

Effects of memory load in a word-naming task: Five failures to replicate

PENNY M. PEXMAN and STEPHEN J. LUPKER
University of Western Ontario, London, Ontario, Canada

Paap and Noel's (1991) recent demonstration that increased memory load facilitates naming of low-frequency irregular words has been taken as strong support for dual-route theories of word naming (e.g., Coltheart, 1978). Others, however, (Jared, personal communication, March 1993; Strain, personal communication, April 1993) have been unable to replicate this effect. In the present research, five different attempts were made to determine the crucial experimental conditions required for successful replication of the Paap and Noel findings. None of these experiments produced results at all similar to those reported by Paap and Noel. A number of explanations for these failures to replicate, in terms of individual differences between Paap and Noel's subjects and the present subjects, were evaluated. Most of these explanations were not supported by the data. The present results call into question the generalizability of the Paap and Noel findings and, hence, the strength of support they provide for dual-route theories.

In recent decades, much research effort has been devoted to understanding the processes involved in reading. In particular, the component of reading referred to as *word recognition* has been extensively examined. In this work, the researchers' main goal has been to disambiguate the processes involved in deriving pronunciation or meaning from a letter string, rather than trying to understand higher level comprehension or integration processes.

The focus of the present research is one aspect of the word recognition process: word naming. Several theories of word naming have been developed, each based on a different conceptualization of the way in which phonological information is derived. These theories can generally be classified as either dual-route or single-route theories. Dual-route theories posit that orthographic information is used in either of two ways to derive a word's name (e.g., Coltheart, 1978; Paap, McDonald, Schvaneveldt, & Noel, 1987; Paap, Newsome, McDonald, & Schvaneveldt, 1982). Single-route theories hold that there is essentially only one way to derive phonological information (e.g., Glushko, 1979; McClelland & Rumelhart, 1981; Morton, 1969; Seidenberg & McClelland, 1989; Taraban & McClelland, 1987; Van Orden, Pennington, & Stone, 1990), although some of these models also acknowledged the possible

existence of a second route (e.g., Seidenberg & McClelland, 1989).

Following a brief introduction to these theories, some recent results that appear to provide strong support for dual-route theories (Paap & Noel, 1991) will be discussed. These results are the starting point for the present investigation.

Dual-Route Theories

Perhaps the most influential theory of word naming has been the dual-route model originally proposed by Coltheart (1978) and more recently expanded on with the activation-verification model of Paap and colleagues (Paap et al., 1987; Paap et al., 1982). According to these models, one route (the "assembly" route) involves readers using knowledge of subword spelling-to-sound relationships in their language to derive a set of possible sounds for subword units, which are then assembled into a complete phonological code. This route is assumed to play a major role in the naming of novel letter strings as well as words that follow the normal spelling-to-sound rules (the "regular" words). The second route (the "lexical" route) involves the direct access of a word's representation in the lexicon followed by the essentially holistic retrieval of the word's phonological code. Since not all English words follow normal spelling-to-sound relationships, those words (the "irregular" words) can only be pronounced by using information from the lexical route.

One of the major reasons for proposing the existence of the two separate routes was that results showed an apparent processing difference between regular and irregular words. For example, Baron and Strawson (1976) showed that subjects read aloud a list of regular words (e.g., *sweet*) faster than they read a list of irregular words (e.g., *sword*), an effect referred to as the *regularity effect*. Further study has revealed that the regularity effect

This research was supported by a Graduate Fellowship to the first author and by Grant A6333 to the second author, both from the Natural Sciences and Engineering Research Council of Canada. A major portion of this paper was presented at the Third Meeting of the Canadian Society for Brain Behaviour and Cognitive Science, Toronto, Canada, July 1993. The authors would like to thank Patrick Brown, Chris Herdman, Debra Jared, Albert Katz, and Ken Paap for their many helpful discussions about both theoretical and methodological issues. Also, two anonymous reviewers provided valuable suggestions regarding an earlier draft of this paper. Correspondence concerning this paper can be addressed to either author, at the Department of Psychology, University of Western Ontario, London, ON, Canada N6A 5C2.

interacts with word frequency (Seidenberg, Waters, Barnes, & Tanenhaus, 1984) in that regular words appear to be named more quickly than irregular words only when the words are of low frequency.

The nature of this interaction of regularity and frequency can be easily explained by the dual-route model if three assumptions are made: (1) that the two routes run in parallel, (2) that speed of processing on the lexical route is a function of word frequency, and (3) that naming times can be affected by any inconsistencies in the outputs of the two routes. For example, with respect to this final assumption, Paap and Noel (1991) describe processing by the two routes as a "horse race," where the faster route is declared the winner and its pronunciation is used. In cases where the two routes produce different pronunciations at approximately the same time, the lexical route tends to be favored, but time is taken up in resolving the competition between routes.

On the basis of these assumptions, the specific form of the interaction can be explained as follows. Low-frequency irregular words are processed somewhat slowly via the lexical route and, thus, competition that must be resolved is often created. Low-frequency regular words are pronounced more quickly than are low-frequency irregular words because either route can provide the correct phonological code. Thus, there is no competition, and the reader can use whatever code is available first. For high-frequency words, the regularity effect does not exist because the lexical route produces the correct pronunciation for all high-frequency words so rapidly that the assembly route can play little role. That is, for high-frequency irregular words, the lexical route yields the correct pronunciation before any competing output is available from the assembly route. Thus, competition will almost never arise.

The existence and nature of regularity effects are, of course, not the only data that provide support for dual-route theories (see Coltheart, Curtis, Atkins & Haller, 1993, Humphreys & Evett, 1985, Patterson & Coltheart, 1987, and Van Orden et al., 1990, for discussions of the current state of empirical support for the dual-route model). As will be discussed below, however, these results form the basis of the Paap and Noel (1991) analysis, which is the specific focus of the present paper.

Single-Route Theories

The predominant challenge to dual-route theories has come from models that postulate that essentially only a single route is necessary for pronunciation of virtually all letter strings. While early versions of this type of model experienced difficulty explaining how irregular words could be named (Gough, 1972) or how a pronunciation could ultimately be synthesized, more recent versions (Brown & Besner, 1987; Seidenberg & McClelland, 1989; Van Orden et al., 1990) appear to be able to explain the naming of all types of letter strings (but see Besner, Twilley, McCann, & Seergobin, 1990, for a contrary opinion).

In the single-route model described by Van Orden et al. (1990), phonological coding is the central process. According to the model, connections between orthographic and phonological units are learned implicitly through experience with different correspondences (the covariant learning hypothesis). The result is that what is created is a "statistical version of regularity in which regularity effects emerge as a function of the relative consistency and frequency with which spelling and phonology covary" (Van Orden et al., 1990, p. 490).

Van Orden et al. describe the word-naming process as a subsymbolic performance that takes place in "phonologic space." A reader's initial encoding of a word is one point in phonologic space, whereas the word's correct phonological code is another point in that space, possibly some distance from the initial encoding. The point that represents the correct code is called the *attractor*. In order to name the word, the initial encoding must be "cleaned up" until it matches the attractor. The distance between the initial encoding and the attractor is the primary determinant of naming latency. Thus, low-frequency words take longer to name than do high-frequency words because they are initially encoded at a greater distance and they require more "cleaning up."

In Seidenberg and McClelland's (1989) model, the process of word recognition is also modeled by the creation of activation patterns in sets of orthographic and phonological units. Between these units, there is a set of hidden units. Connections between units are weighted, and these weights are adjusted to reflect spelling-to-sound correspondences during training by a back-propagation learning algorithm.

Seidenberg and McClelland have shown that the effects of frequency and regularity can be accounted for in terms of ultimate connection weights between units. That is, high-frequency words are named faster than are low-frequency words because words that are seen more frequently have had their connection weights adjusted in the right direction more often. Regular words are named more quickly than are irregular words because they share connections with other orthographically and phonologically similar words. Thus, they are activated more readily than are words with fewer connections. Further simulations using the model demonstrate that it also predicts the frequency \times regularity interaction first reported by Seidenberg et al. (1984).

These types of models represent a sharp departure from other theories of word recognition in terms of mental representation. Previous theories had assumed that words were represented locally in a mental lexicon, whereas the premise of the parallel distributed processing (PDP) models is that information about words is distributed across a network. That is, the PDP models presume that there is no lexicon with specific entries for individual words. More importantly, they deny the existence of separate lexical and assembly routes for spelling-to-sound translation while, at the same time, they appear to be able

to account for most of the data taken as support for dual-route theories.

Paap and Noel (1991)

The focus of the present paper is new support for the dual-route model offered recently by Paap and Noel (1991). The premise of the Paap and Noel experiment was that the lexical and assembly processes for word naming have different attentional demands. That is, both routes are assumed to require attentional resources in order to derive pronunciations, but processing on the assembly route is believed to require more resources for operation because it is an active process involving assembly of phonological codes whereas the lexical route is held to be more automatic, therefore requiring fewer resources.

On the basis of these different attentional requirements, Paap and Noel (1991) made several hypotheses about the effects of memory load on the traditional frequency \times regularity interaction. For example, according to dual-route theories, low-frequency irregular words are named more slowly than are low-frequency regular words because the assembly route and the lexical route produce different, competing pronunciations for low-frequency irregular words. Paap and Noel (1991) hypothesized that under a high memory load, where attentional resources would be limited, the assembly route would slow down relatively more than would the lexical route. As such, naming would mainly be driven by the lexical route. This would tend to eliminate competing phonological codes coming from the assembly route, and the low-frequency irregular words may actually be named more quickly under a high memory load than under a low memory load.

Similarly, according to dual-route theories, low-frequency regular words are often correctly named by the assembly route because of their regularity. Under high memory load, the assembly route would only very rarely provide a phonological code for these words, and, therefore, they would generally be named by the lexical route. As the lexical route would be slow to produce a response due to the low frequency of the words, naming latencies for these low-frequency regular words would be substantially slowed. In essence, the expectation is that the naming latencies for low-frequency regular and irregular words should "come together" under high memory load because the lexical route was providing the phonological code, with little or no competition, for both types of words.

The high memory load would, at the same time, create a large frequency effect for both regular and irregular words. Under normal conditions, the frequency effect for regular words is not large because the assembly route often beats the lexical route in providing the code for the low-frequency words. With a high memory load, virtually all words would be named via the lexical route. Thus, a frequency effect should be produced for regular words that is similar in magnitude to the frequency effect that normally exists for irregular words. These predictions were confirmed exactly by the results reported by

Paap and Noel (1991), and they were taken by the authors as strong new evidence supporting dual-route theories.

Before proceeding, three points should be noted about this argument. First, the Paap and Noel (1991) dual-route analysis does not require that the low-frequency irregular words actually be named faster under high memory load than under low memory load. That is, although the high memory load slows the assembly route, it also slows the lexical route. Thus, there are two counteracting effects at work, and the ultimate effect of a high memory load could be facilitation, inhibition, or no change. The more important aspect of the predictions (and results) is that because the high memory load ties up the assembly route, the regularity effect for low-frequency words should diminish and the frequency effect for regular words should increase.

The second point is that even if Paap and Noel (1991) had failed to find these effects, their results would not have discredited dual-route models. The predictions were based on some assumptions that Paap and Noel made, assumptions that the basic dual-route model does not require—in particular, assumptions concerning the attentional demands of the two routes and the memory task. The fact that the results did come out as predicted, however, does appear to indicate that, within the dual-route framework, those assumptions are valid.

The third point is, as Paap and Noel (1991) argue, there is no obvious way that the Seidenberg and McClelland (1989) single-route model could account for the Paap and Noel data. That is not to say, however, that these results could not be explained within a single-route framework. We will return to this point in the General Discussion.

Since publication of the Paap and Noel results, however, researchers have only had limited success replicating the Paap and Noel findings. For example, Bernstein and Carr (1991) demonstrated some facilitation of naming times for low-frequency irregular words under high memory load and a very slight inhibition effect of naming times for low-frequency regular words under the high load. Thus, in the high-load condition, the naming latencies for low-frequency regular and irregular words did come together somewhat. The frequency effect for regular words did not increase in the high-load condition, however, which is an effect that seems to be required for arguing that, under the high memory load, naming is accomplished mainly via the lexical route.

Research that more closely replicates the Paap and Noel (1991) memory load effects is that of Herdman, Beckett, and Stolpmann (1993). Herdman et al. did find that, in their high-memory-load condition, the frequency effect increased for regular words whereas the regularity effect for low-frequency words decreased. Other researchers, however, (Jared, personal communication, March 1993; Strain, personal communication, April 1993) have failed to produce any results similar to those of Paap and Noel. Furthermore, research by Lukatela and Turvey (1993) produced findings that can only be reconciled with the Paap and Noel data through a com-

plex theory based on thresholds. This lack of consistent replication is troubling when one considers the potential significance of the Paap and Noel findings for the debate between dual-route and single-route theories of word naming.

There are several possible reasons why replication of the Paap and Noel (1991) effect has been so difficult. The most likely reason is methodology. That is, the researchers who have failed to replicate Paap and Noel may have used slightly different methodologies than were used in the original study. If these methodological differences are inconsistent with some of the basic processing assumptions made by Paap and Noel, then replication would not necessarily follow. For instance, it seems crucial for the Paap and Noel effect that subjects consider the memory task to be primary and the naming task secondary. If this were not stressed in the instructions, the result may be a failure to replicate. Similarly, if the digits of the memory task were not encoded by all subjects in a particular way, attentional demands might be somewhat different. That is, some subjects might treat the task as a spatial exercise, where they imaged the digits in order to remember them rather than coding them verbally. If this were the case, the task might not affect the resources normally available for reading, and, thus, naming-task results would not show a reliance on lexical processing. Also, if the memory task were not sufficiently difficult for certain subjects, then they would not find the task as demanding of attentional resources as the subjects in the Paap and Noel study did. If the task were too easy, the contributions of the assembly route would not necessarily be insignificant under the high memory load. This could also lead to a failure to replicate.

The purpose of the present research was to attempt to resolve contradictory findings concerning replication of the Paap and Noel (1991) effects by testing the relevance of the methodological issues described above. Experiment 1A was an exact replication of the Paap and Noel methodology. In Experiment 1B, the instructions were altered slightly to request subjects to use a verbal coding strategy for the digits on the memory task. In Experiment 1C, the importance of making an accurate response on the memory task was emphasized.

These experiments were initial attempts to address the question, "What are the circumstances under which the Paap and Noel results arise?"

EXPERIMENT 1A

Method

Subjects. The subjects in Experiment 1A were 28 undergraduate students at the University of Western Ontario. Average age of the subjects was 19.32 years. In each of these experiments, all subjects considered English to be their native language and had normal or corrected-to-normal vision. Also, unless stated otherwise, all of our subjects received partial course credit in an introductory psychology course for their participation.

Apparatus. An IBM PC computer was used in this experiment to control the presentation of stimuli. Stimuli were presented on a

Packard Bell Enhanced Color Graphic Display monitor (Model No. PB1422EG), where letters and digits were approximately 0.60 cm high and at eye level for subjects. The distance between each subject and the monitor screen was approximately 50 cm. Naming response times were recorded by a Shure Inc. (Model No. 5755) microphone connected to a Ralph Gerbrands Co. (Model No. 800) electronic voice key relay, which was triggered by vocal responses. The subjects were instructed to sit approximately 20 cm from the microphone. The "shift" keys on a standard keyboard were used to make responses in the memory task. These keys were labeled YES and NO, such that a subject's dominant hand was assigned to the YES key.

Stimuli. The stimuli used in this experiment were the same 80 words used in the original Paap and Noel (1991) experiment. These words are presented in Appendix A. Twenty of these words were low-frequency irregular words, 20 were high-frequency irregular words, 20 were low-frequency regular words, and 20 were high-frequency regular words. The regularity/consistency of these words was evaluated by Paap and Noel, and their paper should be consulted for details.

For each type of word, mean frequencies were calculated on the basis of the Kučera and Francis (1967) norms. The low-frequency regular and irregular words had mean frequencies of 3.5. The high-frequency irregular words had a mean frequency of 393. The high-frequency regular words had a mean frequency of 214. In creating the original word list, Paap and Noel had matched each irregular word with a regular word for approximate frequency, length, and initial phoneme.

Procedure. As in the Paap and Noel study, half of the words in each of the frequency \times regularity cells were presented to subjects in the high-memory-load condition, and the other half were presented in the low-memory-load condition. Thus, two groups of subjects were needed to counterbalance assignment of words to load conditions. The subjects were assigned to these groups by order of participation in the study such that odd-numbered subjects were assigned to one group and even-numbered subjects were assigned to the other group. There were 14 subjects in each group.

The subjects were tested individually. The testing session lasted approximately 25 min. The testing room was normally lit. The subjects were instructed that this experiment involved both a naming response and a memory task. They were asked to remember that the memory task was primary and the naming task was secondary. The subjects were asked to make their responses as quickly and as accurately as possible. As in the Paap and Noel study, they completed 20 practice trials before starting the 80 trials that composed the experiment.

Each experimental trial included both the memory task and the naming task. The sequence of each trial was that, initially, a 50-msec 400-Hz beep signal was heard and then a fixation point appeared on the screen for 1 sec. When the fixation point disappeared, either one digit or five digits would appear on the screen for 2 sec. The digits for each trial were randomly selected and presented in random order. The monitor was then blank for a variable interval of either 1 or 2 sec. After this interval, the word to be named appeared on the screen. When the voice key was activated by the response, the word disappeared. Six seconds after the start of the trial, one digit appeared on the screen. The subjects were instructed to make a yes or no response to this digit, according to whether the digit was one of the digits that appeared in the set displayed at the start of the trial. The yes/no response was made by pressing one of the two shift keys on the keyboard. The experimenter sat behind the subject and recorded any naming errors that were made.

After each practice trial, the subjects were given feedback about whether their response to the memory task was correct and also about their response times for the memory task and the naming task. If the subjects made errors during practice, they were given additional instructions.

Results and Discussion

Naming task. For naming responses, a trial was scored as an error if the response time was longer than 1,400 msec or shorter than 250 msec or if the word was mispronounced. Thirty-one percent of the mispronunciations were regularization errors. For instance, the word *sew* was sometimes (19% of trials) pronounced like *few*. The mispronunciations seemed mainly to be due to the subjects being unfamiliar with stimulus words. For instance, the word *wily* seemed to elicit frequent mispronunciations (93% of trials) because the subjects did not know the word and often pronounced it as *willie*. Trials that were scored as naming errors were not included in the analyses of naming latencies. Mean naming latencies are depicted in Figure 1. Figure 1 also includes the data from Paap and Noel (1991) to allow readers to compare the two sets of results.

In this and all subsequent experiments, the data were examined with analyses in which subjects and items were separately treated as random factors. The analyses indicated that the three-way interaction of frequency, regularity, and memory load was nonsignificant [$F_s(1,27) = 2.62$, $MS_e = 1,264.77$, n.s.; $F_i(1,76) = 1.53$, $MS_e = 2,623.50$, n.s.] in both analyses. Significant in both analyses were the effects of frequency [$F_s(1,27) = 65.20$, $MS_e = 2,355.39$, $p < .001$; $F_i(1,76) = 24.17$, $MS_e = 3,252.54$, $p < .001$] and regularity [$F_s(1,27) = 81.69$, $MS_e = 1,003.48$, $p < .001$; $F_i(1,76) = 10.65$, $MS_e = 3,252.54$, $p < .001$], indicating that words were named more quickly when they were of high frequency and/or regular. The interaction of these two factors was significant in the subjects' analysis and was marginally significant in the items' analysis [$F_s(1,27) = 26.12$, $MS_e =$

710.51, $p < .001$; $F_i(1,76) = 2.76$, $MS_e = 3,252.54$, $p < .10$]. This interaction had the standard form where the regularity effect was larger for low-frequency words than for high-frequency words. Results in the subjects' analysis also included a significant effect of memory load, but this effect was only marginal in the items' analysis [$F_s(1,27) = 6.23$, $MS_e = 1,656.47$, $p < .05$; $F_i(1,76) = 3.03$, $MS_e = 2,623.50$, $p < .10$]. The nature of this effect was that the subjects tended to name words more slowly in the high-memory-load condition than in the low-memory-load condition.

As Figure 1 indicates, the pattern of response times for the subjects in Experiment 1A does not resemble the effects observed by Paap and Noel (1991), where low-frequency irregular words were named more quickly under high memory load whereas low-frequency regular words were named more slowly. In fact, if anything, the difference between the low-frequency regular and irregular words actually seemed to increase in the high-load condition, whereas the difference between the low- and high-frequency regular words seemed to shrink.

The mean naming time across all conditions in the Paap and Noel (1991) experiment was approximately 683 msec. In the present experiment, the mean naming time was 641 msec. Thus, the subjects were generally responding more quickly in the present experiment. Given that the present subjects were faster than those of Paap and Noel, a possible reason for the failure to replicate could be that these effects are only manifest with slower subjects. To evaluate this idea, the results for the slowest half of the subjects in each counterbalancing condition (overall mean response time = 693 msec) were calculated. The mean naming latencies for these 14 subjects were 757 and 771 msec for low-frequency irregular words under low and high memory load, respectively, 684 and 692 msec for low-frequency regular words under low and high memory load, 668 msec for high-frequency irregular words under both low and high load, and 637 and 666 msec for high-frequency regular words under low and high memory load. As these results indicate, the naming latencies for the slower subjects also did not show the pattern of facilitation and inhibition found by Paap and Noel. In particular, the latencies for the low-frequency irregular words increased rather than decreased as a function of increased memory load such that the regularity effect for low-frequency words increased slightly. Furthermore, the frequency effect for regular words decreased slightly.

Over all subjects, naming errors occurred on 5.77% of trials. The only significant effects were the main effects of frequency [$F_s(1,27) = 47.29$, $MS_e = 51.88$, $p < .001$; $F_i(1,76) = 6.04$, $MS_e = 374.20$, $p < .05$] and regularity [$F_s(1,27) = 140.97$, $MS_e = 37.18$, $p < .001$; $F_i(1,76) = 12.06$, $MS_e = 374.20$, $p < .001$] and the regularity \times frequency interaction [$F_s(1,27) = 90.44$, $MS_e = 31.92$, $p < .001$; $F_i(1,76) = 7.09$, $MS_e = 374.20$, $p < .01$]. This interaction also took the standard form where, for low-frequency words, many more errors were made for irregular words (17.17%) than for regular words (0.73%),

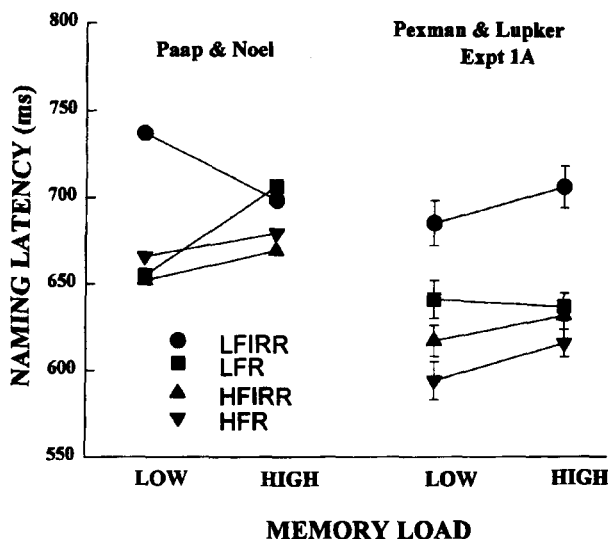


Figure 1. Mean naming latencies (in milliseconds) and standard error bars for subjects in the Paap and Noel (1991) experiment relative to those for subjects in the present (Pexman & Lupker) Experiment 1A, for low-frequency irregular words (LFIRR), low-frequency regular words (LFR), high-frequency irregular words (HFIRR), and high-frequency regular words (HFR), as a function of memory load.

whereas for high-frequency words, the difference in error rates was much smaller between irregular (3.83%) and regular words (1.35%).

Memory task. Both response times and error rates for the memory task were analyzed. In this and all subsequent experiments, both subjects' and items' analyses were carried out. Analyses of these response times showed that there was a significant main effect of memory load [$F_s(1,27) = 56.92, MS_e = 44354.38, p < .001; F_i(1,76) = 127.01, MS_e = 14,198.31, p < .001$], where response times were slower in the five-digit memory-load condition ($M = 843$ msec) than in the one-digit memory-load condition ($M = 630$ msec). In addition, the type of word presented on the naming task did have an effect on memory task response times, with a significant effect of frequency [$F_s(1,27) = 13.42, MS_e = 9,399.94, p < .001; F_i(1,76) = 5.93, MS_e = 15,193.71, p < .05$], such that memory task response times were slower when the stimulus words were of low frequency ($M = 760$ msec) than when the stimulus words were of high frequency ($M = 713$ msec). Furthermore, there was an interaction of frequency and regularity [$F_s(1,27) = 9.72, MS_e = 10,537.21, p < .005; F_i(1,76) = 4.82, MS_e = 15,193.71, p < .05$], such that memory response times were particularly slowed by low-frequency irregular words ($M = 793$ msec) relative to high-frequency irregular words ($M = 703$ msec), low-frequency regular words ($M = 727$ msec), and high-frequency regular words ($M = 723$ msec).

A trial was scored as an error on the memory task only if the subject pressed the wrong key when responding to the probe digit. Over all subjects, 5.27% of memory task responses were incorrect. There is only minimal evidence that this percentage was significantly affected by memory load [$F_s(1,27) = 3.04, MS_e = 58.67, p < .10; F_i(1,76) = 1.66, MS_e = 35.54, n.s.$], although the error rate was smaller in the one-digit memory-load condition (4.45%) than in the five-digit memory-load condition (6.09%). The main effect of frequency was significant [$F_s(1,27) = 14.25, MS_e = 34.64, p < .001; F_i(1,76) = 8.90, MS_e = 43.38, p < .001$] because the subjects made more errors on the memory task on trials in which a low-frequency word was presented (6.80%) than on trials in which a high-frequency word was presented (3.74%). These results are somewhat different from those in the Paap and Noel study. Paap and Noel found that significantly more errors were made in the high-memory-load condition, and they did not find that the type of word presented in the naming task affected memory task performance. No other effects approached significance (all $ps > .10$).

The most important result of Experiment 1A was that the data pattern reported by Paap and Noel (1991) was not replicated. That is, Paap and Noel found that under high memory load, low-frequency irregular words were named much faster and low-frequency regular words were named much slower than under low memory load, essentially eliminating the regularity effect for low-frequency words. In contrast, the results of the present experiment suggest that an increased memory load de-

lays naming of all types of words (with the possible exception of low-frequency regular words).

One possible explanation of these results is that the subjects in Experiment 1A used a different strategy for encoding the five-digit set from that used by Paap and Noel's subjects. For example, it is possible that the present subjects may have encoded the digits in a more visual format than did the subjects in the Paap and Noel study, a format that may have made the high memory load much less attention-demanding. In Experiment 1B, we attempted to avoid use of this type of strategy by altering the task instructions and encouraging the subjects to store the digits in a verbal format.

EXPERIMENT 1B

Method

Subjects. The subjects for Experiment 1B were 20 undergraduates at the University of Western Ontario. None of these subjects had participated in Experiment 1A. Average age of the subjects was 20.30 years.

Apparatus, Stimuli, and Procedure. The apparatus, stimuli, and procedure for Experiment 1B were the same as those used in Experiment 1A, except that the instructions to the subjects were altered slightly—that is, prior to the practice trials, the participants were told to memorize digits on each trial by rehearsing them silently. This instruction was reiterated before the experimental trials began.

Results and Discussion

Naming task. Error criteria were the same as in Experiment 1A. Thirty-six percent of the mispronunciations were regularization errors. The remainder again tended to be due to the subjects being unfamiliar with the word's pronunciation. Mean naming latencies are depicted in Figure 2.

The three-way interaction of frequency, regularity, and memory load was again nonsignificant [$F_s(1,19) = .41, MS_e = 2,233.31, n.s.; F_i(1,76) = .55, MS_e = 3,516.96, n.s.$]. The only significant main effects were the effects of frequency [$F_s(1,19) = 8.92, MS_e = 3,509.05, p < .01; F_i(1,76) = 5.79, MS_e = 8,860.01, p < .05$] and regularity [$F_s(1,19) = 4.68, MS_e = 4,072.09, p < .05; F_i(1,76) = 4.62, MS_e = 8,860.01, p < .05$]. The two-way interaction of frequency and regularity was again significant in the subjects' analysis but not in the items' analysis [$F_s(1,19) = 4.70, MS_e = 2,195.68, p < .05; F_i(1,76) = 1.30, MS_e = 8,860.01, n.s.$]. Most importantly, as seen in Figure 2, under high memory load, the mean naming times for Experiment 1B do not show the same pattern as was observed in the Paap and Noel study.

The mean error rate in the naming task was 3.97%. As in Experiment 1A, the only significant effects were the main effects of frequency [$F_s(1,19) = 46.11, MS_e = 31.83, p < .001; F_i(1,76) = 6.62, MS_e = 197.08, p < .01$] and regularity [$F_s(1,19) = 51.68, MS_e = 32.20, p < .001; F_i(1,76) = 8.06, MS_e = 197.08, p < .01$] and the frequency \times regularity interaction [$F_s(1,19) = 25.57, MS_e = 40.84, p < .001; F_i(1,76) = 5.00, MS_e = 197.08, p <$

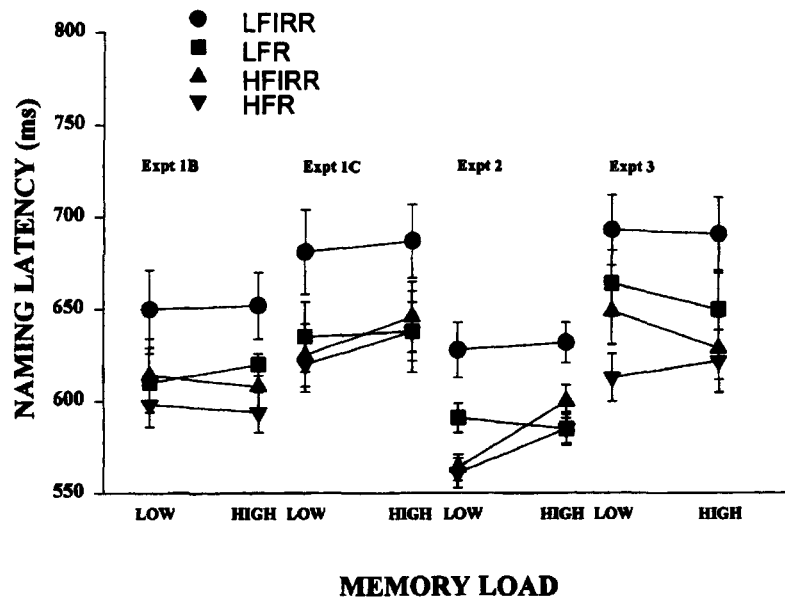


Figure 2. Mean naming latencies (in milliseconds) and standard error bars for subjects in Experiments 1B, 1C, 2, and 3, for low-frequency irregular words (LFIRR), low-frequency regular words (LFR), high-frequency irregular words (HFIRR), and high-frequency regular words (HFR), as a function of memory load.

.05]. The nature of this interaction was that the percentage of naming errors for low-frequency irregular words ($M = 12.51\%$) was much larger than the percentage of naming errors for low-frequency regular words ($M = 1.25\%$), whereas there was a much smaller difference in error rates between the high-frequency irregular ($M = 1.84\%$) and high-frequency regular ($M = 0.50\%$) words.

Memory task. For memory task response times, there was a significant main effect of memory load [$F_s(1,19) = 46.64$, $MS_e = 43,978.45$, $p < .001$; $F_i(1,76) = 158.17$, $MS_e = 13,237.89$, $p < .001$]. Also, the type of word presented in the naming task again influenced response latency in the memory task. Both the main effect of frequency [$F_s(1,19) = 7.23$, $MS_e = 8,380.73$, $p < .05$; $F_i(1,76) = 3.09$, $MS_e = 19,530.37$, $p < .10$] and the regularity \times frequency interaction [$F_s(1,19) = 8.86$, $MS_e = 6,475.83$, $p < .01$; $F_i(1,76) = 3.19$, $MS_e = 19,530.37$, $p < .10$] were significant in the subjects' analysis and marginal in the items' analysis.

The mean error rate in the memory task was 4.87%. The effect of memory load was again nonsignificant [$F_s(1,19) = 2.43$, $MS_e = 41.78$, $p > .10$; $F_i(1,76) = 2.36$, $MS_e = 51.85$, $p > .10$], although the error rate was less in the low-memory-load condition (4.14%) than in the high-memory-load condition (5.60%). Unlike in Experiment 1A, there were no other significant effects (all $ps > .10$).

Although some of the findings of this experiment were similar to the results of the Paap and Noel study, the most important result—the pattern of naming latencies in the high-memory-load condition—was not. Rather, the subjects in the present experiment seemed to be relatively unaffected by increased memory load. In fact, the

mean response times for the naming task were almost identical in the low-memory-load and high-memory-load conditions, for each of the four types of words. Thus, contrary to the hypothesis proposed in the rationale for Experiment 1B, inducing the subjects to code the digits verbally seemed to diminish rather than increase any effects of memory load.

An alternative reason why the data of Experiments 1A and 1B did not resemble those of Paap and Noel is that, although the instructions stressed the importance of the memory task, our subjects may still not have considered it as important as Paap and Noel's subjects did. This suspicion would seem to gain some support from the fact that, in the present experiments, the subjects treated the memory task as a speeded task, just like the naming task. Apparently, speeded responding in the memory task was not stressed in the Paap and Noel study; rather, accuracy was. In fact, although the memory task is a standard "memory-scanning" task (e.g., Sternberg, 1966), Paap and Noel report that there was no response latency difference between a one- and a five-item load (although there was a 3% accuracy difference). Therefore, in Experiment 1C, all reference to speeded responding in the memory task was eliminated to see if de-emphasizing speed and emphasizing accuracy in the memory task would alter the pattern of results in the naming task.

EXPERIMENT 1C

Method

Subjects. The subjects in Experiment 1C were 20 undergraduates at the University of Western Ontario. None had participated in the previous experiments. Average age of the subjects was 19.35

years. Sixteen subjects in this experiment received partial course credit in an introductory psychology course for their participation, and 4 subjects were paid \$4.00 to participate.

Apparatus, Stimuli, and Procedure. The apparatus, stimuli, and procedure for Experiment 1C were the same as those used in Experiment 1B. The only difference in methodology was that, in the present experiment, the subjects were instructed that for the memory task the speed of their response was unimportant and that they should concern themselves only with accuracy. In addition, the feedback during practice trials was altered so that the subjects were not shown their response times for the memory task. Instead, they were shown only their response times for the naming task and whether they were correct on their memory task response.

Results and Discussion

Naming task. Twenty-seven percent of the mispronunciation errors were regularization errors, and the remainder again tended to be due to the subjects being unfamiliar with the words. Mean naming latencies are depicted in Figure 2.

As in Experiments 1A and 1B, the three-way interaction of frequency, regularity, and memory load was non-significant [$F_s(1,19) = 0.00$, $MS_e = 1,345.18$, n.s.; $F_i(1,76) = 0.03$, $MS_e = 3,530.21$, n.s.]. Also, as in the previous experiments, the main effects of frequency [$F_s(1,19) = 17.61$, $MS_e = 1,804.01$, $p < .001$; $F_i(1,76) = 6.27$, $MS_e = 7,898.85$, $p < .05$] and regularity [$F_s(1,19) = 11.90$, $MS_e = 2,448.26$, $p < .01$; $F_i(1,76) = 6.44$, $MS_e = 7,898.85$, $p < .05$] were significant. Their interaction [$F_s(1,19) = 16.85$, $MS_e = 1,005.85$, $p < .001$; $F_i(1,76) = 3.67$, $MS_e = 7,898.85$, $p < .10$] was again significant in the subjects' analysis and marginal in the items' analysis. No other effects reached significance (all $ps > .10$).

As can be seen in Figure 2, the pattern of results for Experiment 1C is similar to that observed in Experiments 1A and 1B and does not resemble the pattern of results found by Paap and Noel. That is, in all these experiments, naming latencies for both the low-frequency irregular and the low-frequency regular words tended to increase slightly under high memory load. Thus, there was no reduction in the regularity effect. Furthermore, there was no evidence of an increase in the frequency effect for regular words—the other important effect that was observed by Paap and Noel. (In fact, just the opposite occurred in every one of these experiments.)

The mean error rate in the naming task was 4.62%. As in the previous experiments, the only significant effects were the main effects of frequency [$F_s(1,19) = 43.09$, $MS_e = 33.18$, $p < .001$; $F_i(1,76) = 5.65$, $MS_e = 255.92$, $p < .05$] and regularity [$F_s(1,19) = 43.01$, $MS_e = 45.30$, $p < .001$; $F_i(1,76) = 7.74$, $MS_e = 255.92$, $p < .01$] and their interaction [$F_s(1,19) = 42.61$, $MS_e = 30.51$, $p < .001$; $F_i(1,76) = 5.14$, $MS_e = 255.92$, $p < .05$], such that the percentage of naming errors for low-frequency irregular words ($M = 14.05\%$) was much larger than the percentage of naming errors for low-frequency regular words ($M = 1.28\%$), whereas there was a much smaller difference in error rates between the high-frequency ir-

regular ($M = 2.31\%$) and high-frequency regular words ($M = 1.00\%$).

Memory task. In the latency analysis, there was a significant main effect of memory load [$F_s(1,19) = 83.89$, $MS_e = 34,452.17$, $p < .001$; $F_i(1,76) = 95.10$, $MS_e = 33,350.52$, $p < .001$]. In addition, there was again evidence that the type of word presented in the naming task affected memory responses: there was a main effect of frequency in the subjects' analysis that was marginal in the items' analysis [$F_s(1,19) = 9.13$, $MS_e = 13,105.18$, $p < .01$; $F_i(1,76) = 2.75$, $MS_e = 25,488.57$, $p < .10$].

The mean error rate in the memory task was 4.18%. Unlike in the previous experiments, the memory load effect was significant [$F_s(1,19) = 7.64$, $MS_e = 45.85$, $p < .01$; $F_i(1,76) = 6.54$, $MS_e = 44.75$, $p < .05$] due to the fact that the error rate was higher in the high-load condition (5.58%) than in the low-load condition (2.79%). The only other significant effect was the frequency effect [$F_s(1,19) = 5.08$, $MS_e = 76.78$, $p < .05$; $F_i(1,76) = 7.47$, $MS_e = 54.03$, $p < .01$].

In Experiment 1C, the goal was to see if eliminating the speeded response in the memory task and instead asking the subjects to respond as accurately as possible on that task without concerning themselves with the speed of their response would change the pattern of results. Consistent with these instructions, the subjects did respond more slowly in the memory task of Experiment 1C (879 msec) than in the memory task of either Experiment 1A or Experiment 1B (707 and 739 msec, respectively) and were slightly more accurate (error rates were 5.27%, 4.87%, and 4.18% in Experiments 1A, 1B, and 1C, respectively). Furthermore, there was less evidence that responses in the memory task were affected by the nature of the word in the naming task (and no evidence that responses were affected by the word's regularity). In spite of the apparent extra importance attached to the memory task in Experiment 1C, response times in the naming task again did not resemble those reported by Paap and Noel. Thus, at the very least, one can conclude that the requirement of speeded responding in the memory task in Experiments 1A and 1B was not the source of the failure to replicate the Paap and Noel results.

EXPERIMENT 2

In general, Experiments 1A, 1B, and 1C showed smaller memory load effects than those observed by Paap and Noel. In fact, the memory load effect on naming latencies was significant only in Experiment 1A and not in Experiments 1B and 1C. Thus, the five-item memory load did not seem to create as large an attentional drain for the subjects in the present experiments as it did for Paap and Noel's subjects. If so, the result may be that our subjects simply were not forced to rely more heavily on the lexical route in the high-memory-load condition. As such, it would not be at all surprising that

our findings would be different from those of Paap and Noel. This hypothesis was examined in Experiment 2.

The purpose of Experiment 2, then, was to increase the difficulty of the high-memory-load condition in order to test the hypothesis that previous failures to replicate the Paap and Noel (1991) findings were due to this condition not being particularly difficult for our subjects. Thus, in Experiment 2, the digit load in the high-memory-load condition was increased from five to seven.

Method

Subjects. The subjects in Experiment 2 were 24 undergraduates at the University of Western Ontario. Average age of the subjects was 20.08 years. None had participated in the previous experiments.

Apparatus, Stimuli, and Procedure. The apparatus for Experiment 2 was the same as that used in Experiment 1A, except that a different PC and monitor were used for presentation of the stimuli. The PC used in Experiment 2 was a Trillium Computer Resources PC (Model No. 316S-80MS), and the monitor was a TTX Multiscan Monitor (Model No. 3435P). The procedure used in Experiment 2 was identical to that of Experiment 1C, except that seven digits were presented instead of five in the high-memory-load condition. The word stimuli for this experiment were the same as those used in the three previous experiments.

Results

Naming task. Forty-three percent of the mispronunciation errors were regularization errors. The remainder again tended to be due to the subjects being unfamiliar with the words. Mean naming latencies are depicted in Figure 2.

The three-way interaction of frequency, regularity, and memory load was again nonsignificant [$F_s(1,23) = 0.00$, $MS_e = 1,002.92$, n.s.; $F_i(1,76) = 0.09$, $MS_e = 1,375.71$, n.s.]. The effect of memory load was significant in the items' analysis but not in the subjects' analysis [$F_s(1,23) = 2.57$, $MS_e = 3,952.27$, $p > .10$; $F_i(1,76) = 4.87$, $MS_e = 1,375.71$, $p < .05$]. The interaction of frequency and memory load was, however, significant [$F_s(1,23) = 11.45$, $MS_e = 1,021.94$, $p < .01$; $F_i(1,76) = 9.38$, $MS_e = 1,375.71$, $p < .05$] in both analyses due to the fact that low-frequency words were affected by memory load much less than were high-frequency words.

The main effects of frequency [$F_s(1,23) = 25.42$, $MS_e = 1,868.14$, $p < .001$; $F_i(1,76) = 10.23$, $MS_e = 4,386.62$, $p < .01$] and regularity [$F_s(1,23) = 20.46$, $MS_e = 1,558.39$, $p < .001$; $F_i(1,76) = 7.93$, $MS_e = 4,386.62$, $p < .05$] were again significant, and their interaction was again significant in the subjects' analysis and marginal in the items' analysis [$F_s(1,23) = 7.17$, $MS_e = 1,825.53$, $p < .05$; $F_i(1,76) = 3.12$, $MS_e = 4,386.62$, $p < .10$].

The mean error rate in the naming task was 3.94%. Again, the only significant effects were the main effects of frequency [$F_s(1,19) = 44.98$, $MS_e = 28.43$, $p < .001$; $F_i(1,76) = 5.11$, $MS_e = 220.32$, $p < .05$] and regularity [$F_s(1,19) = 145.40$, $MS_e = 18.34$, $p < .001$; $F_i(1,76) = 10.48$, $MS_e = 220.32$, $p < .01$] and their interaction [$F_s(1,19) = 42.61$, $MS_e = 24.09$, $p < .001$; $F_i(1,76) = 5.11$, $MS_e = 220.32$, $p < .05$], such that the percentage of naming errors for low-frequency irregular words ($M =$

13.11%) was much larger than the percentage of naming errors for low-frequency regular words ($M = 0.21\%$), whereas there was a much smaller difference in error rates between the high-frequency irregular ($M = 2.50\%$) and high-frequency regular ($M = 0.42\%$) words.

Memory task. Mean response times showed a main effect of memory load [$F_s(1,23) = 65.20$, $MS_e = 106,936.28$, $p < .001$; $F_i(1,76) = 254.32$, $MS_e = 22,775.28$, $p < .001$]. The interaction of memory load and regularity was significant in the subjects' analysis [$F_s(1,23) = 6.37$, $MS_e = 15,737.66$, $p < .05$] and marginal in the items' analysis [$F_i(1,76) = 3.33$, $MS_e = 22,775.28$, $p < .10$]. The nature of this interaction was unusual. That is, under high memory load, memory task responses were slower on trials with regular words ($M = 1,138$ msec) than on trials with irregular words ($M = 1,101$ msec), whereas under low memory load, memory responses were faster on trials with regular words ($M = 714$ msec) than on trials with irregular words ($M = 766$ msec). In the subjects' analysis, there was a significant main effect of frequency [$F_s(1,23) = 4.12$, $MS_e = 27,295.22$, $p < .05$], which was marginal in the items' analysis [$F_i(1,76) = 3.17$, $MS_e = 26,663.95$, $p < .10$].

The mean error rate in the memory task was 9.73%. The only significant effect was that of memory load [$F_s(1,23) = 13.85$, $MS_e = 304.28$, $p < .001$; $F_i(1,76) = 59.02$, $MS_e = 56.87$, $p < .001$], where more trials were incorrect under the seven-digit memory load ($M = 14.35\%$) than under the one-digit memory load ($M = 5.13\%$).

Discussion

The memory load in Experiment 2 was increased from five to seven digits, and the increase in the error rate for the memory task indicates that increasing the memory load in this way did make the task somewhat more difficult. That is, the subjects made considerably more mistakes on the memory task when they had to remember seven digits than when they had to remember five in the previous experiments. Furthermore, there was a significant memory load effect in the items' analysis of the response time data of the naming task. Thus, in terms of some of the peripheral aspects of the data, the results of Experiment 2 resemble the results of the Paap and Noel experiment quite closely. Nonetheless, the subjects' response times in the naming task again did not resemble those in the Paap and Noel experiment. That is, there was no hint that the regularity effect for low-frequency words became smaller nor that the frequency effect for regular words became larger in the high-load condition. If anything, exactly the opposite occurred due to the fact that naming latency actually decreased slightly for the low-frequency regular words when the memory load was increased from one to seven digits.

On the basis of these results, task difficulty does not seem to be the source of the discrepancy between the present data and the Paap and Noel data. The discrepancy also does not seem to be due to the other "instruction-based" issues investigated in Experiments 1B and 1C.

Nonetheless, before concluding that the effect does not generalize to our subjects, we made one further attempt at replication. Herdman et al. (1993) have reported that they successfully replicated the Paap and Noel findings using a completely different set of words. Therefore, in Experiment 3, one final attempt at replication was undertaken, this time using the Herdman et al. stimuli and procedure.

EXPERIMENT 3

Method

Subjects. The subjects in Experiment 3 were 24 undergraduates at the University of Western Ontario. They were paid \$4.00 for their participation. None had participated in the previous experiments. Average age of the subjects was 22.63 years.

Apparatus and Procedure. The apparatus used in Experiment 3 was the same as that used in Experiment 2. The procedure was the same as that used in Experiment 1C.

Stimuli. The stimuli used in Experiment 3 were the words used by Herdman et al. (1993). The mean frequencies for high-frequency regular and irregular words were 1,035.6 and 1,040.9, respectively (Kučera & Francis, 1967). The mean frequencies for low-frequency regular and irregular words were 12.7 and 11.6, respectively (Kučera & Francis, 1967). All of the stimulus words seemed to be distinctly regular or irregular. These words are presented in Appendix B.

Results

Naming task. Sixty-four percent of the mispronunciations in this experiment were regularization errors, a higher percentage than in the previous experiments, because the subjects seemed to be, in general, more familiar with these words and, thus, made fewer errors due to unfamiliarity. Mean naming latencies are depicted in Figure 2.

Contrary to the findings of Paap and Noel (1991) but similar to the results of Experiments 1A, 1B, 1C, and 2, the three-way interaction of frequency, regularity, and memory load was not significant [$F_s(1,23) = 1.99$, $MS_e = 2,405.44$, $p > .10$; $F_i(1,76) = 0.04$, $MS_e = 1,255.81$, n.s.]. There were significant effects of frequency [$F_s(1,23) = 53.67$, $MS_e = 1,936.25$, $p < .001$; $F_i(1,76) = 0.34$, $MS_e = 6,689.72$, n.s.] and regularity [$F_s(1,23) = 14.86$, $MS_e = 2,573.04$, $p < .01$; $F_i(1,76) = 0.07$, $MS_e = 6,689.72$, n.s.] in the subjects' analysis but not in the items' analysis. The interaction of frequency and regularity was significant in the items' analysis [$F_i(1,76) = 5.47$, $MS_e = 6,689.72$, $p < .05$] but not in the subjects' analysis [$F_s(1,23) = 2.13$, $MS_e = 1,125.48$, $p > .10$]. No other effects approached significance (all $ps > .10$).

The error rate in the naming task was 3.26%. There were significant effects of frequency [$F_s(1,19) = 37.91$, $MS_e = 27.27$, $p < .001$] and regularity [$F_s(1,19) = 38.23$, $MS_e = 38.23$, $p < .001$], and a significant interaction of these factors [$F_s(1,19) = 24.49$, $MS_e = 34.98$, $p < .001$] in the subjects' analysis but not the items' analysis [$F_i(1,76) = 1.04$, $MS_e = 130.27$, n.s.; $F_i(1,76) = 0.30$, $MS_e = 130.27$, n.s.; $F_i(1,76) = 1.74$, $MS_e = 130.27$, $p > .10$, respectively], such that the percentage of naming errors for low-frequency irregular words ($M = 10.76\%$)

was much larger than the percentage of naming errors for low-frequency regular words ($M = 0.62\%$), whereas there was a much smaller difference in the error rates between the high-frequency irregular ($M = 1.89\%$) and high-frequency regular ($M = 0.21\%$) words.

Memory task. For response latencies, results indicated a main effect of memory load [$F_s(1,23) = 112.97$, $MS_e = 35,949.29$, $p < .001$; $F_i(1,76) = 243.30$, $MS_e = 13,646.69$, $p < .001$]. As in Experiment 2, the interaction of memory load and regularity was significant in the subjects' analysis [$F_s(1,23) = 4.43$, $MS_e = 11,749.54$, $p < .05$] but not in the items' analysis [$F_i(1,76) = 0.33$, $MS_e = 13,646.69$, n.s.]. In the subjects' analysis, there was a significant main effect of frequency [$F_s(1,23) = 28.30$, $MS_e = 11,146.11$, $p < .001$], but this effect was not significant in the items' analysis [$F_i(1,76) = 0.31$, $MS_e = 33,651.79$, n.s.]. No other effects approached significance (all $ps > .10$).

The mean error rate in the memory task was 3.75%. No effects were significant (all $ps > .10$).

Discussion

The naming-task results from Experiment 3 do not resemble the naming-task results of Paap and Noel (1991) or Herdman et al. (1993), despite the fact that Herdman et al.'s words and methodology were used. Since replication was not achieved with the Paap and Noel stimuli in Experiments 1A, 1B, 1C, and 2 or with the Herdman et al. stimuli in Experiment 3, the bottom line seems to be that this effect simply does not exist in the subject pool from which the present subjects were drawn.

GENERAL DISCUSSION

Paap and Noel (1991) reported a set of results that appears to give strong new support to dual-route theories of word naming. In their experiment, they demonstrated that a high memory load facilitated the naming of low-frequency irregular words and dramatically slowed the naming of low-frequency regular words. With the appropriate assumptions, these results can be predicted by dual-route theories but do not appear to follow in any straightforward way from Seidenberg and McClelland's PDP model. While some have reported successful replications of those results (e.g., Herdman et al., 1993), a number of other researchers have had somewhat less success. Working under the assumption that there is some (minor) methodological difference that caused those failures to replicate, the aim of the present experiments was to find the precise experimental conditions under which the Paap and Noel results could be replicated. Unfortunately, in no instance were we able to obtain data at all consistent with the Paap and Noel results. Rather, across all experiments, the general effect of a high memory load was a slight increase in naming latency for all stimuli.

As noted, it was not necessary to replicate the identical pattern reported by Paap and Noel to have accom-

plished a "successful replication." What is crucial for a successful replication is that (1) naming latencies for the low-frequency regular and irregular words come together under a high memory load, thus creating a smaller regularity effect and (2) naming times for high- and low-frequency regular words diverge, thus creating a larger frequency effect. In essence, both of these effects are based on a high memory load having a large slowing effect on the low-frequency regular words. Quite clearly, this did not occur in any of the experiments reported here. In fact, over all experiments, low-frequency regular words were the words least affected by memory load. The obvious question is why our results were so different from Paap and Noel's.

As noted by Paap and Noel (1991) and in the present paper, dual-route theories only predict the Paap and Noel effects if the assumptions made about the attentional demands of the two routes and the memory task are correct. The fact that Paap and Noel's data came out as predicted suggests that those assumptions were correct for their subjects. Thus, in an attempt to reconcile the two sets of results within the dual-route framework, it would seem to be necessary to argue that while these assumptions were correct for their subjects, they were not correct for ours. In the following, we consider three ways in which, in theory, this could have been true. The first is based on how our subjects may have treated the memory task, and the second and third are based on possible individual differences in reading style and reading ability between our subjects and Paap and Noel's.

Memory Task Performance

One obvious difference between our experiments and Paap and Noel's (1991) was that we usually found memory task effects due to the nature of the word in the naming task, whereas they did not. The argument could be made that the existence of these effects indicates that our subjects were not prioritizing the memory task, and, hence, it may not have been as attention-demanding as for Paap and Noel's subjects.

There are, however, a number of reasons to discount this argument. First, our subjects were actually more accurate on the memory task than were the subjects in the Paap and Noel experiment. This fact would seem to indicate that our subjects were putting even more emphasis on the memory task than Paap and Noel's subjects had.

Second, it seems likely that the Paap and Noel analysis had very little power to detect effects in their memory task. To begin with, Paap and Noel used only 14 subjects, in contrast to the 20–28 used in the present experiments. More importantly, it would seem that only a lack of power could explain why they were unable to detect a memory load effect in their memory task latency data. Although Paap and Noel did not report their memory task latencies, the standard finding in the memory-scanning literature (Sternberg, 1966) is that each new item in the set adds 40+ msec to the overall response time. Such was certainly the case in the present Experiments 1A and 1B. When a stronger encouragement to

pay attention to the memory task was provided in Experiments 1C, 2, and 3, the effects were even larger. It would be surprising if Paap and Noel's effects were substantially smaller than ours. The fact that they were, nonetheless, nonsignificant clearly suggests a lack of power in their memory task analyses. As such, it is not at all surprising that they did not obtain effects of other factors, such as frequency and regularity.

Third, if Paap and Noel's results on the memory task really did indicate absolutely no effect of the nature of the word, it would raise the question of whether their subjects were paying sufficient attention to the naming task. That is, on the basis of previous results using priming paradigms in which a response is required to both prime and target (e.g., Lupker & Williams, 1989), the difficulty of responding to the prime always affects the speed of responding to the target. If the naming task in Paap and Noel's experiment were de-emphasized to such an extent that these types of effects completely disappeared, then the naming latencies may have become meaningless. Since the primary dependent measure in this research is naming latency, subjects should be processing the naming-task stimuli carefully, while maintaining the memory task as the primary task. The result should be at least some spillover into the memory task.

Finally, in order to bring empirical evidence to bear on this issue, we conducted the following analysis. We selected a set of subjects who most clearly showed no effects of the nature of the word in the naming task on memory task performance, and we then evaluated their naming-task data. These subjects had to meet three criteria. First, for their memory task latencies, they had to have a difference smaller than 20 msec between response times for high-frequency and low-frequency words. Second, for the same data, they had to have a difference smaller than 20 msec between response times for regular and irregular words. Third, the subjects had to have no frequency or regularity effects in their memory task error rates. Of the 116 subjects who participated in these studies, we found 28 who met the above criteria. For these subjects, the mean naming response times for low-frequency irregular words in the low- and high-memory load conditions were 622 and 649 msec, respectively. For low-frequency regular words, these means were 594 and 598 msec. For high-frequency irregular words, these means were 583 and 592 msec. For high-frequency regular words, they were 570 and 582 msec. This pattern of results is quite similar to those recorded in each of the present experiments. Thus, it appears that our failure to replicate the Paap and Noel results was not due to our subjects not prioritizing the memory task.

Reading Style

A second possible explanation within the dual-route framework would be based on the idea that subjects bring many strategies to bear in these types of experiments (e.g., Baluch & Besner, 1991; Lupker, Brown, & Colombo, 1993; Tabossi & Laghi, 1992) and, for those subjects who make extensive use of such strategies, the

contributions of the two routes and the effects of attentional demands on the routes become somewhat obscured. Possibly, there were simply more of these types of subjects in the present experiments than in Paap and Noel's (1991) experiments. If so, the argument for analyzing only those subjects who show the standard frequency \times regularity interaction (the "true dual-routers") would be a reasonable one (Bernstein & Carr, 1991; Paap, personal communication, January 1993). The purpose of the Paap and Noel experiment was, after all, to observe the effects of memory load on the standard frequency \times regularity interaction. If subjects do not show this interaction, they could not be expected to show the predicted effects of a high memory load on naming latencies.

To evaluate this idea, subjects of this nature were selected from among the 68 who had participated in Experiments 1A, 1B, and 1C. For subjects to be included, three criteria had to be met. First, the subjects had to have a mean latency for low-frequency irregular words that was at least 15 msec longer than their mean latency for low-frequency regular words. Second, their mean latency for high-frequency words had to be at least 15 msec shorter than that for low-frequency words. Finally, the difference between their mean latencies for high-frequency regular and high-frequency irregular words had to be less than 15 msec. Eighteen subjects met these criteria. The mean naming latencies for these subjects are presented in Table 1.

Analyses of these results showed a significant three-way interaction of frequency, regularity, and memory load in the subjects' analysis [$F_s(1,17) = 5.86$, $MS_e = 1,244.98$, $p < .05$] and a marginal interaction in the items' analysis [$F_i(1,76) = 3.30$, $MS_e = 1,611.38$, $p < .10$]. There were also significant effects of frequency [$F_s(1,17) = 41.17$, $MS_e = 2,060.77$, $p < .001$; $F_i(1,76) = 19.76$, $MS_e = 5,062.24$, $p < .001$] and regularity [$F_s(1,17) = 23.97$, $MS_e = 2,467.16$, $p < .001$; $F_i(1,76) = 17.64$, $MS_e = 5,062.24$, $p < .001$], and, because subjects were selected on that basis, a significant frequency \times regularity interaction [$F_s(1,17) = 36.44$, $MS_e = 827.81$, $p < .001$; $F_i(1,76) = 5.85$, $MS_e = 5,062.24$, $p < .05$]. The effect of memory load was not significant in the subjects' analysis but was marginal in the items' analysis [$F_s(1,17) = 0.39$, $MS_e = 1,826.26$, n.s.; $F_i(1,76) = 3.33$, $MS_e = 1,611.38$, $p < .10$].

As shown in Table 1, the naming latency for low-frequency irregular words is faster in the high-memory-load condition than in the low-memory-load condition. This is similar to the effect demonstrated in the Paap and Noel study. What is not similar to the Paap and Noel (1991) results is that the naming latencies for low-frequency regular words are virtually the same in both the low-memory-load condition and the high-memory-load condition while, at the same time, the high memory load slowed responding substantially for both high-frequency regular and high-frequency irregular words. In fact, the mean latency for low-frequency regular words in the high-memory-load condition is actually

Table 1
Mean Naming Latencies (in Milliseconds) for the 18 Subjects in Experiments 1A, 1B, and 1C Showing a Standard Frequency \times Regularity Interaction, as a Function of Frequency, Regularity, and Memory Load

Word Type	Memory Load		Effect
	Low	High	
Low Frequency			
Irregular	726	706	-20
Regular	648	648	0
High Frequency			
Irregular	622	658	+36
Regular	609	630	+21

faster than the mean latency for high-frequency irregular words in that condition. This effect is not part of the Paap and Noel findings and, furthermore, it would appear to speak strongly against the Paap and Noel explanation. That is, if the high memory load were tying up the assembly route so that the lexical route provided most of the words' names, high-frequency words should be named substantially faster than should low-frequency words. Thus, even with this subject-selection procedure, we still could not find data consistent with the Paap and Noel results.

There is an obvious alternative explanation of the faster naming times for low-frequency irregular words for these 18 subjects in the high-memory-load condition. This explanation is based on the fact that for subjects to show a standard frequency \times regularity interaction under low memory load, they must show a fairly large regularity effect for low-frequency words and a small regularity effect for high-frequency words. Thus, these types of subjects would have to have a relatively slow response time for low-frequency irregular words and somewhat faster response times for the other three types of words. Thus, when considering the high-memory-load conditions, it seems possible that, for these 18 subjects, the tendency would be for them to have faster naming latencies for low-frequency irregular words due to nothing more than the principle of regression toward the mean.

Although this argument is based on results for individual subjects, its viability can at least be examined by looking for the same type of effect in the group data. In order to do this, a median split was done for the 68 subjects tested in Experiments 1A, 1B, and 1C on the basis of their response times to the low-frequency irregular words in the low-memory-load condition. The question was whether the speed-up for these words under high memory load would be observed for the 34 subjects with naming latencies slower than the median in this condition. Twelve of the 18 subjects who showed the frequency \times regularity interaction, and, hence, were part of the earlier analysis, were in the group of 34 subjects who were slower than the median. The mean naming latencies for these 34 subjects are presented in Table 2.

The pattern of means shown in Table 2 is not unlike that for the 18 subjects selected on the basis of having a standard frequency \times regularity interaction (see Table 1).

In particular, when subjects are selected only on the basis of a slow naming latency for low-frequency irregular words under low memory load, they also show a decrease in naming latency for those stimuli in the high-memory-load condition. Thus, it does seem possible that since selecting subjects for their frequency \times regularity interaction in the low-load condition meant that they, by necessity, had relatively slow naming latencies for low-frequency irregular words, their tendency to have faster naming latencies under high memory load may have been a function of regression toward the mean, rather than due to the way they processed words. Taken together, then, these results provide very little support for the idea that our failure to replicate the Paap and Noel results was a function of having only a small number of "true dual-routers" among our subjects.

Reading Ability

As a final possibility (again keeping within the dual-route framework), it could be proposed that our subjects did not show the Paap and Noel effects because they were simply better readers than the Paap and Noel subjects were. Thus, for our subjects, both their assembly and lexical routes were sufficiently automatized that the effects of a manipulation such as increasing memory load would never be sufficiently attention-demanding to produce the Paap and Noel results.

Unless one wanted to argue that our subjects were completely different from Paap and Noel's, however, the implication would still seem to be that our poorer readers should show some evidence of the Paap and Noel effect. As noted, such was not the case when considering the slower half of the subjects from Experiment 1A. To evaluate this hypothesis more completely, we calculated the mean naming latencies for all 116 subjects in our experiments for the purpose of finding the 29 subjects with the longest naming latencies (the slowest 25% of our subjects). We were working under the assumption that naming latency is a reasonable measure of reading ability. The overall mean response times for these subjects were all greater than 679 msec. The mean naming latencies for these subjects were 787 and 803 msec for low-frequency irregular words in the low- and high-memory-load conditions, respectively. For low-frequency regular words, these means were 738 and 735 msec. For high-

frequency irregular words, these means were 729 and 732 msec. For high-frequency regular words, they were 697 and 717 msec. As is obvious, these subjects also did not show the Paap and Noel pattern.

Our analyses have, therefore, provided very little evidence to support any of these alternative explanations of the discrepancy between the present results and the Paap and Noel results. In essence, there does not appear to be any obvious way to reconcile the results of the present experiment with those of the Paap and Noel (1991) experiment within the framework of dual-route theory. This is not to say that no reconciliation would be possible. At present, however, we are at a loss to understand what form that explanation might take.

Single-Route Models

Since our results do not seem to fit well within a dual-route framework, we should next explore the possibility that a better explanation for the discrepancy between the results of the present study and those of Paap and Noel (1991) could be provided by another theoretical framework. As noted, Paap and Noel considered and rejected the possibility that Seidenberg and McClelland's PDP model could explain their data. What we would like to suggest, however, is that with the right assumptions, single-route models can provide not only an adequate explanation of the Paap and Noel data but also a possible reconciliation of those data with the present data. The framework that we will use to discuss these issues is that of Van Orden et al. (1990). This explanation will be based to some extent on individual differences.

Recall that the single-route model described by Van Orden et al. (1990) incorporates phonological coding as its fundamental process. In the model, connections between orthographic and phonological units are learned through experience with different correspondences (the *covariant learning* hypothesis). Van Orden et al. (1990) evoke the concept of "phonologic space" to characterize the word-naming process, where a reader's initial encoding of a word is one point in phonologic space and the word's correct phonological code is represented by another point (called the *attractor*) in that space. In order to derive a pronunciation, the initial "noisy" encoding is cleaned up to match the attractor. The time that it takes this process to be completed is assumed to be mainly a function of the distance of the initial encoding from the attractor. For our purposes, to this would have to be added the assumption that processing time is also a function of the degree to which there are other attractors that are legitimate competitors.

With this additional assumption, a single-route theory may explain the effects of a concurrent memory task on naming latencies in the following way. The effect of having an increased memory load is that it hinders the ability of all competitors to compete. Two possible results could follow. If there are many somewhat viable competitors (as could be true for low-frequency irregular words), the result could be a freeing of the expected winner from the effects of competition. For example, if con-

Table 2
Mean Naming Latencies (in Milliseconds) for 34 Subjects in Experiments 1A, 1B, and 1C With Slowest Response Times for Low-Frequency Irregular Words Under Low Memory Load as a Function of Frequency, Regularity, and Memory Load

Word Type	Memory Load		Effect
	Low	High	
Low Frequency			
Irregular	758	734	-22
Regular	697	680	-17
High Frequency			
Irregular	677	675	-2
Regular	641	661	+20

stant amounts of strength were subtracted from all competitors, the effect may be to eliminate all but the usual winner from the competition. Alternatively, when only one competitor is actually a possible winner (i.e., for most other types of words), the effect is simply to slow down the cleaning-up process, producing slightly longer naming latencies. The argument would then be that, for subjects in the Paap and Noel (1991) experiment, there is some actual competition among possible responses when naming low-frequency irregular words, whereas for our subjects, the process of naming low-frequency irregular words is qualitatively the same as the process of naming all other types of words.

Although this explanation is, of course, quite ad hoc, it does derive some support from the available data. In particular, if competition were a major issue when naming irregular words, a larger regularity effect would be expected than if it were not. The mean regularity effect for low-frequency words for the Paap and Noel subjects in the one-item load condition was over 80 msec. In our experiments, these effects ranged from 30 to 46 msec. In fact, even for our slowest 25% of subjects, the effect was only 49 msec. Thus, even our slowest subjects appeared to have been much less bothered by spelling-to-sound irregularity than did Paap and Noel's subjects. If this argument is correct, then, it might be possible to account for the findings of Paap and Noel as well as the results of the present experiments within the framework of a single-route model.

Conclusions

Paap and Noel (1991) concluded that their findings were evidence that readers name words by two distinct processes. The results of the present experiments, however, are quite different from those of Paap and Noel, and we argue that both sets of results cannot be obviously accounted for by an individual difference explanation within the dual-route framework. Conversely, as discussed above, it is possible that the present results could be reconciled with the results of Paap and Noel within the framework of a single-route explanation (e.g., Van Orden et al., 1990). This explanation describes word naming as a process in which an initially noisy phonological encoding is first derived and then "cleaned up." The speed of the overall process would have to be assumed to be a function of both the noisiness of the initial encoding and the strengths of possible competitors during the clean-up process. In this type of model, assembly and lexical processes are not conceptualized as independent operations. If this explanation stands up to further analysis, it would certainly help support the claim that single-route theories provide a better theoretical framework than do dual-route theories for understanding the word-naming process (Brown, Lupker, & Colombo, 1994).

REFERENCES

- BALUCH, B., & BESNER, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 644-652.
- 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.
- BERNSTEIN, S. E., & CARR, T. H. (1991, November). *Dual route theories of reading: When does a concurrent task facilitate pronunciation?* Paper presented at the meeting of the Psychonomic Society, San Francisco.
- BESNER, D., TWILLEY, L., MCCANN, R. S., & SEERGOBIN, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review*, *97*, 432-446.
- BROWN, P., & BESNER, D. (1987). The assembly of phonology in oral reading: A new model. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 471-490). Hillsdale, NJ: Erlbaum.
- BROWN, P., LUPKER, S. J., & COLOMBO, L. (1994). Interacting sources of information in word naming: A study of individual differences. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 537-554.
- COLTHEART, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). New York: Academic Press.
- COLTHEART, M., CURTIS, B., ATKINS, P., & HALLER, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589-608.
- GLUSHKO, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception & Performance*, *5*, 674-691.
- GOUGH, P. B. (1972). One second of reading. In F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 331-358). Cambridge, MA: MIT Press.
- HERDMAN, C. M., BECKETT, B., & STOLPMANN, K. (1993, July). *Code-specific interference in phonological recoding: Evidence supporting a dual-route approach*. Paper presented at the meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Toronto.
- HUMPHREYS, G. W., & EVETT, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *Behavioral & Brain Sciences*, *8*, 689-740.
- KUČERA, H., & FRANCIS, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LUKATELA, G., & TURVEY, M. T. (1993). Similar attentional, frequency, and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 166-178.
- LUPKER, S. J., BROWN, P., & COLOMBO, L. (1993, November). *Strategic control in a naming task: Changing routes or changing deadlines*. Paper presented at the meeting of the Psychonomic Society, Washington, DC.
- LUPKER, S. J., & WILLIAMS, B. A. (1989). Rhyme priming of pictures and words: A lexical activation account. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 1033-1046.
- MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, *88*, 375-407.
- MORTON, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*, 165-178.
- PAAP, K. R., McDONALD, J. E., SCHVANEVELDT, R. W., & NOEL, R. W. (1987). Frequency and pronounceability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 221-243). Hillsdale, NJ: Erlbaum.
- PAAP, K. R., NEWSOME, S. L., McDONALD, J. E., & SCHVANEVELDT, R. W. (1982). An activation-verification model of letter and word recognition: The word-superiority effect. *Psychological Review*, *89*, 573-594.
- PAAP, K. R., & NOEL, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research*, *53*, 13-24.
- PATTERSON, K., & COLTHEART, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and per-*

- formance XII: The psychology of reading* (pp. 421-448). Hillsdale, NJ: Erlbaum.
- SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523-568.
- SEIDENBERG, M. S., WATERS, G. S., BARNES, M. A., & TANENHAUS, M. K. (1984). When does irregular spelling or pronunciation influence recognition? *Journal of Verbal Learning & Verbal Behavior*, *23*, 383-404.
- STERNBERG, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652-654.
- TABOSSI, P., & LAGHI, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition*, *20*, 303-313.
- TARABAN, R., & MCCLELLAND, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory & Language*, *26*, 608-631.
- VAN ORDEN, G. C. (1987). A rows is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181-198.
- VAN ORDEN, G. C., PENNINGTON, B. F., & STONE, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488-522.

APPENDIX A

Experiments 1A, 1B, 1C, and 2: Stimuli
(From Paap & Noel, 1991)

Low Frequency		High Frequency	
Irregular	Regular	Irregular	Regular
bury	buds	been	best
caste	canes	both	book
comb	coil	come	came
crow	curl	done	dark
glove	grade	door	deep
lure	lump	foot	flat
lute	lode	give	game
pour	pops	good	gain
ruse	rump	have	high
sew	sock	most	more
sans	sage	move	miss
sues	suck	said	same
sown	sobs	says	seem
wand	wade	sure	soon
warn	weed	touch	train
warp	wick	want	wall
wasp	weld	warm	wage
wily	wilt	were	well
wool	woke	word	west
worm	wink	work	week

APPENDIX B

Experiment 3: Stimuli
(From Herdman et al., 1993)

Low Frequency		High Frequency	
Irregular	Regular	Irregular	Regular
bowl	beam	are	best
bush	bean	both	block
caste	bunt	break	came
deaf	carve	come	dark
doll	deed	do	did
flood	dusk	does	group
gross	float	done	him
mow	grape	give	less
pint	peel	great	main
pear	plank	lose	out
plaid	pump	move	page
sew	sank	pull	place
soot	slat	put	sing
spook	slam	said	sit
swamp	stab	says	soon
sweat	swore	shall	stop
wad	wane	want	week
wand	weld	watch	which
wool	wig	were	while
worm	wink	work	with

(Manuscript received January 12, 1994;
revision accepted for publication October 31, 1994.)