested the prediction of the verification hypothesis that false positives should be least likely when corresponding category exemplars such as ROSE are very high in frequency. (Note again that the second analysis tested for an effect of the frequency of category exemplars such as ROSE, words that never appeared in the experiment.)

The effect of stimulus frequency on error rates. As can be seen in the upper portion of Table 4, the mean false positive error rate to homophone foils (30.5%) was much higher than the error rate to yoked control items (2.5%) [$F(1,9) = 18.77, p < .05$, for subjects, and $F(1,16) = 36.894, p < .05$, for items]. However, neither the effect of stimulus foil frequency nor the frequency group × type of foil interaction was significant. This was true for both subject [$F(3,27) = 1.40, p = .26$, and $F(3,27) = 2.11, p = .12$, respectively] and item analyses [$F(3,16) = .43, p = .74$, and $F(3,16) = 1.19, p = .34$, respectively].

The effect of category exemplar frequency on error rates. The lower portion of Table 4 also shows the high false positive error rate to homophone foils (33.2%) over controls (2.5%) [$F(1,9) = 19.18, p < .05$, for subjects, and $F(1,15) = 70.44, p < .05$, for items].[3] In this case, however, the percentage of false positives to homophone foils decreases across Groups VL, L, H, and VH (error rates of 55%, 40%, 22%, and 16%, respectively), but no effect of exemplar frequency is apparent in the error rates to control foils (2%, 4%, 2%, and 2%, respectively). The effect of frequency group [$F(3,27) = 7.48, p < .05$, for subjects, and $F(3,15) = 5.90, p < .05$, for items] and the frequency × homophony interaction [$F(3,27) = 5.39, p < .05$, for subjects, and $F(3,15) = 5.66, p < .05$, for items] were both significant, indicating an effect of the frequency of corresponding category exemplars such as ROSE upon the rate of false positives to homophone foils such as ROWS.

"No" latencies. The data for this analysis were taken from yoked trials in which each particular subject correctly rejected both a homophone foil and its corresponding yoked control. For example, if a trial containing a homophone foil such as ROWS resulted in a false positive, the trial that contained its yoked control was excluded from the analysis of "no" RTs, and vice versa. The remaining data were subjected to two ANOVAs corresponding to the two arrangements by frequency counts. Table 5 presents the mean "no" latencies for each of the possible frequency groupings. Neither the effect of frequency group nor the frequency group × homophony interaction was significant in either of the two analyses [frequency of ROWS—upper portion of Table 5—$F(3,27) = .12$, for

Table 4
Percentage of False Positive Responses to Homophone Foils Such As ROWS and Control Foils for Frequency Groups VL, L, H, and VH in Experiment 3 When Frequency Groups Are Assigned Using Stimulus Homophone Foil Frequency Counts and Category Exemplar (ROSE) Frequency Counts

	VL		L		H		VH		
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean
Frequency Group Assigned by Stimulus Frequency Count									
Homophones	24	8.8	30	7.5	42	8.7	26	7.3	30.5
Controls	6	4.3	4	2.7	0	0	0	0	2.5
Frequency Group Assigned by Category Exemplar Frequency Count									
Homophones	55	13.8	40	7.9	22	6.3	16	5.8	33.2
Controls	2	2.0	4	2.7	2	2.0	2	2.0	2.5

Table 5
Correct "No" RTs (in Milliseconds) to Homophone Foils Such As ROWS and Control Foils for Frequency Groups VL, L, H, and VH in Experiment 3 When Frequency Groups Are Assigned Using Stimulus Homophone Foil Frequency Counts and Category Exemplar (ROSE) Frequency Counts

	VL		L		H		VH		
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean
Frequency Group Assigned by Stimulus Frequency Count									
Homophones	1077	98	983	96	1063	151	1081	145	1051
Controls	897	88	949	80	904	116	801	38	888
Frequency Group Assigned by Category Exemplar Frequency Count									
Homophones			1046	112	1108	123	1007	80	1053
Controls			887	112	892	62	927	86	902

subjects, and $F(3,16) = 1.4$, $p = .28$, for items, frequency of ROWS × homophony—$F(3,27) = 1.01$, $p = .40$, for subjects, and $F(3,16) = .71$, for items; frequency of ROSE—lower portion of Table 5—$F(2,18) = .35$, for subjects, and $F(2,12) = .12$, for items, frequency of ROSE × homophony—$F(2,18) = .58$, for subjects, and $F(2,12) = .68$, for items]. (The RT analysis by the frequency count of ROSE included only three frequency groups because too few correct responses were given to the items in Group VL.)

In both alternative analyses, however, mean "no" latencies to homophone foils such as ROWS (1,051 msec) were significantly longer than "no" latencies to yoked control foils (888 msec) [$F(1,9) = 14.51, p < .05$, for subjects, and $F(1,16) = 8.84$, $p < .05$, for items]. (These means and F values come from the analysis of the effect of the frequency of ROWS because it included all of the possible data points that pertain to this comparison.)

Discussion

The bypass hypothesis, a fundamental hypothesis of dual-process theory (Coltheart, 1978), is not supported by the results of Experiment 3. The likelihood of mistaking ROWS for the flower ROSE in the categorization task was not found to be a function of the frequency of the target, ROWS. In fact, contrary to the bypass hypothesis's prediction (that the rate of false positives would decrease as a function of stimulus frequency), the error rate to the VH frequency homophone foils was slightly higher than the error rate to the VL frequency homophone foils. In contrast, the predictions of the alternative verification hypothesis are supported by the results of Experiment 3: False positive error rates to foils such as ROWS decreased as the frequency (and presumably the familiarity) of sound-alike category exemplars such as ROSE increased.

Although the results of Experiment 3 seem to provide unequivocal support for the verification hypothesis, there are several possible confounds of category exemplar word frequency that should be considered. Specifically, it is possible that this seeming effect of category exemplar fre-

quency is actually an effect of category exemplar typicality or production frequency, since all three of these variables have been shown to be correlated (Uyeda & Mandler, 1980).

Concerning production frequency and the present data, I obtained production frequency norms for the categories of Experiment 3 by using Battig and Montague's (1969) method with 62 undergraduates at the University of California at San Diego. This method involves presenting subjects with a category name and asking them to write down as many exemplars as possible of that category within 30 sec. For each exemplar, a tally is then made of how many subjects produced that exemplar, and that score is the exemplar's production frequency. A correlational analysis using these norms failed to yield a significant correlation between production frequency and rate of false positives [$r = -.22$, $t(18) = -.98, p = .34$], and a partial correlational analysis found no residual effect of production frequency ($pr = 0$) that is independent of the effect of category exemplar word frequency.

Concerning typicality, Van Orden et al. (in press) conducted a categorization experiment very similar to Experiment 3; their experiment used stimuli that enabled the isolation of the respective effects of typicality and production frequency. All of their homophone foils were chosen to sound like exemplars that appear in the typicality norms of Uyeda and Mandler (1980) and the production frequency norms of Battig and Montague (1969). In correlational analyses of their data, virtually no effects of either typicality or production frequency were found, although a strong correlation was found between category exemplar (ROSE) frequency and false positive error rates to homophone foils such as ROWS. Thus, these analyses not only rule out the potential confounding effects of typicality and production frequency upon error rates to homophone foils, but also replicate the effect of category exemplar frequency. (To date, Greg Stone and I have replicated this effect in two other categorization experiments and have found an analogous effect upon the rate of false positives to nonword homophone foils in a lexical decision task. Additionally, Alice Healy and I have

found an analogous effect upon the rate of proofreading "miss" errors using homophone foils in a proofreading task. Clearly, the effect is robust.)

Before discussing "no" latency results, I must note that almost all current dual-access theories include a further assumption in addition to the bypass hypothesis concerning word frequency. This threshold hypothesis was borrowed from signal detection theory (Green & Swets, 1966) by Morton (1969) for his logogen theory of word perception. For the categorization task, this hypothesis predicts that the frequency of category exemplars such as ROSE will affect the rate of false positives to foils such as ROWS. The phonological and orthographic similarity of the stimulus ROWS to the category exemplar ROSE will provide perceptual evidence for misfiring the logogen for ROSE; this misfiring is more likely to occur the more frequently the logogen for ROSE has fired in the past (which is directly related to the word frequency of ROSE). Notice that this prediction is exactly opposite to what occurred in Experiment 3. The threshold hypothesis predicts that the *highest* percentage of false positives should be observed for foils that sound like very high-frequency category exemplars—exemplars with the lowest thresholds. In Experiment 3, however, the *lowest* percentage of false positives was observed for foils that sound like very high-frequency category exemplars.

The results of the "no" RT analysis replicated the finding of Meyer and Gutschera (1975) and Meyer and Ruddy (1973) that mean "no" RTs to homophonic foils are prolonged relative to control foils in the categorization task. However, close inspection of the distributions that underlie these means revealed that the RT distribution for homophone foils was not generally shifted toward longer "no" latencies. Rather, these two distributions were nearly identical at their fast ends but differed in their high (slow) RT tails, where there were more outlier RTs for homophone foils. Thus, if these distributions were truncated just below their long RT tails, their recomputed means might not differ. This point is illustrated in Table 6, which shows the mean times for the relevant conditions as the underlying distributions are repeatedly truncated using successively decreasing cutoff RTs. (In order to preserve the within-subject comparison ensured by item yoking, the corresponding yoked control trial was deleted whenever a trial was lost for a homophone item, and vice versa.) Notice there that the initial difference of 163 msec is eventually diminished to 3 msec with a cutoff of 800 msec.

Van Orden et al. (in press) found the same similarity between distributions of "no" RTs to homophone foils and control foils. Their categorization experiment included many more subjects, which yielded a much larger sample of "no" latencies, and, just as in Experiment 3, they found a huge overlap between the two distributions, with the only difference being in the slow RT tails.

I did not test for effects upon "yes" latencies in Experiment 3 because there were too few data points and there was no control over factors such as typicality that influence "yes" RTs in categorization tasks. However, as I noted earlier, Van Orden et al. (in press) conducted an experiment very much like Experiment 3; in their experiment, they compared correct "yes" latencies to category exemplars such as ROSE and false positive "yes" latencies to foils such as ROWS. Those distributions of correct and false positive "yes" latencies were nearly identical, with only a small difference in the slow RT tails, similar to what I (and they) found for "no" latencies.

Both the present latency results and the results of Van Orden et al. (in press) suggest that neither false positive "yes" RTs nor correct "no" RTs are generally shifted toward longer RTs. This observation is inconsistent with at least one possible conception of verification: that it is included or not included in word identification on a trial-by-trial basis. That is, it does not seem to be a strategy that is invoked only on trials that include homophones, necessarily resulting in generally longer RTs to those stimuli (e.g., see Rubenstein et al., 1971). Alternatively, either word identification always includes verification, irrespective of the presence or absence of homophone foils, or verification is a general strategy for all trials invoked when homophone foils are present in an experiment, but not otherwise. It is impossible at present to distinguish between these two possibilities because the primary evidence for verification models comes from experiments that all include pseudohomophone foils in lexical decision tasks (e.g., see Becker, 1976, 1980; Becker & Killion, 1977; but cf. Schvaneveldt & McDonald, 1981). In either case, however, it is reasonable to assume that the verification procedure sometimes finds itself stuck between acceptance and rejection criteria, possibly resulting in additional cycles of memory retrieval and spelling check. This would explain the occasionally exaggerated "yes" and "no" RTs to homophone foils.

Before I begin a general discussion, please note several observations extraneous to the present experimental manipulations, but relevant to interpreting the results of these experiments. First, in the debriefing that followed all three experiments, I asked the subjects about their experience with the task. No subjects in Experiments 1 and 3 reported any confusion about whether the appropriate response to homophone foils was "no." Consistent with these reports, all of these subjects correctly rejected some portion of the homophone foils. (This was not true for the subjects in the pattern-masking conditions of Experiment 2: Fewer than half of them reported having seen any homophone foils.)

Table 6
Successive Mean Correct "No" RTs (in Milliseconds) to Homophone Foils Such As ROWS and Control Foils, and Their Differences, As the Underlying RT Distributions Are Repeatedly Truncated Using Successively Faster Cutoff RTs

	Cutoff RT				
	None	2000	1000	900	800
Homophones	1051	911	750	715	651
Controls	888	861	723	702	648
Difference	163	50	27	13	3

A second point concerns verification. In Experiment 3, I framed the predictions of the verification hypothesis solely in terms of the relative "completeness" of knowledge about the spellings of category exemplars. Although that hypothesis correctly predicted the outcome of Experiment 3, I have other evidence that suggests that readers in the present experiments usually knew the correct spelling of the category exemplars, but that this knowledge was not available to influence the outcome of verification. I have conducted several spelling experiments in which subjects like those in Experiments 1–3 were asked to spell category exemplars such as ROSE that correspond to the homophone foils used in the present experiments. Except for the VL frequency category exemplars of Experiment 3, almost none of these exemplars were ever misspelled. If correct performance in the spelling test indicates complete knowledge of spelling, the high rates of false positives to foils such as ROWS cannot be attributed entirely to incomplete spelling knowledge. Consequently, veridical verification may require both complete and *readily available* spelling knowledge, and completeness and availability may both be a function of familiarity.

GENERAL DISCUSSION

The results of the three experiments reported here are inconsistent with several common assumptions about word identification. Contrary to the ubiquitous delayed phonology hypothesis—that phonological mediation is generally delayed relative to direct access—all of the present experiments found an effect of homophony upon the rate of false positive categorization errors. This effect upon "yes" responding cannot be dismissed as a latecomer to word perception, as could previous effects upon generally slow "no" latencies. Additionally, in Experiment 2, the effect of homophony is found under pattern-masking conditions, in which we might assume that the most readily available codes would predominate. Thus, these results support a theory of word identification in which phonological codes are an early, important source of constraint.

Contrary to the bypass hypothesis—that phonology is bypassed for high-frequency words—the effect of homophone phonology in the categorization task does not decrease as stimulus frequency increases. Using a more extreme manipulation of familiarity, Van Orden et al. (in press) also observed this null result. They compared the rate of false positives to word-homophone foils such as BEATS (for the category VEGETABLE) with the rate of false positives to nonword-homophone foils such as SHEAP (for the category A FOUR-FOOTED ANIMAL) that were controlled for spelling similarity to category exemplars and category exemplar frequency. In two experiments, which used both different items and subjects, they found virtually identical error rates for responses to these two types of foils. (For this argument, pronounceable nonwords are viewed as extremely low-frequency words, and a comparison between word and nonword homophone

foils is a comparison between stimuli of different frequencies.) Thus, the null effect of stimulus frequency in Experiment 3 is not especially surprising, given the similar negative findings of Van Orden et al. (in press).

The present failure to find evidence for the action of direct access for high-frequency words is especially paradoxical within dual-process theory because proponents of that theory often take the ambiguous phonology of homophonic words to be prima facie evidence for the direct-access route (e.g., see Ellis, 1984). Presumably, without direct access, readers would lack the ability to disambiguate homophones, but the verification hypothesis offers an alternative to this assumption.

The null finding of Experiment 3 and the Van Orden et al. (in press) null findings join several other failures, reviewed by Jorm and Share (1983), to support the bypass hypothesis. In fact, to my knowledge, the only finding that directly supports the bypass hypothesis is that of Doctor and Coltheart (1980). The verification hypothesis offers an alternative explanation of Doctor and Coltheart's observation that children with more reading experience are less likely to make false positives to homophone foils in a sentence verification task—with experience, children gain knowledge of the correct spellings of words and use this knowledge to reject homophonic impostors. This verification hypothesis also correctly predicted the frequency effect observed in Experiment 3, that subjects made fewer false positive responses to homophone foils that sounded like familiar category exemplars.

The extent to which phonology affects performance in the present experiments is underscored by the simple verification model's relatively comprehensive account of the results of these experiments, even though it lacked a mechanism of direct access. This is not to say that I deny the possibility of direct access. I am, however, concerned that current theories overly rely upon a separate, sometimes poorly specified, mechanism of direct access because previous methods have failed to elicit phonology's role in word identification (see related discussions in Bauer & Stanovich, 1980; Besner, Davies, & Daniels, 1981; Glushko, 1981; and McCusker et al., 1981). If the present analysis is correct, it is clear how small changes in previous categorization task methods resulted in big changes in our ability to see phonology's influence upon word identification. What may be less clear, however, is how any mechanism of phonological coding can operate with the general efficiency of direct access. My final comments include a brief description of a mechanism that, I believe, operates in just such a fashion.

Consistency, Frequency, and Phonological Coding

A critical finding concerning the mechanism of phonological coding is that reading performance is affected by the correspondence between spelling and sound (Andrews, 1982; Baron & Strawson, 1976; Barron, 1980; Bauer & Stanovich, 1980; Glushko, 1979, 1981; Gough & Cosky, 1977; Stanovich & Bauer, 1978). One of the ways in

which this correspondence is captured empirically is in the consistency/inconsistency distinction (Andrews, 1982; Bauer & Stanovich, 1980; Glushko, 1979, 1981). In naming tasks, performance on consistent words is faster than performance on inconsistent words. However, this consistency effect is found only for low-frequency words (Andrews, 1982; Backman, Bruck, Hebert, & Seidenberg, 1984; Seidenberg, 1985; Seidenberg et al., 1984; Waters et al., 1984).

This interaction between frequency and consistency can be interpreted in several ways. The bypass hypothesis explains this interaction if it is assumed that high-frequency words bypass phonology and thus bypass the possibility of consistency effects. But the bypass hypothesis is not supported otherwise by the results of Experiment 3 (see also Jorm & Share, 1983). Seidenberg and his colleagues (Seidenberg, 1985; Seidenberg et al., 1984; Waters & Seidenberg, 1985; Waters et al., 1984) have proposed a variant of the delayed phonology hypothesis that also explains this interaction. They assume that, in general, delayed phonological activation will influence performance only on more slowly recognized lower frequency words. But the results of all of the present experiments (homophony effects upon the percentage of false positives, homophony effects under pattern-masking conditions, and the failure to find any effect of stimulus frequency in Experiment 3), as well as the results of Van Orden et al. (in press; false positive "yes" RTs to homophone foils are not generally slower than correct "yes" RTs to category exemplars, and two replications of the null effect of stimulus frequency in Experiment 3) are best explained by nondelayed phonological activation that is a primary influence upon word identification, irrespective of word frequency.

Covariant Learning and Phonological Coding

Alternatively, I am attracted to a third interpretation of the interaction between frequency and consistency. I propose that this effect can be attributed entirely to the mechanism of phonological coding, a computational mechanism that transforms orthographic codes into phonological codes. Specifically, this interaction arises because the mechanism that acquires associations between orthographic features and phonological features is sensitive to a statistical regularity implicit in the consistency/inconsistency distinction. That is, the mechanism that associates orthographic and phonological features benefits from a consistent covariance between these two types of features across many words, and this benefit is expressed in reading performance by faster responses (see Lewicki, 1986, for a related demonstration of covariant learning). I also assume, however, that learning is asymptotic in this hypothetical mechanism, or, put differently, overlearning can compensate for the initial disadvantage resulting from inconsistency. Thus, phonological codes of very familiar words, whether they be consistent or inconsistent, are all computed with equal efficiency.

The delta rule is one of the best understood of the learning algorithms that work in this prescribed fashion (see especially Rumelhart, Hinton, & Williams, 1986, and Stone, 1986; but for related discussions and applications, see Anderson, 1983; Kohonen, 1977, 1984; McClelland & Rumelhart, 1985; Rumelhart, Hinton, & McClelland, 1986; Rumelhart & McClelland, 1986; and Sutton & Barto, 1981). I have used the delta rule to simulate the learning that precedes artificial (simulated) instances of reading aloud for a very small neighborhood of consistent and inconsistent words. In that simulation, this error-correction learning rule repeatedly adjusted a memory matrix of associative weights between each orthographic feature and every phonological feature so that they came to reflect the covariation between orthographic input patterns and phonological output patterns (see also the closely related artificial intelligence studies of Rosenberg & Sejnowski, 1986, and Sejnowski & Rosenberg, 1986).

Adaptive filters such as the delta rule produce an output vector that can be evaluated against an ideal output vector (Kohonen, 1984). For the simulation, the elements of the output vector represented phoneme identities and positions analogous to McClelland and Rumelhart's (1981) scheme for representing letter identities and positions in their interactive-activation model. "Frequency" corresponded to the number of learning trials for a particular word. "Response time" corresponded to the Euclidean distance (error) between the model's output vector and an ideal output vector. Assuming that some mechanism enhances the ideal phonological features in the rough output vectors (e.g., interfeature competition, as in McClelland & Rumelhart, 1981; resonance, as in Grossberg & Stone, 1986; or the "brain-state-in-a-box" feedback of Anderson, 1977, 1983; Anderson & Mozer, 1981), then it is also reasonable to assume that the time function of this mechanism generally is directly related to the length of the error vector. In the simulation, the lengths of error vectors for high- and low-frequency, consistent and inconsistent words were a qualitative match for the interaction between frequency and consistency that is observed in readers' naming times.

This covariant learning hypothesis may also explain effects of spelling-sound correspondence other than consistency effects. For example, we might expect the frequency of particular grapheme-phoneme correspondences to affect naming time because frequency of covariance will combine with consistency of covariance to determine the associative weights between orthographic and phonological features. Rosson (1985) reported this effect of grapheme-phoneme correspondence frequency. In her experiments, words and nonwords that contained high-frequency grapheme-phoneme correspondences (her "strong rule" stimuli) were read aloud more quickly than words and nonwords that contained low-frequency grapheme-phoneme correspondences (her "weak rule"

stimuli). Additionally, as would be expected from the asymptotic learning assumption, she did not find this effect for high-frequency word stimuli (Experiment 2).

Covariant Learning and Lexical Coding

A mechanism of covariant learning can also accomplish direct access in the same way that it accomplishes phonological coding. That is, if we extend the previous assumptions about the association of phonological features (with orthographic features) to the association of other linguistic features, this mechanism of phonological coding becomes a mechanism of direct access. By this view, any linguistic features that frequently covary with orthographic features will become associated. The consequence of covariant learning for any subsequent instance of lexical coding will be that, initially, a representation of the spelling of a word will activate most strongly those linguistic features (i.e., semantic, syntactic, and phonological features) that covary to the highest degree with its orthographic features. This set of active linguistic features, the subset of linguistic features that are most likely to be functional for every occurrence of a particular word, is the lexical representation of that word. Notice that the covariant learning hypothesis implies that lexical codes are composed of morphophonological features, because, to various degrees, both morphological and phonological features covary with orthographic features.

I am reassured of the usefulness of this analysis by the fact that others have been led to similar conclusions, although their theoretical and empirical analyses converge from a variety of approaches to the problem of lexical representation (e.g., see Chomsky, 1970; Fowler, Napps, & Feldman, 1985; Lima & Pollatsek, 1983; Murrell & Morton, 1974; Snodgrass & Jarvella, 1972; Taft & Forster, 1975; Taft, 1979, 1981). I am also confident that the covariant learning hypothesis may be viably combined with the verification hypothesis, possibly after the fashion of Grossberg and Stone (1986). They described how verification can be accomplished using an adaptive resonant filter—a theoretical cousin to the delta rule.

REFERENCES

AARONSON, D., & FERRES, S. (1983). A model for coding lexical categories during reading. *Journal of Experimental Psychology: Human Perception & Performance*, 9, 700-725.

ALLPORT, D. A. (1977). On knowing the meaning of words that we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and performance IV* (pp. 505-533). New York: Academic Press.

ANDERSON, J. A. (1977). Neural models with cognitive implications. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 27-90). Hillsdale, NJ: Erlbaum.

ANDERSON, J. A. (1983). Cognitive and psychological computation with neural models. *IEEE Transactions on Systems, Man, & Cybernetics*, 13, 799-815.

ANDERSON, J. A., & MOZER, M. (1981). Categorization and selective neurons. In G. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory* (pp. 213-236). Hillsdale, NJ: Erlbaum.

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

BACKMAN, J., BRUCK, M., HEBERT, M., & SEIDENBERG, M. (1984).

Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology*, 38, 114-133.

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

BARON, J., & STRAWSON, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception & Performance*, 2, 386-393.

BARRON, R. W. (1980). Visual-orthographic and phonological strategies in reading and spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 195-213). London: Academic Press.

BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, 80(3, Pt. 2).

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

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

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

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

BESNER, D. (1984). Specialized processors subserving visual word recognition: Evidence for local control. *Canadian Journal of Psychology*, 38, 94-101.

BESNER, D., DAVELAAR, E., ALCOTT, D., & PARRY, P. (1984). Wholistic reading of alphabetic print: Evidence from the FDM and the FBI. In L. Henderson (Ed.), *Orthographies and reading: Perspectives from cognitive psychology, neuropsychology, and linguistics* (pp. 121-135). Hillsdale, NJ: Erlbaum.

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

BOWER, T. G. R. (1970). Reading by eye. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading* (pp. 134-146). New York: Basic Books.

CARPENTER, P. A., & JUST, M. A. (1981). Cognitive processes in reading: Models based on readers' eye fixations. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 177-213). Hillsdale, NJ: Erlbaum.

CHOMSKY, C. (1970). Reading, writing, and phonology. *Harvard Educational Review*, 40, 287-309.

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

COLTHEART, M., DAVELAAR, E., JONASSON, J. T., & BESNER, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). New York: Academic Press.

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

DANKS, J. H., & HILL, G. O. (1981). An interactive analysis of oral reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 131-153). Hillsdale, NJ: Erlbaum.

DAVELAAR, E., COLTHEART, M., BESNER, D., & JONASSON, J. T. (1978). Phonological recoding and lexical access. *Memory & Cognition*, 6, 391-402.

DOCTOR, E. A., & COLTHEART, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory & Cognition*, 8, 195-209.

ELLIS, A. W. (1984). *Reading, writing and dyslexia: A cognitive analysis*. Hillsdale, NJ: Erlbaum.

FOWLER, C. A., NAPPS, S. E., & FELDMAN, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory & Cognition*, 13, 241-255.

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

GLUSHKO, R. (1981). Principles for pronouncing print: The psychol-

ogy of phonology. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 61-84). Hillsdale, NJ: Erlbaum.

GOODMAN, K. S. (1969). Analysis of reading miscues: Applied psycholinguistics. *Reading Research Quarterly*, 5, 9-30.

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

GOUGH, P. B., & COSKY, M. J. (1977). One second of reading again. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory 2* (pp. 271-288). Hillsdale, NJ: Erlbaum.

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

GROSSBERG, S., & STONE, G. (1986). Neural dynamics of word recognition and recall: Priming, learning, and resonance. *Psychological Review*, 93, 46-74.

HAWKINS, H. L., REICHER, G. M., ROGERS, M., & PETERSON, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, 2, 380-385.

HENDERSON, L. (1982). *Orthography and word recognition in reading*. New York: Academic Press.

HENDERSON, L., & CHARD, J. (1980). The reader's implicit knowledge of orthographic structure. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 85-116). New York: Academic Press.

JOHNSTON, J. C. (1981). Understanding word perception: Clues from studying the word-superiority effect. In O. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 65-84). Hillsdale, NJ: Erlbaum.

JOHNSTON, J. C., & MCCLELLAND, J. L. (1980). Experimental tests of a hierarchical model of word identification. *Journal of Verbal Learning & Verbal Behavior*, 7, 560-572.

JORM, A. F., & SHARE, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics*, 4, 103-147.

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

KOHONEN, T. (1977). *Associative memory: A system-theoretical approach*. Berlin: Springer-Verlag.

KOHONEN, T. (1984). *Self-organization and associative memory*. Berlin: Springer-Verlag.

KOLERS, P. A. (1970). Three stages of reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading* (pp. 90-118). New York: Basic Books.

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

LEWICKI, P. (1986). Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 12, 135-146.

LIMA, S. D., & POLLATSEK, A. (1983). Lexical access via an orthographic code? The basic orthographic syllabic structure (BOSS) reconsidered. *Journal of Verbal Learning & Verbal Behavior*, 22, 310-332.

MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.

MCCLELLAND, J. L., & RUMELHART, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, 114, 159-188.

MCCUSKER, L. X., HILLINGER, M. L., & BIAS, R. G. (1981). Phonological recoding and reading. *Psychological Bulletin*, 89, 217-245.

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

MEYER, D. E., & GUTSCHERA, K. D. (1975, November). *Orthographic versus phonemic processing of printed words*. Paper presented at the meeting of The Psychonomic Society, Denver.

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

MORRIS, W. (Ed.). (1976). *The American Heritage dictionary of the English language*. Boston: Houghton Mifflin.

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

MURRELL, G. A., & MORTON, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102, 963-968.

NORRIS, D., & BROWN, G. (1985). Race models and analogy theories: A dead heat? *Cognition*, 20, 155-168.

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

PERFETTI, C. A. (1985). *Reading ability*. New York: Oxford University Press.

REICHER, G. M. (1969). Perceptual recognition as a function or meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275-280.

ROSENBERG, C. R., & SEJNOWSKI, T. J. (1986). *The effects of distributed vs. massed practice on NETtalk, a massively-parallel network that learns to read aloud*. Unpublished manuscript.

ROSSON, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory & Cognition*, 13, 90-99.

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

RUMELHART, D. E., HINTON, G. E., & MCCLELLAND, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 45-76). Cambridge, MA: MIT Press.

RUMELHART, D. E., HINTON, G. E., & WILLIAMS, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 318-362). Cambridge, MA: MIT Press.

RUMELHART, D. E., & MCCLELLAND, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89, 60-94.

RUMELHART, D. E., & MCCLELLAND, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Applications* (pp. 216-271). Cambridge, MA: MIT Press.

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

SEIDENBERG, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, 19, 1-30.

SEIDENBERG, M. S., WATERS, G. S., BARNES, M., & TANENHAUS, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, 23, 383-404.

SEJNOWSKI, T. J., & ROSENBERG, C. R. (1986). *NETtalk: A parallel network that learns to read aloud* (Tech. Rep. No. JHU/EECS-86/01). Baltimore, MD: The Johns Hopkins University, Electrical Engineering and Computer Science.

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

SNODGRASS, J. G., & JARVELLA, R. J. (1972). Some linguistic determinants of word classification time. *Psychonomic Science*, 27, 220-222.

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

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

STONE, G. O. (1986). An analysis of the delta rule and the learning of statistical associations. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 444-459). Cambridge, MA: MIT Press.

SUTTON, R. S., & BARTO, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, **88**, 135-170.

TAFT, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7, 263-272.

TAFT, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning & Verbal Behavior*, 20, 289-297.

TAFT, M., & FORSTER, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning & Behavior*, 14, 638-647.

UNDERWOOD, G., & THWAITES, S. (1982). Automatic phonological coding of attended printed words. *Memory & Cognition*, 10, 434-442.

UYEDA, K. M., & MANDLER, G. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods & Instrumentation*, 12, 587-595.

VAN ORDEN, G. C. (1984). *A ROWS is a /ROZ/ is a ROSE: Visual, phonological, developmental and strategic influences upon lexical access*. Unpublished doctoral dissertation, University of California at San Diego.

VAN ORDEN, G. C., JOHNSTON, J. C., & HALE, B. L. (in press). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.

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

WATERS, G. S., SEIDENBERG, M. S., & BRUCK, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory & Cognition*, 12, 293-305.

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

WHEELER, D. (1970). Processes in word recognition. *Cognitive Psychology*, 1, 59-85.

WINER, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

NOTES

1. The error analyses in all the experiments reported here were also computed using the arcsin transformation of the proportion of errors (Winer, 1971), and virtually identical results were found.

2. A significant partial correlation ($pr = .48, p < .05$) was found between a rough estimate of spelling controls' phonological similarity (PS) to their corresponding category exemplars and the false-positive error rates to these spelling controls (in the pattern-masking conditions of Experiment 2) that is independent of the effect of OS, the estimate of orthographic similarity. In contrast, the complementary partial correlation between OS and false-positive error rates (holding PS constant) was not significant ($pr = .27, p > .10$). PS was simply the difference between the number of phonemes in a spelling control that were also in its corresponding category exemplar and the number of phonemes in that spelling control that were not in the corresponding category exemplar.

3. There were no false positive categorization errors to the homophone foil TURN (for the category A BIRD). Its 0% error rate differed by over two standard deviations ($SD = 20.7$) from the mean error rate of 44% for this frequency group. All other homophone foils in Group VL (by the frequency of TERN) yielded error rates of 50% or greater. To find the source of this difference between TURN and the other foils in its frequency group, I conducted a "usage test" in which 32 University of California at San Diego undergraduates attempted to use category exemplars such as TERN and ROSE in a sentence. The closest competitor for "least known category exemplar" was used correctly by 72% of the students. If the lexical entry for TERN is not readily available, as it may not be if most subjects do not even know that a TERN is A BIRD, then subjects are unlikely to falsely categorize the foil TURN as A BIRD. Consequently, the data from trials in which target word TURN and its yoked control word appeared were discarded from the analysis of the effect of category exemplar (TERN's) frequency. (These data were not discarded from the analysis of the effect of stimu-

lus foil frequency because the very high-frequency TURN's low error rate is consistent with the bypass hypothesis.)

APPENDIX A

The experiments reported here required a control for the similarity in spelling that is common between pairs of homophones. For this purpose, I adapted an estimate from Weber (1970). Weber's measure of graphic similarity (GS) is computed as follows:

$$GS = 10([(50F + 30V + 10C)/A] + 5T + 27B + 18E)$$

F = number of pairs of adjacent letters in the same order shared by word pairs:

HOUSE/HORSE   F = 2

EVERY/VERY   F = 3

V = number of pairs of adjacent letters in reverse order shared by word pairs:

WAS/SAW   V = 2

C = number of single letters shared by word pairs:

SPOT/PUFF   C = 1

FAMILY/FUNNY   C = 2

A = average number of letters in the two words:

EVERY/VERY   A = 4.5

T = ratio of number of letters in the shorter word to the number in the longer:

EVERY/VERY   T = 4/5

B = 1 if the first letter in the two words is the same; otherwise, B = 0.

E = 1 if the last letter in the two words is the same; otherwise, E = 0.

Weber's (1970) method of measuring GS implies differences in how similar words are to themselves. For example, a comparison of GS computed between MEET and itself with GS computed between MEAT and itself indicates that MEET is spelled more closely to MEET (GS = 1050) than MEAT is to MEAT (GS = 975). I assumed that all words are equally similar in spelling to themselves. To accommodate this assumption, the index of orthographic similarity used here, OS, was defined by the following ratio:

$$OS = (GS \text{ of target foil and category exemplar})/$$
$$(GS \text{ of category exemplar and itself})$$

For example, OS between MEET and MEAT is:

$$OS = 10\{[50(1) + 30(0) + 10(3)]/4 + 5(1) + 27(1) + 18(1)\}/$$
$$10\{[50(3) + 30(0) + 10(4)]/4 + 5(1) + +27(1) 18(1)\}$$
$$= 700/975$$
$$= .72$$

APPENDIX B
The Similarly Spelled and Less Similarly Spelled Homophone Foils and Their Respective Yoked Spelling Control Foils Used in Experiments 1 and 2

Category Name	Homophone	Spelling Control
Similarly Spelled Foils		
A FEATURE OF A PERSON'S ABDOMEN	NAVAL	NOVEL
A FEATURE OF AN OCEAN SHORE	BEECH	BENCH
A SMALL STREAM	CREAK	CHEEK
ORGANIZED GROUP OF PEOPLE	TEEM	TERM
PART OF A DRESS	SEEM	SLAM
PART OF A HORSE'S BRIDLE	RAIN	RUIN
PART OF A MOUNTAIN	PEEK	PECK
TYPE OF FOOD	MEET	MELT
A KITCHEN UTENSIL	BOLL	BOIL
A BIBLICAL RELIGIOUS LEADER	PROFIT	PROTEST
Less Similarly Spelled Foils		
A DEER	DOUGH	DOUBT
PART OF A PERSON'S FACE	KNOWS	SNOBS
TYPE OF HOTEL ROOM	SWEET	SHEET
A SERVANT	MADE	MAIN
A MEMBER OF A CONVENT	NONE	NINE
AN ANCIENT MUSICAL INSTRUMENT	LOOT	LOST
A FLOWER	ROWS	ROBS
SOMETHING CAUSED BY GRAVITY	WAIT	WRIT
A WILD ANIMAL	BORE	BORN
A BREAKFAST FOOD	SERIAL	VERBAL

APPENDIX C
The Homophone and Control Foils from Experiment 3 Arranged in Frequency Groups Using the Frequency Counts of Corresponding Sound-Alike Category Exemplars (Upper Portion) and the Stimulus Homophone Foils (Lower Portion)

Frequency Group	Category Name	Homophone	Control
Corresponding Sound-Alike Category Exemplars			
VL	PART OF A HORSE'S BODY	MAIN	DEEP
	AN EDIBLE PLANT	LEAK	HIVE
	A BIRD	TURN	FEEL
	A VEGETABLE	BEAT	FLOW
	A PERSON OF MEDIEVAL TIMES	SURF	TINT
L	A WILD ANIMAL	BORE	JUMP
	PART OF A WOMAN'S DRESS	SEEM	PLAY
	A PLACE TO SHOP	MAUL	ROTE
	AN ASSESSMENT OF OPINION	POLE	BELL
	PART OF A HORSE'S HARNESS	RAIN	DUST
H	A KITCHEN UTENSIL	BOLL	DUPE
	PART OF A MOUNTAIN	PEEK	WANE
	A GROUP OF PEOPLE	TEEM	SCOW
	A TYPE OF FOOD	MEET	FALL
	A SERVANT	MADE	LIKE
VH	A UNIT OF TIME	WEAK	CURT
	PART OF THE HUMAN BODY	FEAT	WOLF
	A SPHERICAL OBJECT	BAWL	KINK
	PART OF A HOUSE	HAUL	CHAT
	PART OF THE HUMAN BODY	HARE	CUBE

APPENDIX C (Continued)

Stimulus Homophone Foils

VL	A PLACE TO SHOP	MAUL	ROTE
	A KITCHEN UTENSIL	BOLL	DUPE
	PART OF A MOUNTAIN	PEEK	WANE
	A GROUP OF PEOPLE	TEEM	SCOW
	A SPHERICAL OBJECT	BAWL	KINK
L	AN EDIBLE PLANT	LEAK	HIVE
	A PERSON OF MEDIEVAL TIMES	SURF	TINT
	PART OF THE HUMAN BODY	FEAT	WOLF
	PART OF A HOUSE	HAUL	CHAT
	PART OF THE HUMAN BODY	HARE	CUBE
H	A VEGETABLE	BEAT	FLOW
	A WILD ANIMAL	BORE	JUMP
	AN ASSESSMENT OF OPINION	POLE	BELL
	PART OF A HORSE'S HARNESS	RAIN	DUST
	A UNIT OF TIME	WEAK	CURT
VH	PART OF A HORSE'S BODY	MAIN	DEEP
	A BIRD	TURN	FEEL
	PART OF A WOMAN'S DRESS	SEEM	PLAY
	A TYPE OF FOOD	MEET	FALL
	A SERVANT	MADE	LIKE