Masters Theses 1911 - February 2014

1976

Phonemic processing in reading printed words :: effects of phonemic relationships between words on semantic categorization response time.

Margaret L. Mcmahon
University of Massachusetts Amherst

PHONEMIC PROCESSING IN READING PRINTED WORDS:
EFFECTS OF PHONEMIC RELATIONSHIPS BETWEEN WORDS
ON SEMANTIC CATEGORIZATION RESPONSE TIME

A Thesis Presented

By

MARGARET L. MCMAHON

Submitted to the Graduate School of the
University of Massachusetts in partial
fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 1976

Department of Psychology

PHONEMIC PROCESSING IN READING PRINTED WORDS: EFFECTS OF PHONEMIC RELATIONSHIPS BETWEEN WORDS ON SEMANTIC CATEGORIZATION RESPONSE TIME

A Thesis Presented

By

MARGARET L. MCMAHON

Approved as to style and content by:

Alexander Pollatsek, Ph. D., Chairperson of Committee

Charles Clifton, Ph. D., Member

Jerome L. Myers, Ph. D., Member

Bonnie Strickland, Ph. D., Department Head

Department of Psychology

August, 1976.

Acknowledgement

I would like to thank my committee members for their generous assistance. Alexander Pollatsek deserves special appreciation for his technical assistance, patience, and intellectual challenge. And I thank my husband Bruce Rideout, who contributed greatly to this work through constant aid and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The present research is part of a long term attempt to understand the reading process. The general question addressed is whether some form of silent speech is a significant subprocess in comprehending the meaning of text. The term silent speech is used in the broad sense. As described by Conrad (1972), the phenomena of silent speech range from full articulation (e.g., lip movement) without sound, through articulation observed with the aid of electromyography (EMG), to auditory imagery which may not be detectable using EMG.

Many investigators (e.g., Hochberg, 1970) have regarded the translation of a seen word into an "inner utterance" or some representation of a heard word, as a crucial step in the silent reading process. The contrasting point of view holds that reading is, or can be, a strictly visual affair with no mediation by mechanisms used to process speech. Bower (1970), for example, argues that reading often has no auditory component for skilled readers.

Clarification of the role of auditory processing in silent reading has important implications for the field of reading education. As Hardyck and Petrinovich (1970) remarked, the remedial reading literature stresses the importance of eliminating subvocal speech during reading. Because extensive subvocalization is thought to correlate with slow reading (Hardyck and Petrinovich, 1969), it is generally assumed that

better, or at least faster, reading results from elimination of the silent speech habit. It is standard practice in reading classes to attempt to suppress the subvocalization tendency in adult readers and to avoid the use of initial teaching methods which are suspected to lead to silent speech (Edfeldt, 1960).

Very little is actually known, however, about the function of subvocalization in silent reading. It is possible that for some readers phonemic processing is a necessary part of deriving meaning from printed text. Since techniques sensitive enough to measure reliably vocal muscle activity during silent reading were developed (Faaborg-Andersen and Edfeldt, 1958), a few studies have shown that activity in the speech musculature increases with conceptual difficulty of the text read (Edfeldt, 1960; Hardyck and Petrinovich, 1970). Hardyck and Petrinovich (1970) found that subjects who were required to suppress subvocalization (maintain laryngeal muscle activity at nonreading relaxation levels) during reading, performed significantly less well on a comprehension test than control subjects reading the same, difficult material.

While these EMG studies do not demonstrate that silent speech (or any phonemic processing) is necessary for comprehension, their findings are consistent with the hypothesis that phonemic processing of printed words has some semantic function. An understanding of auditory (or phonemic) processing is necessary for a complete theory of reading, and as a

basis for evaluating methods directed at eliminating subvocalization. Both measureable peripheral muscle activity and inferred central phonemic encoding, as indices of silent speech, will be useful in research on the relation between deriving meaning from printed language and auditory processing. The present study is concerned with inferred phonemic encoding.

A number of experiments (reviewed by Meyer, Schvaneveldt, and Ruddy, 1974, and by Barron, 1975) have attempted to assess the role of phonemic encoding in reading, using information processing techniques. Meyer _et al_. describe three basic theories addressed by these experiments. Strictly speaking, these theories are hypotheses about _word-recognition_, the process whereby information about a single printed word is accessed in lexical memory. But they are often applied in general accounts of reading, i.e., in theories of comprehending printed text.

The _graphemic-encoding hypothesis_, as described by Meyer _et al_. (1974), holds that a printed word is recognized directly from a visual representation, which is used to locate stored information about the meaning of the word. Bower's (1970) view of reading as a strictly visual process exemplifies this theory.

The _phonemic-encoding hypothesis_ holds that recognition involves converting a visual representation of a word into a phonological representation. Thus, the phonemic code is required in accessing lexical memory. Phonemic encoding

theories may or may not be explicit about the nature of the auditory code—the vocal apparatus is not implied as the locus of phonemic encoding.

A third hypothesis holds that lexical memory can be accessed through both visual and phonological representations of a printed word. This dual-encoding hypothesis integrates the phonemic and graphemic hypotheses. Whereas the first two hypotheses specify the mode of the representation which accesses memory, the dual-encoding hypothesis allows for retrieval processes based on both types of code to occur in parallel.

While recent research of Meyer and his colleagues (Meyer and Ruddy, 1973; Meyer, Schvaneveldt, and Ruddy, 1974) favors the dual-encoding view, a few experiments have been reported in support of the graphemic theory (Bower, 1970; Baron, 1973). Bower (1970) argued for the possibility of purely visual reading. In his experiment he instructed a group of bilingual subjects to translate passages of Greek text into English. Passages of Greek words spelled in the usual fashion were used in one condition, while in a second condition the passages were modified by replacing some of the items with phonemically equivalent Greek pseudowords. The use of pseudowords was analogous to changing an English text by replacing words like PHONOGRAPH, with pseudowords like FONOGRAF. Bower found that the modified passages required twice as much time for translation. This finding presents

apparent difficulties for a *phonemic-encoding hypothesis*, since, if it is the phonemic code which is used in accessing memory, a phonological equivalent should serve as well as the usual, orthographically correct word. The pseudoword should not disrupt the translation process. Thus, Bower concluded that auditory encoding is not involved in the normal processing of printed words.

As Meyer *et al*. (1974) observed, Bower's (1970) results can be interpreted in a number of ways. Graphemic differences between words and pseudowords may affect operations prior to phonemic encoding. For instance, the construction of an initial graphemic representation of the word may be influenced by the familiarity of the orthographic structures. Thus, the pseudowords may be characterized by bigrams and trigrams which are infrequently seen in the written language, and which are more difficult to encode than words. Furthermore, a slower than normal transformation from graphemic to phonemic representation may be taking place for pseudowords. The operations that convert a letter string into a phonological representation may take longer when the string is graphemically anomalous. Thus, Bower's finding does not preclude phonemic encoding.

In a recent experiment which supported the *graphemic-encoding hypothesis*, Jonathan Baron (1973) asked subjects to decide whether or not various printed phrases "made sense" (Experiment I). He found that, when subjects classified

visually anomalous phrases (e.g., MY KNEW CAR, OUR NO CAR) as not making sense, the phonemic characteristics of the phrase had no effect on response time. When the phrase did not look meaningful, it did not matter whether it was phonemically congruent (e.g., MY KNEW CAR), or both phonemically and graphemically incongruent (e.g., OUR NO CAR). Baron argued that meaning is accessed directly from visual representations, since the phonemic properties of the string had no effect on reaction time. But the finding that graphemic incongruity is quickly detectable, and affects response time for meaningfulness, does not eliminate the possibility that phonemic encoding takes place. Graphemically anomalous phrases may be rejected as not making sense on the basis of their visual unfamiliarity alone. Furthermore, the phonemically congruent phrases (e.g., MY KNEW CAR) resulted in a relatively high error rate, which suggests that phonemic encoding did play a role in subjects' decisions about the phrases.

In Baron's (1973) Experiment II, subjects judged whether the printed phrases "sounded as if they made sense." It was found that, in judging whether a phrase sounded meaningful, subjects took less time to classify stimuli which were graphemically, as well as phonemically, normal. For example, MY NEW CAR, which is graphemically and phonemically congruent, was more quickly classified as sounding sensible than I NEW HIM, which is phonemically, but not graphemically congruent. The strictly visual properties of the stimuli thus played a

role in determining whether they sounded meaningful. From these data, Baron argued that words are recognized visually. But the fact that the looks of the phrase influences response time does not preclude phonemic encoding. It is possible that the unfamiliar visual pattern slows phonemic encoding.

It seems more likely that dual encoding takes place, and that responses to stimuli such as I NEW HIM are slow because, while phonemic processing indicates that the phrase makes sense, graphemic processing indicates that it is anomalous. This conflict might delay the response, if, given a negative outcome of processing in either mode, a subject checks before responding "Yes" or "No", to decide whether the mode in which the stimulus does not make sense is relevant to the task.

While the experiments reported to support the *graphemic-encoding hypothesis* (Bower, 1970; Baron, 1973) demonstrate the role of graphemic encoding in responding to visual stimuli, they do not provide a strong argument against a phonemic stage in the processing of printed words.

Phonemic encoding was inferred by Eriksen, Pollack, and Montague (1970) from the effects of phonological properties of stimuli on response time. The attribute varied was the number of syllables in a word or number, and the response was overt vocalization. Vocalization latency (from presentation of a word or number, to activation of a voice key by the first sounds of a subject's pronouncing response) was found to increase with the number of syllables to be pronounced.

One difficulty with this study, for the present purpose, is that the processes involved in organizing a vocal response to name the stimulus may differ from the processes involved in word recognition per se.

Stuart Klapp (1971) conducted an experiment which extends the generality of the Eriksen *et al*. (1970) finding, by introducing a task in which subjects do not pronounce the stimuli. Klapp's subjects were presented two numbers simultaneously, and instructed to say "yes" if the two stimuli were identical, and "no" if they were different. A second experiment required a manual response to the same situation. Klapp found that response latency increased with the number of syllables in the name of the number. He also found (in a third experiment) that number of syllables had an effect on time to decide that two words (e.g., CLEAR-COURT; COVER-COLOR) were different.

Unfortunately, the sets of numbers used by Eriksen *et al*. (1970), and by Klapp (1971) differed on dimensions other than number of syllables. As Klapp noted, most of the two-syllable numbers contained the digit *0*, while all of the four-syllable numbers contained the digit *7*. So some properties of the numbers (e.g., familiarity of *20*, *30*, *40*, etc.) were confounded with number of syllables. Although Klapp, in Experiment III, used words of identical length and frequency classification to overcome this difficulty, graphemic and phonemic properties may still have been confounded

in these stimuli presented in the *same-different* task. Comparing the two words of each pair, and counting the number of letter positions filled by different letters, one obtains an index of dissimilarity for the pair. CLEAR-CLEAR would have an index of *0*, and CLEAR-COURT would have an index of *4*, the members of the pair differing in four letters. Klapp's one-syllable words were found to differ in more letters (3.7 out of 5 letters, on the average) than the two-syllable words (3.0, out of 5). The one-syllable words may therefore be easier to classify as different, because they are more different (less confusable) graphemically. Phonemic encoding, then, was not clearly demonstrated in either Eriksen *et al.*'s or Klapp's study.

An experiment (Rubenstein, Lewis, and Rubenstein, 1971) which supported the *phonemic-encoding hypothesis* (that the phonological representation is used in accessing the internal lexicon), is also subject to the familiar complaint of confounding between phonemic and graphemic properties of stimuli. Rubenstein *et al.* used a *lexical-decision* task, which required subjects to judge whether strings of letters were words or nonwords. They found that unpronounceable nonwords (e.g., BRAKV) were more quickly rejected than pronounceable nonwords (e.g., BLEAN). Among pronounceable nonwords, those which were homophonic with English words (e.g., BRUME) took longer to reject than those which were not homophonic. So reaction time varied directly with phonemic approximation to

English. Rubenstein *et al*. also reported that positive responses were slower for English words that were homophonic with other English words (e.g., MAID, or MADE), than for nonhomophones (e.g., BATH). The phonemic properties of the words affected response time.

These results are compatible with the *phonemic-encoding hypothesis*. If a phonemic representation is used in lexical search, then unpronounceable nonwords should be quickly rejected because they violate phonological rules, making the search unnecessary. Nonwords which are homophonic to English words will take more time to reject than nonhomophones, because the homophones make a phonemic match in the comparison process, and must be checked for spelling before the response is made. Finally, the results for words are compatible with the *phonemic-encoding hypothesis* because homophonic words (e.g., MAID), which take longer to classify than other words, have two or more spellings that potentially must be checked, and spelling checks add to reaction time.

A difficulty with this interpretation of the outcome of Rubenstein *et al*.'s (1971) experiment is that the graphemic properties of the stimuli vary along with the phonemic properties. It is thus questionable to make inferences on the basis of the effects of phonemic properties. BRAKV, for example, might be rapidly rejected on the basis of violation of orthographic rules or regularities, as well as on the basis of phonemic illegality. It is also possible that non-

words homophonic to English _look_ most like English words. By a similar line of argument, Meyer _et al_. (1974) have shown that it is possible to explain (with certain assumptions) the Rubenstein _et al_. results in terms of the _graphemic-encoding hypothesis_. A number of experiments in this field have thus yielded inconclusive results, due to the difficulty of varying phonemic properties of stimuli, without varying their graphemic properties.

Meyer, Schvaneveldt, and Ruddy (1974) have provided a convincing demonstration of the role of phonemic encoding in visual word-recognition. Their experiments were based on the _lexical-decision task_, in which subjects must judge whether strings of letters are words. Meyer _et al_. presented a pair of letter strings on each trial. In Experiment I, subjects saw the members of a pair simultaneously, and responded "Yes" if both members were words. Reaction time was thus measured for the pair of stimuli as a unit. In Experiment II, the two letter-strings were presented successively, and the dependent variable was the decision time to the second string. The independent variable was the degree of phonemic similarity between the members of a pair.

Meyer _et al_. (1974) found that pairs whose members were phonemically similar (e.g., BRIBE-TRIBE; FENCE-HENCE) were recognized somewhat more quickly than their control pairs, whose members were dissimilar (e.g., BRIBE-HENCE; FENCE-TRIBE). It should be noted that members of the phonemically similar pairs are also similar graphemically, in that their

spelling differs only in the first letter position. Control pairs are both phonemically and graphemically dissimilar.

A second finding was that pairs whose members were phonemically dissimilar (but graphemically similar) resulted in significantly longer reaction times than their graphemically dissimilar (and non-rhyming) controls. Thus, pairs whose members were spelled almost alike (e.g., COUCH-TOUCH; FREAK-BREAK), but were pronounced differently, took longer to recognize than their controls (e.g., COUCH-BREAK; FREAK-TOUCH), which were dissimilar on both graphemic and phonemic dimensions.

In sum, Meyer *et al*. (1974) found that when graphemic similarity of members within pairs was held constant, the phonemic similarity (rhyming *vs*. not rhyming) between two words had an effect on reaction time to both a pair of words as a unit (Experiment I), and to the second of two successively presented words (Experiment II). This finding is inconsistent with a *graphemic-encoding hypothesis*, since, if subjects recognized words directly from their visual representations, the performance of lexical decisions would not depend on the phonemic relation within a pair. The *graphemic-encoding hypothesis* predicts that, if similarity of spelling does not vary, it will make no difference whether or not words of a pair rhyme. Meyer *et al*. thus conclude that their experiments support the *phonemic*- and *dual-encoding hypotheses*, but do not distinguish between them. Their salient

point is that the apparent influence of phonemic properties of preceding words on the phonemic encoding of subsequent words indicates the involvement of phonological representation in the word-recognition process.

Although Meyer _et al_. (1974) demonstrated that printed words are encoded phonemically sometimes and avoided the confounding of graphemic and phonemic properties of stimuli which characterized previous experiments (e.g., Rubenstein _et al_., 1971; Eriksen _et al_., 1970), their results are only suggestive with respect to the normal reading process. It is not evident that the subject performing a _lexical-decision task_ deals with the meanings of the words, which is a large part of what the typical silent reader is after in a typical silent reading task. The present research was intended to extend Meyer _et al_.'s demonstration of phonemic encoding in _lexical-decision tasks_, to a meaning-oriented task, which is presumably a closer approximation to normal reading. In the following section the view that conclusions based on _lexical-decision tasks_ do not automatically generalize to meaningful processing of printed text will be elaborated, and then a rationale for predicting phonemic effects on the semantic categorization task will be given.

It is possible (theoretically) that a subject could decide whether an isolated string of letters is a word or not, without accessing its meaning. A simple strategy would be to pronounce the letter string subvocally and notice whether the

name sounds like any real word or not (a spelling check would be needed to eliminate pseudowords). Such a strategy would not necessarily involve semantic analysis, and it would be quite likely to produce phonemic effects in an experiment like Meyer et al.'s (1974). While it may be the case that meaning is accessed in decisions about wordness, the question is open, and a process such as phonemic encoding which seems to be involved in decisions about words need not be involved in meaning access.

One argument against this view is that one must know that a word is a word before one can know anything about its meaning, and that consequently the processes involved in the lexical decision comprise a subprocess of access to meaning. Meyer and Ellis (1970), however, showed that some semantic category decisions are made reliably faster than lexical decisions. This finding indicates that decisions about wordness, at least as they take place in lexical-decision tasks such as those used by Meyer and Ellis and Meyer et al. (1974), are not always a component of meaning access for visually presented words. So the demonstration of phonemic encoding in the lexical decision task does not imply that such encoding takes place in reading, or in simple meaning decisions.

It is an empirical question, then, whether the phonemic encoding demonstrated by Meyer et al. (1974) applies to reading, or any processing of the printed word for meaning. An experiment by Meyer and Ruddy (1973), which supports the dual-

encoding hypothesis, suggests that subjects sometimes make use of phonological representations in performing the categorization task. For example, subjects who were asked to decide whether or not a word belonged to a given category (e.g., FRUIT), took longer to reject pseudomembers (like PAIR), which sound like members, than to reject nonmembers (like TAIL). In this case, if processing were strictly visual, PAIR and TAIL would have taken the same amount of time to reject, since neither looks like a fruit. PAIR may have taken longer because, sounding like a fruit, it must be checked for correct spelling, and then rejected. A phonemic match with PEAR must have caused the delay. While this result implies phonemic encoding, it is based on negative responses to a specialized set of words—homophones. It is also subject to the criticism that phonemic properties of the words were confounded with graphemic properties, since PAIR may look more like a fruit than TAIL does. A demonstration that phonemic properties within pairs of graphemically similar words (as compared to control pairs) influence encoding of words in a semantic categorization task, would provide clearer evidence that phonological representations are activated in meaning-oriented tasks.

Thus, the present research was designed to examine the effects of graphemic and phonemic relations (rhymes vs. nonrhymes, as in Meyer et al., 1974) within pairs of words, on reaction time in deciding whether printed words belong to

given categories. Performance on phonemically similar pairs like BEER-DEER (relative to appropriate control pairs) was compared to performance (relative to controls) on phonemically dissimilar pairs like YEAR-BEAR, when the subject decided whether the test word was, for example, an ANIMAL. While this categorization task probably does not involve quite the same processes used to access word meaning in silent reading, it goes beyond lexical decisions by introducing meaning decisions, and presents possibilities for studying the effects of semantic variables, such as context, on phonemic processing in a meaning-oriented task.

The model shown in Figure 1 is the encoding-bias model

Insert Figure 1 about here

of Meyer et al. (1974). It models a lexical-decision task in which two strings are presented simultaneously. The model assumes that processing of the strings is serial, and that a string of letters does not always have a unique pronunciation. According to the encoding-bias model, the initial phonemic representation of a string may be rejected, and the string recoded. This is likely to occur for the second string of a pair, when the grapheme/phoneme correspondence rules used in encoding the first string are applied in encoding the second string, and the members of the pair are phonemically dissimilar (e.g., YEAR-BEAR).

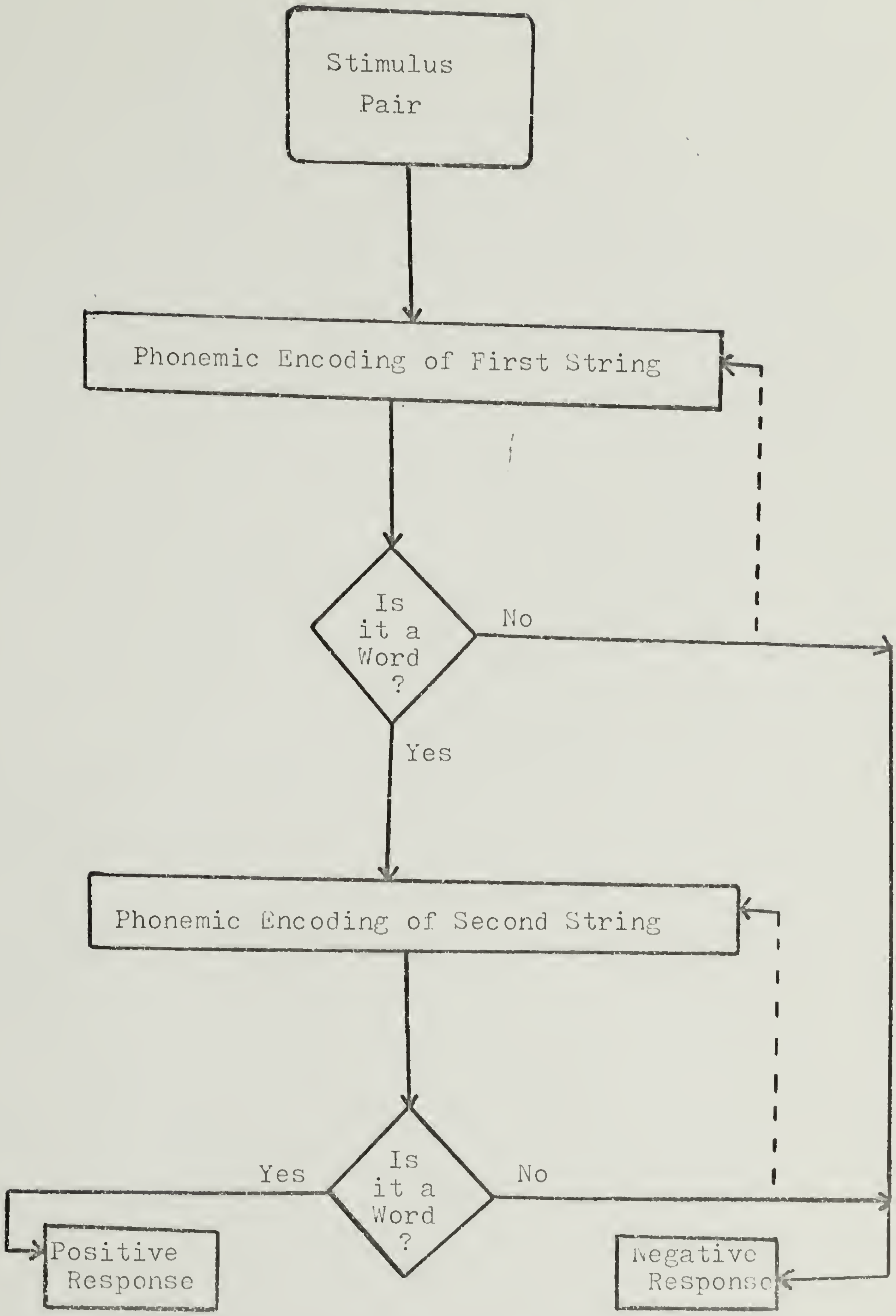

Figure 1. An encoding-bias model from Meyer et al. (1974), to explain the effects of graphemic and phonemic relations on visual word recognition.

It is reasonable to expect that encoding-bias also occurs in phonemic processing in the categorization task. Figure 2

Insert Figure 2 about here

---

is an integration of the <u>encoding-bias model</u> of Meyer <u>et al</u>. (1974) and the <u>dual-retrieval model</u> of Meyer and Ruddy (1973). The dual-retrieval model was proposed to account for findings which indicated a role for both graphemic and phonemic encoding in their categorization tasks. It was based on parallel processing of phonemic and graphemic representations in category search. The adaptation (in Figure 2) of Meyer and Ruddy's model shows how encoding bias might take place within the categorization process. It is presented to convey the plausibility of the hypothesis that the phonemic effects observed by Meyer <u>et al</u>. (1974) extend to the semantic categorization task used in the present experiment.

According to the model in Figure 2, the subject searches the specified category for the test word, and the outcome of the search determines the response. Graphemic and phonemic retrieval (search) processes begin with a common graphemic encoding stage. Then graphemic and phonemic category searches are executed in parallel, and a response occurs when either process finds a successful match or completes the search unsuccessfully. During the graphemic retrieval process, graphemic representations of the category members are searched

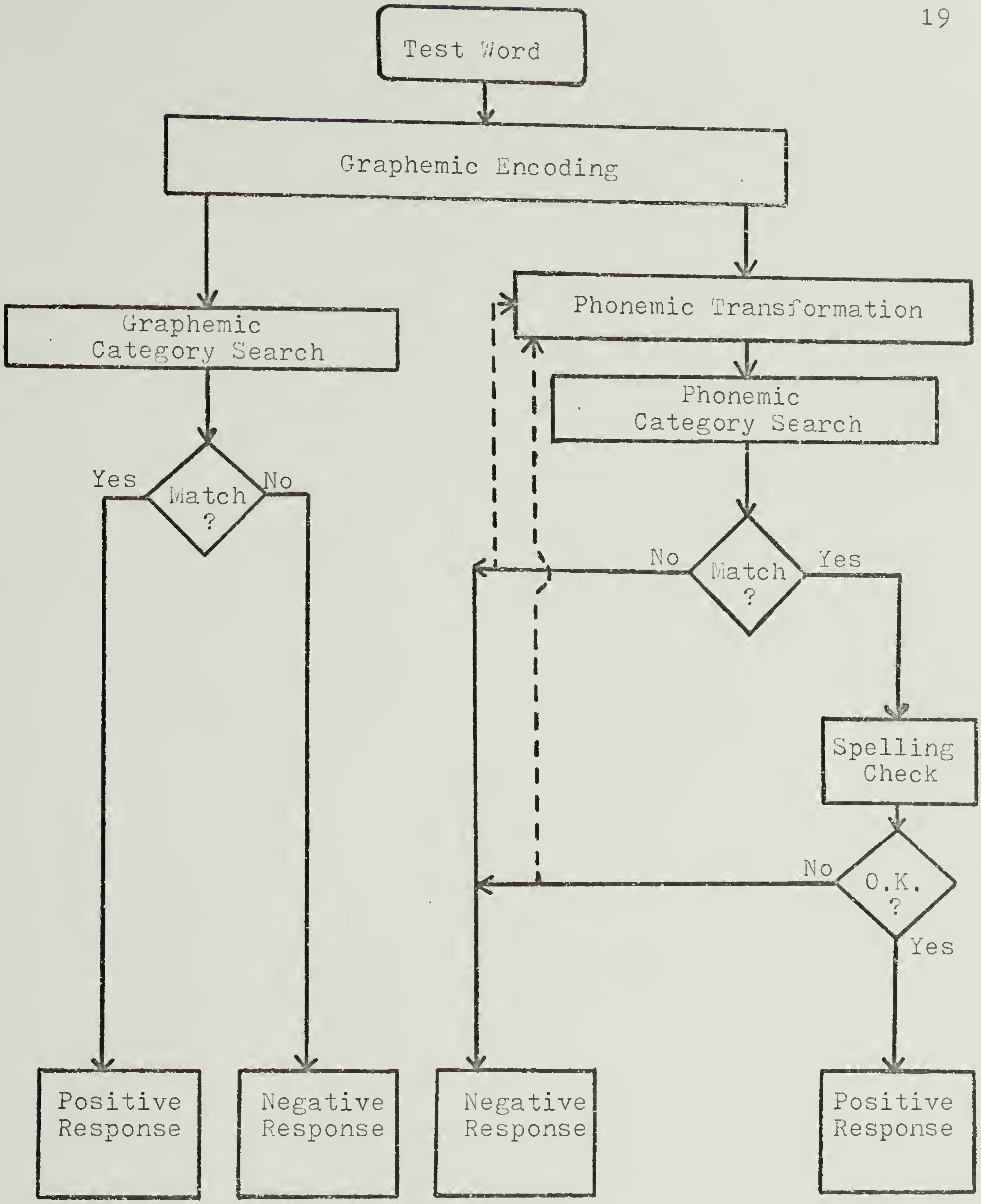

Figure 2. A dual-encoding model for categorization, with possibility of phonemic recoding.

and compared with the graphemically encoded test word. If the word is a member, then a match occurs, and a positive response is made. If a match does not occur, then a negative response is made.

The phonemic retrieval-process is simultaneous with the graphemic retrieval-process. If the graphemic process is still in progress, then a positive or negative response occurs whenever the necessary stages of the phonemic process are finished. In this process the phonemically encoded test word is compared with phonemic representations of the category members. If the test word is a member, then there is a match, and a spelling check is made followed by a positive response.[1] If no match is made (or if the outcome of the spelling check is negative), and if the word has more than one possible phonological representation, then the encoding and decision operations are repeated, as shown by the dashed line in Figure 2. This is likely to occur on pairs like YEAR-BEAR, for which graphemic similarity within the pair

---

[1]The spelling check is necessary because a stimulus may be a pseudomember of a category, i.e., it may sound like a member but have a different spelling. The spelling check, which determines whether the stimulus is spelled like a category member, involves comparing the graphemic representation of the test word with the graphemic representation paired in memory with the matching phonemic representation. A positive outcome occurs only when the two graphemic representations are identical. This process is based on the assumption from Rubeinstein _et al._'s (1971) theory that long-term memory contains, for each word, a phonemic representation paired with a graphemic representation (Meyer and Ruddy, 1973).

biases the subject toward inappropriate application of the same grapheme/phoneme correspondence rules to both members. The repetition continues until either a positive outcome occurs, or all of the alternatives have been checked exhaustively. If none of the possible representations of the word is found in category search, the processing terminates, and a negative response is made.

This dual-encoding model would predict that the second members of pairs which do not rhyme but are graphemically similar (e.g., YEAR-BEAR; HOME-SOME) would take longer to categorize (relative to their controls) than the second members of rhyming, graphemically similar pairs (relative to controls). The reaction time to rhyming pairs and controls (for positive responses) includes: 1) initial graphemic encoding; 2) the minimum of a) the graphemic search with positive outcome, b) the phonemic search with positive outcome plus spelling check; and 3) positive response execution. The reaction time to non-rhyming, graphemically similar pairs includes an extra step in phonemic processing: 1) initial graphemic encoding; 2) the minimum of a) the graphemic search with positive outcome, b) the phonemic search with positive outcome, which could occur normally, but which, if encoding bias is in effect, is likely to involve phonemic search with negative outcome followed by additional phonemic search with positive outcome plus spelling check; and 3) positive response execution. The model shows that phonemic encoding bias could

operate in basically the same way for a categorization task as for the lexical decision task. Although it shows that phonemic effects can be expected to extend to semantic categorization, as that process is conceived by Meyer and Ruddy (1973), the model in Figure 2 will not be used further in the present study for two reasons. It is not useful in making predictions for negative responses, which will comprise a large part of our data. Secondly, the model would not be acceptably plausible unless it were elaborated considerably to take into account the likely possibility that subjects could recognize their erroneous phonemic representations as nonwords, which they would be in most cases. When this occurred category search would be superfluous. Thus, Figure 2 is not a useful predictive model for the present experiment.

It was hypothesized that the effects of phonemic similarity on reaction time using a categorization task, are comparable to those observed by Meyer *et al*. (1974) with the lexical decision task. The prediction is consistent with the *dual-encoding* and *phonemic-encoding* hypotheses. It conflicts with the *graphemic-encoding hypothesis*, i.e., the theory that words are read directly from visual representations. The *graphemic-encoding hypothesis* would predict no effect from the manipulation of strictly phonemic properties.

In conclusion, although the *dual-encoding hypothesis* or the *phonemic-encoding hypothesis* seems likely to hold for categorization tasks as well as lexical decision tasks, there

are some reasons for expecting the effect of phonemic relations within pairs to be smaller for the categorization task than for the lexical-decision task. First, a large number of the test stimuli used by Meyer *et al*. (1974) were nonwords. Since people normally "sound out" printed words which they do not recognize (as in learning to read), the presence of such stimuli may have established a tendency to pronounce the strings silently. The phonemic encoding bias may become reduced when subjects see only English words throughout the experiment. In other words, the procedure of Meyer *et al*. may have elicited more than the normal amount (or intensity) of phonemic encoding.

Secondly, Meyer *et al*. (1974) pointed out that the presence of verbal context might influence the relative importance of visual *vs*. phonological representations. If context facilitates the recognition of words, the facilitation might affect graphemic and phonemic processes differentially. In particular, performance on tasks with context might depend more on visual representations than does performance on tasks with no context. This would be expected if strictly visual word recognition were conceived of as a process more direct than word recognition with a phonemic stage, but more dependent on facilitation by contextual information. The specification of the category in the present experiment provides semantically related context when the test word is an instance of the category. This is a small amount of context, but it

could conceivably reduce the role of phonemic encoding for Yes responses.

Method

Subjects

The subjects were 20 students attending the University of Massachusetts at Amherst summer session. Five of 20 original subjects were eliminated or replaced. One replacement was required because of an error made by the experimenter in selecting stimulus tapes. Four subjects were replaced because they produced either more than 6% errors or an average reaction time greater than 700 msec. These four subjects also appeared unusually tired and unmotivated at the time of the experiment.

Apparatus

The experiment was controlled by an HP2114B computer with a millisecond clock, connected to a display oscilliscope and a response panel with finger keys for the right and left hands.

Procedure

Subjects were run individually in a one-hour session which included a short instruction period, two practice blocks of 36 trials each, and six test blocks of 36 trials each.

After instructions were given, the subject was seated in a semi-darkened room facing the display scope. At the start of each block of trials the subject pressed a key to indicate readiness for the first trial. A trial consisted of the presentation of a category name, and the sequential presentation (and response to) the two words of a pair. At the start of each trial the subject saw the name of a category (e.g., ANIMAL) centered on the display for 1 sec. One sec. after the category name was removed two crosses, one above the other, appeared in the center of the scope. The subject was instructed to fixate on the space between the two crosses. The fixation crosses were presented for 500 msec. and then removed from the screen. The screen remained blank for another 500 msec., after which the first word of a pair was presented in the space between the two fixation points. The word was centered horizontally on the screen and subtended approximately 2° of vertical angle and 1° of horizontal angle per letter. The words used ranged from 3-8 letters in length.

The subject had to judge whether or not the word was a member of the specified category. A positive decision was indicated by pressing the "yes" key with the right index finger, and a negative decision was indicated by pressing a "no" key with the left index finger. As soon as the response occurred the first word was removed and the two crosses (fixation pointers for the second word) appeared again for 500 msec. The removal of the fixation points was again followed

by 500 msec. of blank screen, and the presentation of a word. The word was the second word of the pair assigned to the trial, and it appeared in the same position on the scope as the first word, and was centered horizontally.

The subject again had to decide whether or not the word belonged to the specified category. The same set of response keys was used for this decision, which removed the second word from the screen. Another trial began 1 sec. after the response to the second word. Subjects were instructed to respond as quickly and accurately as possible to each word. If an error occurred in response to either word of a pair on a given trial, the word "error" appeared in the lower left corner of the screen as soon as the trial ended, i.e., as soon as the response to the second word of the pair was made. The feedback added one second to the intertrial interval.

A different random order of presentation of pairs within trial blocks was used for each subject. Reaction time (RT) was measured from onset of each presented word to the key press response. Subjects received informal feedback on mean RT during rest periods of approximately 2 minutes, between trial blocks.

Materials

Seventy-two pairs of words were constructed for the practice blocks. These pairs were similar to the stimuli used in the test blocks, but were not used on any of the test

blocks. The entire set of test stimuli used in the experiment consisted of 432 pairs of words. The appendix is a list of these pairs. The pairs were chosen according to the phonemic and graphemic properties of their members. Only unambiguous members of the six categories (ANIMAL, FOOD, BODYPART, PERSON, SHELTER, CLOTHING) were selected as instances. Ambiguous noninstances, and pair members which were semantically related to each other were avoided. Phonemic similarity was defined as a rhyming relation between the members of a pair. Phonemically dissimilar pairs had non-rhyming members. Graphemically similar pairs had members which were spelled alike, but differed in their initial letters (e.g., BRIBE-TRIBE). In a few exceptional cases, however, the graphemically similar words differed in more than one letter (e.g., CLEAR-SWEAR). These exceptions are marked with asterisks in the appendix.

One fourth of the test pairs were graphemically and phonemically similar (e.g., HARE-MARE, MIGHT-TIGHT). These rhyming pairs were labeled _type 1_. Another fourth of the test pairs were graphemically similar but phonemically dissimilar (e.g., GULL-BULL, FAR-WAR). These nonrhyming pairs were labeled _type 3_. The _type 2_ and _type 4_ pairs were control pairs for types 1 and 3 respectively, and comprised the remaining two quarters of the test stimuli. In obtaining type 2 pairs, for instance, the second members of type 1 pairs (e.g., GOAT-BOAT, DEER-BEER) were interchanged to form

control pairs which are both phonemically and graphemically dissimilar (e.g., GOAT-BEER, DEER-BOAT). The pairs were interchanged randomly within semantic category and response type (e.g., Yes-Yes, Yes-No, etc.). The interchanging of type 1 pairs to obtain type 2 pairs was also subject to the restriction that no resultant control pair have members which are obviously related semantically, or which begin with the same letter. Type 4 pairs were derived from the type 3 pairs in the same manner as the type 2 pairs were derived from the type 1 pairs. Thus, half of the 432 test pairs were controls, which were made up of the same words as the rhyming and nonrhyming graphemically similar words, but were neither graphemically nor phonemically similar. No word appeared in more than one pair of a given type, and the words used in types 1 and 2 were not used in types 3 and 4.[2]

Table 1 presents examples of the four stimulus types,

Insert Table 1 about here

---

and a description of their graphemic and phonemic relations. The subdivision of stimulus materials according to category membership status is also shown. Since a category is specified on each trial, the correct response to each member of a

[2]The Yes-Yes pairs were exceptional in that members of type 1 and type 3 pairs were interchanged to form control pairs, as seen in the appendix. This procedure was necessitated by the scarcity of Yes-Yes pairs.

Table 1

Relative frequencies of Types of stimulus pairs, with respect to Category membership,* and Graphemic and Phonemic relations.

Type of Stimulus Pair	Graphemic Relation	Phonemic Relation	Examples	Correct Response	Relative Frequency
instance-instance (1)	similar	similar	HARE-MARE GOOSE-MOOSE	Yes-Yes	.028
instance-instance (2)	dissimilar	dissimilar	DONKEY-MOOSE GULL-MARE	Yes-Yes	.028
instance-instance (3)	similar	dissimilar	GULL-BULL DONKEY-MONKEY	Yes-Yes	.028
instance-instance (4)	dissimilar	dissimilar	GOOSE-MONKEY HARE-BULL	Yes-Yes	.028
instance-noninstance (1)	similar	similar	GOAT-BOAT DEER-BEER	Yes-No	.056
instance-noninstance (2)	dissimilar	dissimilar	GOAT-BEER DEER-BOAT	Yes-No	.056
instance-noninstance (3)	similar	dissimilar	BEAR-YEAR WOLF-GOLF	Yes-No	.056
instance-noninstance (4)	dissimilar	dissimilar	BEAR-GOLF WOLF-YEAR	Yes-No	.056
noninstance-instance (1)	similar	similar	VAT-RAT STOUT-TROUT	No-Yes	.056
noninstance-instance (2)	dissimilar	dissimilar	VAT-TROUT STOUT-RAT	No-Yes	.056
noninstance-instance (3)	similar	dissimilar	GASP-WASP LOW-COW	No-Yes	.056
noninstance-instance (4)	dissimilar	dissimilar	GASP-COW LOW-WASP	No-Yes	.056
noninstance-noninstance (1)	similar	similar	PITCH-DITCH MIGHT-TIGHT	No-No	.110
noninstance-noninstance (2)	dissimilar	dissimilar	PITCH-TIGHT MIGHT-DITCH	No-No	.110
noninstance-noninstance (3)	similar	dissimilar	COUCH-TOUCH COST-POST	No-No	.110
noninstance-noninstance (4)	dissimilar	dissimilar	COUCH-POST COST-TOUCH	No-No	.110

*The specified category for all example pairs in Table 1 is *Animal*.

pair is determined by whether or not the word is an instance of the category. Thus, there are four patterns of category membership within each of the four stimulus pair types, requiring four different response patterns. The four patterns of category membership (instance-instance, instance-noninstance, noninstance-instance, and noninstance-noninstance), their correct response patterns, and the relative frequency of each stimulus type x response type combination are also shown in Table 1. In sum, the test stimuli consisted of 16 groups of word pairs made from four stimulus pair types, in four yes/no (or instance/noninstance) combinations, referred to as response types.

One sixth of the total set of test pairs was assigned to each of the six categories: ANIMAL, FOOD, BODYPART, PERSON, SHELTER, and CLOTHING. The relative frequencies (as in Table 1) of stimulus types 1, 2, 3, and 4, and of the four response types were preserved within categories.

For purposes of assigning stimuli to subjects the set of 432 test pairs was divided into two subsets, _A_ and _B_. Subset _A_ consisted of those 216 pairs assigned to the categories ANIMAL, FOOD and BODYPART. Subset _B_ was made up of pairs representing PERSON, SHELTER and CLOTHING categories. The categories were randomly assigned to the subsets. Word length and frequency data are reported by subset, and the frequency data refer only to the second members of the pairs. The test words ranged from 3-8 letters in length. Words forming

graphemically and phonemically similar pairs had mean lengths of 4.4 letters (subset A) and 4.6 letters (subset B). The mean lengths of words forming graphemically similar but phonemically dissimilar pairs were 4.6 letters (subset A) and 4.9 letters (subset B). The modal word length was four letters, for each subset.

Word frequencies were obtained from the Kucera and Francis (1967) norms for American English. Frequencies of the test words (second members of pairs) ranged from less than one, to over 3,000 occurrences per million. The median frequencies were as follows: 16 (subset A) and 14 (subset B) per million for words forming graphemically and phonemically similar pairs, and 13.5 (subset A) and 15 (subset B) occurrences per million for words forming graphemically similar, but phonemically dissimilar pairs. Members of types 1 and 2 were thus equated approximately in average length and frequency with those of types 3 and 4.[3]

Design

All subjects were presented the same 72 pairs of words during the practice blocks. An incomplete block design was used to assign stimuli to subjects for the test blocks. The subjects were divided into two groups, Group 1 and Group 2.

---

[3]No further mention of subsets A and B is made, because a change in the experimental design eliminated the need to use them.

Group 1 was presented one half of the type 1 and type 3 pairs from each of the six categories. Group 2 was assigned the other half of the type 1 and type 3 pairs from the six categories. The reverse assignment was made for the control pairs. Group 1 was presented the type 2 and type 4 pairs derived from the type 1 and type 3 pairs which had been assigned to Group 2. Group 2 was assigned the other half of the type 2 and type 4 pairs. In this way, every type 1 and type 3 word presented to a subject was presented to a different subject in a control pair, and no subject saw the same word twice. Table 2 presents the assignment of stimulus pairs to Group 1 and Group 2 by semantic category, response type, and stimulus type. As shown in Table 2, each subject

---

Insert Table 2 about here

---

responded to 216 pairs, 54 of each stimulus type. The exact number of pairs within stimulus types, assigned to each group for each category and response type is given in the cell entries of Table 2. It can be seen that all subjects were assigned the same number of pairs from each category, within each response type and stimulus type.

Test blocks were constructed in accordance with the relative frequencies of the entire set. Each category occurred on one sixth of the trials of each block, and the percentage of positive responses required overall was 33%.

Table 2

Number of stimulus pairs of each type for the two groups of subjects, within category and response type*.

GROUP 1

Type of Stimulus Pair	ANIMAL				FOOD				BODYPART				PERSON				SHELTER				CLOTHING				Total
	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	
Type 1	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Type 2	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Type 3	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Type 4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Total	4	8	8	16	4	8	8	16	4	8	8	16	4	8	8	16	4	8	8	16	4	8	8	16	216

GROUP 2

Type of Stimulus Pair	ANIMAL				FOOD				BODYPART				PERSON				SHELTER				CLOTHING				Total
	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	YY	YN	NY	NN	
Type 1	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Type 2	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Type 3	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Type 4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	1	2	2	4	54
Total	4	8	8	16	4	8	8	16	4	8	8	16	4	8	8	16	4	8	8	16	4	8	8	16	216

*The response types are represented as follows: Yes-Yes (YY), Yes-No (YN), No-Yes (NY), and No-No (NN).

Results

Reduction of Data

Although subjects responded twice to each pair of words used in the experiment (once to the first word of the pair and once to the second), only the RT to the second word was included in the analysis.[4] Furthermore, the RT to the second word was included only when responses to both members of the pair were correct and the RT fell between 150 msec and 1500 msec. The analysis of variance was performed on mean RTs for the four stimulus pair types within category x response type cells. The means were based, not on individual subjects, but on pairs of subjects (SSs). Subjects from the two groups defined above under Design were paired as follows: each of the ten subjects from Group 1 was paired with a different subject from Group 2, yielding ten SSs. Since Group 1 had performed on half of the stimuli used in the experiment and Group 2 had responded to the other half, the pairing of a Group 1 subject with a Group 2 subject yielded an SS or composite subject which had RTs to all of the stimulus pairs.

Creation of SSs produced a complete block design (similar to that of Meyer et al., 1974), since for each datum provided by one of the subjects on a pair of phonemically and/or

---

[4]The RT to the first word was not of interest because it could not be influenced by the phonemic relationship between the two words. The second word was not presented until the response to the first had been made.

graphemically similar words (an _original_ pair), the other subject of the _SS_ provided the control datum. It was thus possible to conduct the experiment without having subjects serve as their own controls (which would necessitate responding to the same word more than once). Table 3 shows examples

Insert Table 3 about here

of the kinds of pairs (and their sources) which were presented to two complementary subjects to form an _SS_. It can be seen that a subject who performed on certain type 1 and type 3 pairs from subset _A_ of the stimuli (made up of ANIMAL, FOOD, and BODYPART categories) and certain type 2 and 4 (control) pairs from subset _B_ (PERSON, SHELTER, CLOTHING), was paired (at random) with a subject who performed on the corresponding type 2 and type 4 pairs from subset _A_ and the corresponding type 1 and type 3 pairs from subset _B_. Each _SS_, then, saw equal numbers of pairs of all types and had an _RT_ for each stimulus pair used in the experiment.

Within _SS_s a mean was obtained for each stimulus pair type x category x response-type cell. For example, one of the means for $\underline{SS}_1$ consisted of four _RT_s from each of the two subjects, obtained on the four type 1 (rhyming) pairs for which the category was ANIMAL and the response called for was No on both the first and second word of the pair. The four word-pairs were not the same for the two subjects, but

Table 3

Examples of stimulus pairs as they were combined from two subjects to make an SS. All examples are of the Yes-No type. Each column represents a set of stimuli for an individual subject, giving examples from three of the six semantic categories for each stimulus type.

SS$_i$

subject$_j$ (Group 1)			subject k (Group 2)		
Stimuli from subset A			**Stimuli from subset A**		
Type 1	Animal	GOAT-BOAT	Type 2	Animal	GOAT-BEER
		DEER-BEER			DEER-BOAT
	Food	STEW-CREW		Food	STEW-GUTTER
		BUTTER-GUTTER			BUTTER-CREW
	Bodypart	BACK-TACK		Bodypart	BACK-REIN
		VEIN-REIN			VEIN-TACK
Type 3	Animal	BEAR-YEAR	Type 4	Animal	BEAR-GOLF
		WOLF-GOLF			WOLF-YEAR
	Food	WAFFLE-BAFFLE		Food	WAFFLE-BUDDING
		PUDDING-BUDDING			PUDDING-BAFFLE
	Bodypart	BONE-DONE		Bodypart	BONE-WAND
		HAND-WAND			HAND-DONE
Stimuli from subset B			**Stimuli from subset B**		
Type 2	Person	KING-BROOK	Type 1	Person	KING-RING
		CROOK-RING			CROOK-BROOK
	Clothing	VEST-FLIRT		Clothing	VEST-TEST
		SHIRT-TEST			SHIRT-FLIRT
	Shelter	HALL-WENT		Shelter	HALL-FALL
		TENT-FALL			TENT-WENT
Type 4	Person	FATHER-WARMER	Type 3	Person	FATHER-RATHER
		FARMER-RATHER			FARMER-WARMER
	Clothing	HOOD-DROVE		Clothing	HOOD-MOOD
		GLOVE-MOOD			GLOVE-DROVE
	Shelter	CAVE-LOWER		Shelter	CAVE-HAVE
		TOWER-HAVE			TOWER-LOWER

the independent variables had the same values. Although individual RTs were sometimes eliminated due to errors or to exceeding the high or low RT cutoffs, there were sufficient data remaining in every case to provide an SS mean for each of the 96 cells for each SS.

Analysis of Variance Factors

The analysis of variance factors within SS were homophony, pairing, response type and category. The type of stimulus pair—whether a pair was type 1, 2, 3, or 4—was defined by two factors: homophony and pairing, each with two levels. Homophony distinguished stimuli of types 1 and 2 from stimuli of types 3 and 4. The stimulus materials consisted originally (before control pairs were made) of a set of pairs whose members rhymed and looked alike (type 1) and a set of pairs whose members looked alike but did not rhyme (type 3). One level of homophony is made up of rhymes plus the control pairs derived from them; the second level is made up of all the original non-rhymes plus the control pairs derived from them. The homophony factor thus distinguishes two completely different sets of words used in the experiment. It might be described as: phonemically-and-graphemically-similar-pairs-and-their-controls *vs.* graphemically-similar-but-phonemically-dissimilar-pairs-and-their-controls.

The second factor, pairing, refers simply to whether an item was one of the original pairs or a control pair. Pair-

ing distinguishes stimuli of types 1 and 3 from stimuli of types 2 and 4. Pairs of types 2 and 4 were constructed by randomly interchanging the second members of pairs of type 1, and then those of type 3, to make graphemically and phonemically dissimilar pairs. Taken together, the factors homophony and pairing produce a 2 x 2 factorial design whose cells are types 1-4, as shown in Table 4. As explained earlier in terms of differences in RT between stimuli of types 1 and 3, relative to their controls, it is the interaction between homophony and pairing which is of major interest in the present results.

Insert Table 4 about here

A third variable, response type, was the four possible combinations of correct responses to the two words of a pair. Because many of the Yes-Yes pairs could not be constructed to meet the criteria for graphemic and phonemic similarity and category membership they were eliminated from the analysis of variance. Three levels of response type were included —Yes-No, No-Yes, and No-No.

The fourth factor, category, was a random effect variable, and was composed of the six semantic categories used in the experiment. The six levels of this factor could also be assigned stimulus pairs at random instead of by semantic category, for purposes of determining the reliability of ex-

Table 4

The four types of stimulus pairs shown as a 2 x 2 factorial design with homophony and pairing as factors

Homophony	Pairing: (a) Original pairs	Pairing: (b) Control pairs
(a) Graphemically and phonemically similar pairs	Type 1 GOAT-BOAT DEER-BEER	Type 2 GOAT-BEER DEER-BOAT
(b) Graphemically similar but phonemically dissimilar pairs	Type 3 BEAR-YEAR WOLF-GOLF	Type 4 BEAR-GOLF WOLF-YEAR

perimental effects over random subsets of words (cf. Meyer *et al.*, 1974; Clarke, 1973). When this was done this factor was referred to as *words* instead of *categories*.

Data Analysis: Reaction Times

The mean *RT*s and error rates for the four types of stimuli at each of the response types are presented in Table 5.

Insert Table 5 about here

More detailed data appear in Table 7, which provides *RT* means within semantic categories. The average *RT* over all conditions and *SS*s was 593 msec. The results of the analysis of variance performed on the *RT* data from the ten *SS*s appear in Table 6.[5] The ten subject pairs produced mean *RT*s which dif-

---

Insert Table 6 about here

---

fered significantly from one another, $F\ (9,45) = 65.66$, $p < .001$. Although the task thus yields individual differences in *RT* they seem to occur rather consistently across experimental treatments. The only significant interaction observed, involving subjects, was *SS* x response type. This interaction

[5] Several of the effects involving response type, pairing and homophony were tested using quasi-F ratios. These test statistics and their degrees of freedom were calculated as suggested by Myers (1966), pp. 281-283.

Table 5

Mean RT and proportion of errors of ten SSs on the four types of stimuli and four response types, collapsed over semantic category.

		Pairing					
		Original			Control		
Response type	Homophony	Pair type	Mean RT	Propor. errors	Pair type	Mean RT	Propor. errors
Yes-Yes	Phonem. similar	1	482	.09	2	479	.01
	Phonem. dissimilar	3	485	.13	4	559	.07
Yes-No	Phonem. similar	1	593	.03	2	595	.03
	Phonem. dissimilar	3	625	.04	4	630	.03
No-Yes	Phonem. similar	1	556	.08	2	576	.05
	Phonem. dissimilar	3	574	.09	4	578	.11
No-No	Phonem. similar	1	576	.02	2	592	.02
	Phonem. dissimilar	3	610	.006	4	606	.01

Table 6

Analysis of Variance of *RTs* for subject pairs (*SS*), Response Type (*R*), Pairing (*P*), Homophony (*H*) and Category (*W*)

Source of Variance	Degrees of Freedom	Mean Square	Error Term MS	Error df	F	p
SS	9	161,226.1	SSW	45	65.66	< .001
R[a]	2	97,231.1	SSR+RW-SSRW	9	3.26	<.10
P	1	9,464.9	SSP+PW-SSPW	6	0.98	>.20
H	1	88,877.8	SSH+HW-SSHW	2	9.68	<.10
W	5	11,790.3	SSW	45	4.80	<.005
SSR	18	9,777.4	SSRW	90	3.85	<.001
SSP	9	4,353.1	SSPW	45	1.43	>.20
RP	2	1,143.7	SSRP+RPW-SSRPW	7	0.15	> .20
SSH	9	885.0	SSHW	45	0.17	>.20
RH	2	8,232.9	SSRH+RHW-SSRHW	10	0.72	>.20
PH	1	5,150.4	SSPH+PHW-SSPHW	6	0.47	> .20
SSW	45	2,455.6				
RW	10	22,612.6	SSRW	90	8.91	<.001
PW	5	8,315.0	SSPW	45	2.73	<.05
HW	5	13,654.0	SSHW	45	2.55	<.05
SSRP	18	3,369.9	SSRPW	90	0.96	>.20
SSRH	18	2,704.0	SSRHW	90	1.03	>.20
SSPH	9	4,721.4	SSPHW	45	1.37	>.20
RPH	2	2,341.6	SSRPH+RPHW-SSRPHW	7	0.26	>.20
SSRW	90	2,536.8				
SSPW	45	3,041.9				
RPW	10	7,698.0	SSRPW	90	2.20	<.05
SSHW	45	5,356.3				
RHW	10	11,348.6	SSRHW	90	4.31	<.001
PHW	5	9,777.7	SSPHW	45	2.83	<.05
SSRPH	18	3,466.3	SSRPHW	90	0.73	>.20
SSRPW	90	3,492.8				
SSRHW	90	2,630.2				
SSPHW	45	3,455.4				
RPHW	10	10,359.4	SSRPHW	90	2.19	<.05
SSRPHW	90	4,729.1				

[a]Three levels of *R* were included in this analysis: Y-N, N-Y, and N-N.

is probably due to variation among subjects in strategic reactions to the response-type contingencies in the stimulus materials (e.g., that No responses were called for after No and Yes (first word) responses, more often than were Yes responses).

The main effect of homophony was nearly significant, *F* (1,2) = 9.68, p < .10. This indicates that type 1 (rhyming) pairs and their controls, taken together, elicited shorter *RTs* than the type 3 pairs and their control stimuli. There was no main effect of pairing, *F* (1,6) < 1. Thus the *RTs* on type 1 and type 3 pairs combined did not differ significantly from the control pairs (types 2 and 4) combined. Considering homophony x pairing cells, there was essentially no difference between the response time to phonemically dissimilar (type 3) pairs and their controls (type 4). This is the comparison in which Meyer *et al*. (1974) found the large inhibition effect of nonrhyming pairs whose members looked alike. Also contributing to the major interaction observed by Meyer *et al*. was a 20 msec. (nonsignificant) facilitation effect of rhyming (type 1) pairs relative to their controls. The facilitation effect of phonemically similar pairs in the present study was 13 msec., which was not significant, *F* (1,6) = 1.27, p < .20. In sum, the manipulation of phonemic relationships within pairs of words produced no significant differences in *RT* (over *SS*, response type, and category) to the second words of experimental pairs, relative to control

pairs.

The homophony by pairing interaction, which would indicate differential effects of phonemically similar and dissimilar pairs (relative to their controls) was thus quite small, and not significant (*F* < 1). The total interaction, i.e., the difference between type 2 and type 1 *RTs* plus the difference between type 3 and type 4 *RTs*, was 11 msec. This interaction appears in Figure 3, which shows the means for the four types

Insert Figure 3 about here

of stimulus pairs, for both *RT* and proportion of errors. These interactions do not closely resemble those observed by Meyer *et al.*, and which were predicted in this experiment.

The effect of response type was close to significant, $F(2,9) = 3.26$, $p < .10$. As Table 6 indicates, response type interacted significantly with *SS*, category, and with various combinations of pairing, homophony and category. It did not, however, interact with pairing x homophony, and as Table 5 shows, there were only a few slight instances within the various levels of response type, in which the results occurred as predicted. The *RTs* for these cases were analyzed by two post hoc contrasts which were performed to test the facilitating effect of phonemically similar pairs within the No-Yes and No-No response types. For No-Yes pairs the 20 msec. difference between rhyming pairs and their control

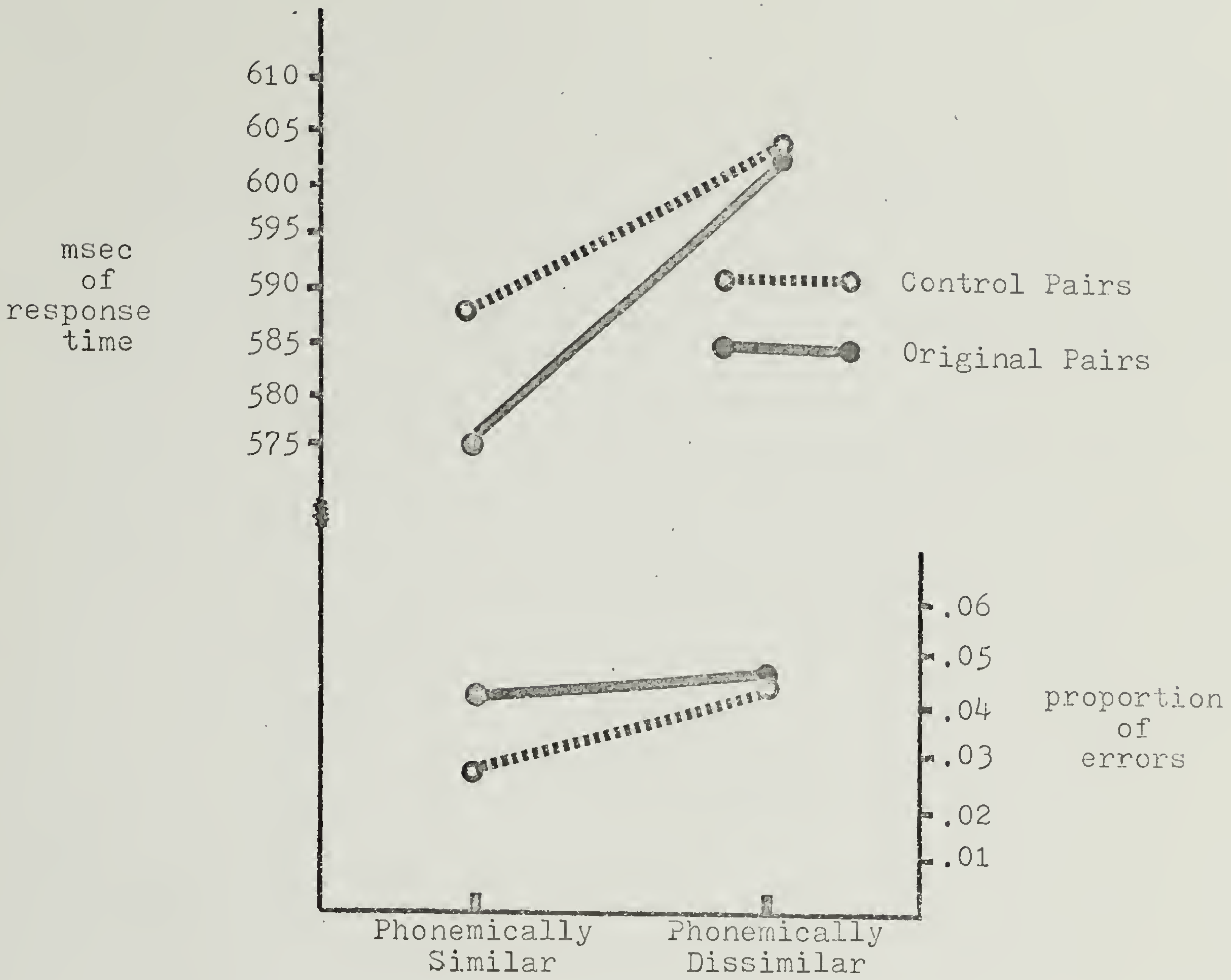

Figure 3. Mean reaction time and proportion of errors of 10 SSs to phonemically similar and dissimilar pairs and control pairs; the interaction of homophony and pairing, collapsed over response type, category, and SS.

pairs was not significant, F (1,7) = 1.32, $p > .20$. For the No-No response category, since the difference between non-rhyming pairs and their controls was also in the predicted direction, the interaction between homophony and pairing was tested using both phonemically similar and phonemically dissimilar pairs and their controls. This interaction was not significant, F (1,7) = 1.32, $p > .20$. Thus, the overall lack of effect of the manipulation of phonemic relations in pairs of words held true within the different response combinations as well as across them.

The main effect of category on RT was significant, F (5,45) = 4.80, $p < .005$. The overall means for the six categories are presented in Table 7 along with the pairing x homophony x response-type means for each of the six categories. As Table 6 shows, category interacted significantly

---

Insert Table 7 about here

---

with every possible variable or combination of variables in the experiment. Table 7 shows that the interaction of category with pairing x homophony was such that two categories, CLOTHING and SHELTER, resulted in pairing x homophony interactions in the expected direction, of 75 and 25 msec., respectively. These interactions were not, however, consistent over response types within categories, so there is no strong indication that the phonemic effect was present, but

Table 7

Mean RTs of the ten SSs for the six categories, and for the four stimulus pair types at each response type and category level, in msec.

Homophony		-	Phonemically Similar		Phonemically Dissimilar	
Pairing		-	Control	Original	Control	Original
Category	Overall Mean	Response Type				
Animal	584	Y-N	561	605	632	611
		N-Y	569	517	589	542
		N-N	589	585	612	593
		mean	573	569	611	582
Food	600	Y-N	591	622	714	603
		N-Y	574	535	581	611
		N-N	596	566	583	622
		mean	587	574	626	612
Bodypart	578	Y-N	536	553	602	626
		N-Y	571	567	626	634
		N-N	543	549	565	571
		mean	550	556	598	610
Person	592	Y-N	600	603	606	590
		N-Y	577	623	546	566
		N-N	581	596	604	616
		mean	586	607	585	591
Clothing	595	Y-N	711	589	612	646
		N-Y	549	523	557	563
		N-N	601	543	635	616
		mean	620	552	601	608
Shelter	605	Y-N	571	589	613	671
		N-Y	616	572	568	525
		N-N	642	617	635	641
		mean	610	592	605	612
Overall	593					

emerged only under some semantic categories. The fluctuations may be due to differential effectiveness of specific stimulus pairs which may have clustered by chance in a few category by response type cells. In other words, it is suspected that there is nothing intrinsically special about these two categories, but they may have been populated by superior specimens of the types of stimuli we created, if in fact the fluctuations referred to are not random and are due to phonemic effects.

A second analysis of variance was carried out using random subsets of words as a factor instead of semantic categories, in order to demonstrate that the results were reliable over word pairs. This was accomplished by designating the pairs presented on a given trial block as a subset or level of the factor words, resulting in six subsets of stimulus pairs which had been randomly assigned to their levels, with certain restrictions (see Method section for assignment of pairs to trial blocks). The pattern of results of the second analysis was the same as that of the first. The main effects homophony, response type and SS were again significant. Words was also significant, apparently due to the general decrease in RT over trial blocks. The pairing x homophony interaction was again not significant, but the F value was much larger (F (1,6) = 2.03). Words tended not to interact with other variables as category had done, but the pairing x homophony x words interaction approached significance,

$\underline{F}$ (5,45) = 2.01, $p < .10$. The words analysis demonstrates the reliability of the basic results over random subsets of words. It would have been more informative had positive results been obtained.

The proportions of errors on responses to the second words of stimulus pairs for response-type x homophony x pairing cells are presented in Table 5. The overall error rate was low, less than 5%. Somewhat different patterns of errors were found for different response categories. As Table 5 shows, the error rates for the No-Yes response type were consistently higher than those for the Yes-No and No-No response types, which were quite low (between .6% and 4%). The higher error rates on No-Yes trials suggest that the somewhat lower RTs for this response type can be accounted for by speed/accuracy tradeoff. It is not surprising that this would occur for No-Yes trials and not for Yes-No and No-No trials, because stimuli which required the No-Yes sequence of responses were relatively infrequent. Two thirds of the stimuli requiring a No response to the first word of the pair required a No response to the second word. A Yes response to the second word, moreover, was called for on only one third of the trials, overall. The higher error rate for No-Yes trials was thus to be expected as a reflection of subjects' expectancies based on the relative frequencies of stimuli of various response types. The error rates on the Yes-Yes trials were highly variable. It is impossible to interpret

these proportions of errors due to the small number of Yes-Yes trials and the unusualness of the items used on these trials (with respect to their semantic, graphemic and phonemic properties).

The error rates within homophony x pairing cells are of interest because differences which occur there may indicate predicted differential difficulty of the four stimulus pair types which was not observed in the RT data. Figure 3 shows the proportions of errors for the type 1-type 4 stimulus types. A small interaction is apparent, contributed by the relatively low error rate for type 2, relative to type 1 pairs. This difference is small, but points to a speed/accuracy tradeoff interpretation of the small rhyme facilitation effect observed in the RT data. This interaction is in the opposite direction from the predicted phonemic effect.

Returning to Table 5, which presents error rates within response types, the proportions of errors within the Yes-No and No-No response types are seen to be quite low and homogeneous. They do not indicate differential difficulty of the four stimulus pair types. Within the No-Yes response type the phonemically similar pairs produced a higher error rate (8%) than their control pairs (5%). This result accounts for the overall interaction in the error proportions of Figure 3, discussed above. It is contrary to the prediction of facilitation of rhyming pairs, and again suggests a speed/accuracy tradeoff—for the 20 msec. facilitation observed in the RTs

of this response type. This was the largest rhyme facilitation effect on RT observed within response type. The only remaining suggestion of a phonemic effect is the 16 msec. (nonsignificant) facilitation of rhyme pairs on No-No trials.

A small difference in error rate was also found between the phonemically dissimilar pairs (9%) and their controls (11%), of the No-Yes response type. There is no evidence here of a speed/accuracy tradeoff. The direction of the difference, however, is contrary to that predicted on the basis of phonemic effects. If graphemically similar nonrhyming pairs had an inhibiting effect, relative to their controls, the error rate would be higher for the original pairs than for the graphemically dissimilar controls.

In summary, the proportions of errors of rhyming and nonrhyming graphemically similar pairs differed very little from those of their control pairs. The differences which occurred involved error rates of the No-Yes response type. Neither facilitation of type 1 pairs compared to type 2 pairs nor inhibition of type 3 pairs compared to type 4 pairs characterized the error data.

Supplementary Analyses

A supplementary analysis was conducted to insure that any major discrepancy between the present results and those of Meyer et al. (1974) was not due simply to differences in the stimulus words used in the two experiments. Since some

of the stimuli used in the present experiment were taken from lists of those used by Meyer et al., it was possible to find the mean RTs in the present data which corresponded to the subset of stimulus pairs which was common to the present experiment and to the Meyer et al. experiment. Although 54 such pairs of types 1 and 3 were available (33 of type 1; 21 of type 3), the corresponding type 2 and type 4 pairs differed from those used in the Meyer experiment. This happened because we did not make the same random interchanges of the second members of the original pairs (to produce control pairs) as did Meyer et al. In the following results, then, the RTs for type 1 and type 3 pairs are based on stimulus pairs actually used by Meyer et al., but the control pairs did not have the same first word and often had a first word extraneous to the Meyer et al. list. Furthermore, all of these pairs were of the Yes-Yes response type in Meyer et al.'s study, whereas in our experiment they were almost all used for No-No responses. The mean RTs for the 51 N-N pairs of this subset of items were: type 1, 561 msec.; type 2, 590 msec.; type 3, 619 msec.; and type 4, 598 msec. The facilitation effect for rhymes was 29 msec., the inhibition effect for phonemically dissimilar pairs was 21 msec. and the interaction between pairing ahd homophony was 50 msec. These effects were substantially larger than those produced with the full set of word pairs.

An analysis of variance was performed on the mean RTs

for SSs, collapsed over categories. The pairing x homophony interaction in these data was significant, $F(1,9) = 6.01$, $p < .05$. Thus, the phonemic properties of these pairs had an effect when graphemic similarity was held constant. This finding was upheld by the pattern of errors in the four pairing x homophony cells. The facilitation of type 1 pairs relative to their controls appeared in the error rates; the proportion of errors for type 1 pairs was .012, while for type 2 pairs it was .027. Inhibition of response by nonrhyming graphemically similar pairs did not emerge as clearly in the error data for this subset of stimuli. But there was a very slight difference in the right direction. The proportions of errors were .022 for type 3 pairs and .017 for type 4 pairs, based on 18 stimulus pairs of each type. In general, the occurrence of errors was rare for the Meyer *et al.* subset of our stimulus pairs. The error data are thus too scanty to be highly informative. Nevertheless they might have rendered the RT data ambiguous, and they did not. The phonemic effect predicted in this experiment was thus observed in the data from this 51-pair subset of the stimuli.

Discussion

In their research using the lexical decision task Meyer *et al.* (1974) found evidence for a dual-encoding theory and phonemic-encoding theory, as opposed to a graphemic-encoding

theory of visual word recognition. These theories hold that phonological representations play a role in visual word recognition. They were supported by Meyer et al.'s finding that phonemic properties of words affect recognition response times even when graphemic properties are held constant. That is, pairs of non-rhyming words which looked alike took longer to recognize (relative to control pairs), than did pairs of rhyming words (relative to controls) which also looked alike. Phonemic similarity had an effect independent of graphemic similarity. In the present study it was hypothesized that this finding would generalize to a task involving access of word meaning—the semantic categorization task. Specifically, it was predicted that categorization RT (to the second member of a graphemically similar pair of words) would be greater for non-rhyming pairs relative to graphemically dissimilar controls, than for rhyming pairs relative to graphemically dissimilar controls, across various semantic categories and Yes-No response combinations. In our design this effect would be manifested as an interaction between homophony (whether a pair was derived from a rhyme or non-rhyme set) and pairing (whether a pair was graphemically similar or a control pair). The expected interaction was not found. Although the differences among the cell means occurred in approximately the same pattern as in Meyer et al., both these differences and the key interaction were very small, and the results of the corresponding statistical tests did not approach

significance.

A substantiated negative finding from this experiment would cast doubt on the phonemic and dual-encoding theories as they apply to the process of deriving meaning from written words. It would suggest that when semantic processing is required, the role of phonological representations in reading printed words is insignificant. Such evidence would bear not only on hypotheses about phonemic encoding in meaningful word processing, but also on assessment of the appropriateness of findings based on the lexical-decision task for use in building theories of normal reading. There are, however, a number of differences between the experimental procedures used here and those used by Meyer _et al_. (1974). These arose necessarily in the application of Meyer's paradigm to the semantic task, but are not inherent in the distinction between lexical decisions and semantic categorization. One or more of these differences might have accounted for the difference in results. We have, in fact, discovered that one such factor—the specific stimulus pairs used—played a significant role in producing the discrepancy in findings between the two experiments. Our analysis of that factor will be discussed below, following a brief consideration of other procedural differences which might bear on comparison of the present results and those of Meyer _et al_.

The most salient feature differentiating our procedure from that of Meyer _et al_. (1974) is the use of words _vs_.

nonwords. Meyer et al. used a large proportion of nonwords because their subjects were to judge whether or not a string of letters was a word. The semantic decision task called for words as noninstances of the categories. Since Meyer et al.'s phonemic encoding effect was established on data from the real words used in their experiment there is no reason to expect the switch to real words as noninstances as well as positive instances to influence the results substantially. But the possibility (mentioned earlier) remains that the phonemic effect in responses to the words in the lexical-decision task was induced by generalization of subjects' tendency to attempt to pronounce (to sound out silently) unfamiliar letter strings. We have no evidence that this is the case, but if it were, elimination of nonwords from the stimuli would eliminate or reduce the phonemic encoding effect.

Another procedural difference between the two experiments was in the type of response made by the subject in producing the informative portion of the data. In testing their hypotheses Meyer et al. used RTs based on Yes responses to the second members of stimulus pairs which followed Yes responses to first members. All No responses were made to nonwords, which are ambiguous with respect to pronunciation and thus rule out specification of phonemic similarity and dissimilarity. In the present experiment the difficulty of finding words with particular phonemic and graphemic properties within a small number of semantic categories made it

impossible to use many Yes-Yes pairs, and those that were used tended to be rather peculiar (e.g., CURD-LARD as graphemically similar foods). It was necessary to use No-No pairs in obtaining one half of the analyzable data; and the RT to the second member of a pair (the dependent variable) was based on a No response in two thirds of the data. Thus, the findings of the lexical-decision study were based on Yes responses, while the findings of the semantic categorization study were based primarily on No responses.

There may be some aspect of the process of responding No on the categorization task which would hide the phenomenon under study. If, as Meyer et al. note, the phonemic encoding bias is time dependent, then the latency to the first member of a stimulus pair might determine whether or not the phonemic relationship between the stimuli has any effect. The biasing effect may be less likely to occur the longer the reaction time to the first word of the pair, since this time separates the successive encoding operations. Since No responses take more time than Yes responses, the encoding bias would then be less likely to occur on No-No and No-Yes trials than on Yes-No trials, because the first words of those trials call for No responses. But the smallest pairing x homophony interaction occurred on the Yes-No trials. It is therefore unlikely that making No responses to the first words of pairs eliminated the encoding bias. Although no clear argument is available it might also have been possible that No responses

to the second word of a pair would not show the phonemic effect. The only trials which do not involve No responses are the Yes-Yes trials, for which we have no useful data. The response-type problem would present a major difficulty in interpreting our results, had we not observed the predicted phonemic effect on a subset of the No-No trials. It apparently is possible to obtain phonemic encoding bias using negative responses.

The incidental characteristics of the specific task used to investigate semantic processing must also be considered as factors which might be responsible for the negative findings. Is there any aspect peculiar to the categorization task which might have degraded the influence of phonemic properties of words? The task may call for or initiate some cognitive activity (in addition to the intended semantic processing) which tends to interfere with the encoding bias presumed to underlie the Meyer et al. effect. Observation of the effect is apparently made possible by a dependence between successive phonemic encoding operations. This dependence may hold only when those operations are strictly contiguous temporally. If so, any interpolated activity (between onset of the response to the first word and reading of the second word) would disrupt the encoding bias. Interpolated activity seems more likely to occur on the categorization task (as employed here) than on the lexical decision task because the subject must keep in mind which of six categories

is relevant on any given trial. Some subjects in fact spontaneously reported having rehearsed (some time between the two responses to a pair) the name of the category specified for the ongoing trial. Since the category was subject to change on every trial, this possibility could have been reduced by presenting all items associated with a given category on one trial block.[6] A large number of categorizations with no change in category would no doubt minimize the tendency subvocally to rehearse the category name during the trials. Although it is not possible to assess the occurrence or effects of such rehearsal given the procedure used, the emergence of the phonemic encoding effect on the subset of pairs also used in the Meyer et al. experiment indicates that this factor does not operate consistently or strongly enough to abolish it completely.

Meyer et al. (1974) pointed out that tasks used to study the processing of words vary in the amount of context available to the subject, and that this factor could influence the relative importance of visual vs. phonological representations. The semantic categorization task provides a bit more context than does the lexical-decision task, in that (for positive instances) the category term itself is semantically related to the isolated word responded to. If this amount of context had a demonstrable influence in lessening the

---

[6]The categories were interspersed in order to avoid priming effects.

phonemic effect, it would imply that such effects would be negligible in normal reading. The pattern of results observed, however, does not indicate context effects. Trials involving Yes responses (to the second word), for which context related to the test word is present, did not provide evidence for a role of phonological representations. This would occur if context lessened the phonemic effect. Trials involving No responses, however, for which context related to the test word is not present, also failed to show phonemic encoding. So no evidence was found that semantic context influenced the amount of phonemic processing. Unfortunately the effect of context cannot be tested on the Meyer *et al.* words, since all but three of the pairs were No-No stimuli in this experiment.

A final possibly interfering characteristic of the categorization task is the variety of semantic relations between the category term and the stimulus pairs which must be classified discreetly as membership or nonmembership. The difficulty of finding pairs of words (especially those involving category members) which looked alike but did not rhyme led to the use of some unusual words and instances of categories, even for response types other than Yes-Yes. The large main effect of homophony may reflect this problem. It is well established that typicality of subordinates affects categorization time (Collins and Loftus, 1975). Having to decide whether *tower* is a shelter or whether *budding* is a food might

have given the subjects pause, and time contributed by distraction would be likely to mask the phonemic effect.[7] Since it was the nonrhyming set of stimuli which tended to be unusual, the fact that the data showed some trend toward a rhyme facilitation effect, but no evidence of a nonrhyme inhibition effect implicates the typicality problem. This is more convincing because in the Meyer _et al_. study it was the inhibition effect which was the larger, and statistically significant. It may be possible to explore this possibility through item analysis.

A second semantic factor which might have obscured the phonemic encoding effect by distracting subjects is related to the frequent change of category in this experiment. It is possible that categorization judgments interfered with one another as follows. Despite efforts to avoid such occurrences, members of a category did sometimes appear as noninstances of another category. These were usually vaguely possible members or homophones of possible members. For example, _dwarf_ appeared as a noninstance of clothing and _dear_ appeared as a noninstance of food. If, for a given subject, these items appeared shortly after a person or animal trial, respectively, they might have had a distracting influence which could overpower the inhibition or facilitation of encoding

[7]It is also possible that the difficulty of these pairs made them more susceptible to graphemic facilitation, which counteracted phonemic inhibition.

bias. Although it is difficult to establish clear selection criteria for these semantic phenomena, an attempt is being made to isolate them for the purpose of further analysis of the data remaining after their removal.

The factors described above should be kept in mind in comparing the results of the present experiment with those of Meyer et al. (1974). There were, in addition, some procedural differences in administration of the tasks which also deserve mention, but seem less important to the interpretation of the findings. The experiments differed in the type of subjects, the incentive conditions employed, the use of white background noise (by Meyer et al.), and the arrangement and fingering of response keys. There were also minor differences in the timing of stimulus presentation and feedback.

The one difference between this experiment and Meyer et al.'s for which we have a clear indication of a significant role in producing the discrepancy in results was the specific stimulus pairs used. When those stimulus pairs used in Meyer et al.'s experiment, as well as in the present experiment, were analyzed separately from the total set, the interaction predicted on the basis of phonemic encoding bias was found to be significant. This finding indicates that the phonemic encoding effect does extend to the semantic task, but that it depends on the words used or on the particular relationships within the particular pairs. Since the phenomenon studied then appears to generalize across tasks but not

across stimulus materials, it is important to consider carefully the characteristics of the pairs used by Meyer et al. in contrast to those of our general pool. This may explain the discrepancy in results, and illuminate the nature of the phenomenon as well. We can begin by considering those aspects of the pairs that Meyer et al. pointed out as possible factors for further research. We will also try to discover whether any of the factors described above (especially the semantic factors) might interact with the set of words used to reduce the role of phonological representations.

Meyer et al. (1974) mentioned two factors which should bear on the effects of graphemic and phonemic relations between words on recognition. The first is the extent to which the word follows the grapheme-phoneme correspondence rules of English. If a word is quite anomalous in this respect it may be processed by direct visual recognition. This factor will not be considered in the attempt to understand why the words used in this experiment did not produce the predicted effect, because the vast majority of the words used follow grapheme-phoneme correspondence rules, as the writer understands them.

The second factor may well have played a role in determining the negative results. Meyer et al. stated that their encoding bias model predicts that,

> the effects of graphemic and phonemic relations on recognition may be influenced by word order. For example, suppose that the word BLOW is processed immediately before the word PLOW. Here the more

> common pronunciation of the LOW-ending occurs in the first word. This fact, together with the graphemic similarity of the two words, may bias the S sufficiently that he always initially encodes PLOW to rhyme with BLOW. As a result, it would take longer to recognize PLOW in the above pair than in the graphemically dissimilar pair like LEMON-PLOW, where there is less bias toward the wrong encoding. In contrast, suppose that PLOW is processed before BLOW. Here BLOW might not take longer to recognize than a graphemically dissimilar word, since a bias toward applying the more common grapheme-phoneme correspondence rules could overcome an erroneous bias toward rhyming. Thus, the difference between RTs for graphemically similar and dissimilar words should depend on both the phonemic relation and the order of presentation.

Meyer *et al.*'s "informal examination" of their stimuli suggested to them that a majority of the second members of their nonrhyming, graphemically similar (type 3) pairs had less common pronunciations than the corresponding first members. If the order of these words were reversed, they would expect lower RTs relative to control pairs.

Informal examination of the type 3 N-N pairs used in this experiment, but not in Meyer *et al.*'s experiment indicates that about 58% are pairs for which either the pronunciation of the second word is more typical for the spelling pattern, or the pronunciation of the two words seemed about equally common (the writer could not decide which was more common). A possible contribution to the failure to observe the nonrhyme inhibition effect, then, was the frequency of pairs used which are not clear cases of more common pronunciation of the first word than the second word.[8]

In perusing the stimulus pairs it was noticed that three other potential problems existed in the type 3 N-N stimuli used exclusively in this experiment: 1) a few of the pairs had members which differed very little in pronunciation (e.g., VASE-BASE, LEASE-TEASE);[9] 2) two pairs were less similar graphemically than the criterion of differing first letters only (CLEAR-SWEAR, TREAD-PLEAD); 3) there were a number of pairs whose first or second words may have been strange enough to elicit long RTs regardless of whether they appeared in a type 3 pair or a control pair. Inhibition effects on RT

---

[8]It should be noted, however, that it would not be unreasonable to expect the ordering of the words in a pair to have an effect opposite to that proposed by Meyer *et al*. They suppose that the first word must have the more common pronunciation in order to effectively bias encoding of the second word. One might suppose that the *second* word must have the more common pronunciation in order to effect encoding bias. This would be true if subjects had a strong tendency to apply the more common grapheme/phoneme correspondence rules to all words. If this were the case the incorrect rules would be applied to the uncommon words in control pairs as well as graphemically similar pairs. The subject's preexisting bias might obscure any experimentally induced bias. When the first word was of uncommon pronunciation, however, the subject would not already be biased to pronounce the test word incorrectly and the encoding of the first word according to unusual rules could induce a bias which was not already present and which would thus not occur on control pairs. The inhibition effect would thus occur for the pairs which had the more commonly pronounced word (given the spelling) second in order. Since an argument can be made for either order, it would be helpful to introduce order of the words in pairs as an experimental variable and examine its effects.

[9]Cases of pairs which seemed to the writer to differ little in pronunciation tended to share the characteristic that the difference in pronunciation between the two words was based on consonant rather than vowel sounds.

might have been washed out by the distraction value of pairs such as LITER-MITER, BOOTY-SOOTY, FETAL-METAL, and FREIGHT-SLEIGHT, which were resorted to as the language was exhausted. Unusual first words could have a complicated effect if they influenced the RTs to the control words with which they were paired, and which may have been controls for normally effective pairs. Instances of these three problems, taken together, account for about another 17% of the N-N stimuli not used by Meyer et al. Three quarters of these pairs thus seem to have some deficiency relative to the ideal type 3 pair. Although the judgments of the above factors were quite subjective the percentages are large enough to make the failure of the phonemic effect to generalize to our stimulus materials comprehensible.

Conclusions

The findings of the present study provide some evidence that the phonemic encoding effect observed by Meyer et al. (1974) extends to the semantic categorization task. But since the positive finding was limited to a subset of the stimulus materials used, it must be substantiated by further experimentation. In the light of the factors discussed above it seems likely that a more sensitive experiment could be designed, combining procedural changes such as blocking of category trials with more stringent criteria for selection of stimulus pairs. It should thus be possible to demonstrate

phonemic effects on the categorization task, using stimuli other than the specific pairs used by Meyer _et al_. (1974).

At present we may tentatively conclude that phonemic properties of words do affect reaction time (over and above graphemic effects) when subjects make category judgments. As mentioned earlier this finding is interesting because semantic categorization has something in common with normal reading which lexical decisions do not necessarily entail—access of specific semantic information associated with a word. As such it is stronger support for a phonemic- or dual-encoding theory of word recognition in reading than is the finding based on the lexical decision task.

This stronger support, however, is still far from compelling as an indication of processes involved in reading. An obvious shortcoming for its generalizability to reading is that people do not engage in explicit categorization of words they encounter in text—they read them. Categorization, since it involves accessing semantic memory, may be based on the same processes used in reading words in sentences or phrases, but it may not. The categorization task is thus a limited tool for exploring reading processes. A second difficulty in applying our results to reading is not specific to the categorization task. It may be invalid to generalize findings based on processing isolated words to reading which involves the use of semantic and syntactic context. The presence of rich verbal context could minimize

or eliminate phonemic encoding through facilitation of graphemic access to word meanings, as explained above. The effects of context on phonemic encoding is thus an important problem for future study. One possible approach would be to attempt to produce phonemic encoding bias during actual reading by manipulating the phonemic relations within sequences of words, and measuring fixation time as a dependent variable. This would, however, require forcing the subject to read every word (for example, by putting large spaces between the words, a method currently being used in McKonkie's laboratory), which would partially disrupt normal reading. The construction of stimulus materials would be a major problem for this approach.

From another point of view, one might expect to find more pronounced phonemic processing during normal reading than during processing of isolated words. Kleiman (1975) has suggested that "speech recoding" occurs after lexical access and facilitates the temporary storage of words necessary for sentence comprehension. Speech recoding is needed for the "working memory stage", which involves processing which determines information about the syntactic form of the sentence and the interrelations of word meanings. Kleiman reported results that supported the "working memory hypothesis." His experiments showed that phonemic processing occurred when subjects made decisions about the semantic acceptability of sentences, but not when they made category or synonymity

judgments. Occurrence of speech recoding was assessed by degree of interference (reflected in reaction time) created by shadowing random digits while performing the judgment tasks. Kleiman's paper is significant because it makes a distinction between processing sentences and processing isolated words. Further research on his hypothesis is needed.

It is noteworthy that Kleiman (1975) did not obtain evidence of phonemic processing using a categorization task. This finding conflicts with the tentative positive finding of the present study. A possible explanation of the discrepancy, which might be of concern in future research, is that the two experiments measure different levels or kinds of phonemic processing. Processes revealed by establishing encoding bias in the Meyer _et al_. (1974) task or the categorization task may not be subject to interference by shadowing. Sentence processing could involve somewhat different (or more intensive) phonemic processing which is susceptible to Kleiman's interference technique. Another important problem in the investigation of phonemic processing in reading will be differentiating and discovering the relations among the varieties of phonemic processes included in broad definitions of silent speech. The approach in current use is simply to invent a new term for phonemic processing each time a new technique for assessing it is applied.

Finally, the large individual differences in reading skill suggest that application of the task used in this ex-

periment to groups of differing reading ability might yield differing results. In particular, it would be interesting to find out whether developmental differences in processing words could be observed using the phonemic encoding-bias paradigm. In conclusion, although research aimed at discovering whether phonemic processing occurs in silent reading may prove fruitful, progress in this area will probably not provide an answer to this question, but a specification of what kinds of phonemic processing do or do not occur during particular kinds of reading, under certain conditions and among certain types of subjects.

REFERENCES

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

Barron, R. (in press) Access to the lexical meanings of written words. Frank B. Murray (Ed.) Development of the Reading Process.

Bower, T.G.R. Reading by eye. In H. Levin and J.P. Williams (Eds.), Basic Studies in Reading. New York: Basic Books, 1970.

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

Collins, A. and Loftus, E. A spreading-activation theory of semantic processing. Bolt, Beranek and Newman, Inc., Report, 1975.

Conrad, R. Speech and reading. In Kavanagh, J. and Mattingly, I. (Eds.), Language by Ear and by Eye. Cambridge, Mass.: MIT Press, 1972.

Edfeldt, A. Silent Speech and Silent Reading. Chicago: University of Chicago Press, 1960.

Eriksen, C., Pollack, M. and Montague, W. Implicit speech: Mechanism in perceptual encoding. Journal of Experimental Psychology, 1970, 84, 502-507.

Faaborg-Andersen, K. and Edfeldt, A. Electromyography of intrinsic and extrinsic laryngeal muscles during silent speech: Correlation with reading activity. Acta Otolaryngologia, 1958, 49, 478-482.

Hardyck, C. and Petrinovich, L. Treatment of subvocal speech during reading. Journal of Reading, 1969, 12, 361-368, 419-422.

Hardyck, C. and Petrinovich, L. Subvocal speech and comprehension level as a function of the difficulty level of reading material. Journal of Verbal Learning and Verbal Behavior, 9, 647-652, (1970).

Hochberg, J. Components of Literacy: Speculations and exploratory research. In H. Levin and J. Williams (Eds.), Basic Studies on Reading. New York: Basic Books, 1970.

Klapp, S. Implicit speech inferred from response latencies in same-different decisions. Journal of Experimental Psychology, 1971, 91(2), 262-267.

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

Kucera, H. and Francis, W. Computational Analysis of Present-Day American English. Providence: Brown University Press, 1967.

Meyer, D. and Ellis, G. Parallel processes in word-recognition. Paper read at the meeting of the Psychonomic Society, San Antonio, November, 1970.

Meyer, D. and Ruddy, M. Lexical-memory retrieval based on graphemic and phonemic representations of printed words. Paper presented at the meeting of the Psychnomic Society, Saint Louis, Missouri, 1973.

Meyer, D., Schvaneveldt, R. and Ruddy, M. Functions of graphemic and phonemic codes in visual word recognition. Memory and Cognition, 1974, in press.

Myers, J. Fundamentals of Experimental Design. Boston: Allyn and Bacon, Inc., 1966.

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

Appendix

STIMULUS PAIRS AND THEIR CONTROLS FOR SUBJECTS OF GROUP 1 AND GROUP 2

Group 1: Pairs

Yes-Yes (1)
HARE-MARE
THIEF-CHIEF
HAM-JAM
MUFF-CUFF
LIP-HIP
CAVERN-TAVERN

Yes-Yes (3)
GULL-BULL
TASTER-MASTER
SCONE-HONEY*
ADOBE-ABODE*
SCALP-CALF*
BLOUSE-TROUSER*

Yes-No (1)
GOAT-BOAT
DEER-BEER
MAID-RAID
SON-TON
STEW-CREW*
BUTTER-GUTTER
DRESS-PRESS
BELT-MELT
BACK-TACK
VEIN-REIN
LOBBY-HOBBY
PORCH-TORCH

Yes-No (3)
CROW-PROW
WOLF-GOLF
FARMER-WARMER
LORD-WORD
WAFFLE-BAFFLE
PUDDING-BUDDING
SWEATER-GREATER*
SCARF-DWARF*
BONE-DONE
HAND-WAND
HOVEL-NOVEL
DORM-WORM

No-Yes (1)
VAT-RAT
STOUT-TROUT*
BUDGE-JUDGE
PURSE-NURSE
GUN-BUN
LAKE-CAKE
BUMPER-JUMPER
MOAT-COAT
RUT-GUT
PAIR-HAIR
DOOM-ROOM
TAMPER-CAMPER

Group 2: Controls

Yes-Yes (2)
GULL-MARE
TASTER-CHIEF
SCONE-JAM
BLOUSE-CUFF
SCALP-HIP
ABODE-TAVERN

Yes-Yes (4)
HARE-BULL
THIEF-MASTER
HAM-HONEY
CAVERN-ABODE
LIP-CALF
MUFF-TROUSER

Yes-No (2)
DEER-BOAT
GOAT-BEER
SON-RAID
MAID-TON
BUTTER-CREW
STEW-GUTTER
BELT-PRESS
DRESS-MELT
VEIN-TACK
BACK-REIN
PORCH-HOBBY
LOBBY-TORCH

Yes-No (4)
WOLF-PROW
CROW-GOLF
LORD-WARMER
FARMER-WORD
PUDDING-BAFFLE
WAFFLE-BUDDING
SCARF-GREATER
SWEATER-DWARF
HAND-DONE
BONE-WAND
DORM-NOVEL
HOVEL-WORM

No-Yes (2)
STOUT-RAT
VAT-TROUT
PURSE-JUDGE
BUDGE-NURSE
LAKE-BUN
GUN-CAKE
MOAT-JUMPER
BUMPER-COAT
PAIR-GUT
RUT-HAIR
TAMPER-ROOM
DOOM-CAMPER

Group 2: Pairs

Yes-Yes (1)
GOOSE-MOOSE
SAILOR-TAILOR
POTATO-TOMATO
SMOCK-FROCK*
CHIN-SHIN
HOME-DOME

Yes-Yes (3)
DONKEY-MONKEY
DIVER-GIVER
CHIVE-OLIVE*
BASTION-STATION*
WAIST-WRIST*
HEADGEAR-HEADWEAR*

Yes-No (1)
ROOSTER-BOOSTER
FISH-WISH
KING-RING
CROOK-BROOK
LIME-DIME
CORN-HORN
SHIRT-FLIRT*
VEST-TEST
GLAND-BLAND
RIB-FIB
TENT-WENT
HALL-FALL

Yes-No (3)
HORSE-WORSE
MOTH-BOTH
DAUGHTER-LAUGHTER
BATHER-RATHER
MINT-PINT
BROTH-SLOTH*
HOSE-LOSE
HOOD-MOOD
NAVEL-RAVEL
HEAD-BEAD
BARN-WARN
HUT-PUT

No-Yes (1)
WIG-PIG
ROAD-TOAD
HITCH-WITCH
TOP-COP
SAFER-WAFER
REEF-BEEF
DIE-TIE
SONNET-BONNET
TRAIN-BRAIN
PECK-NECK
SORT-FORT
DODGE-LODGE

Group 1: Controls

Yes-Yes (2)
DONKEY-MOOSE
DIVER-TAILOR
CHIVE-TOMATO
HEADGEAR-FROCK
WAIST-SHIN
BASTION-DOME

Yes-Yes (4)
GOOSE-MONKEY
SAILOR-GIVER
POTATO-OLIVE
HOME-STATION
CHIN-WRIST
SMOCK-HEADWEAR

Yes-No (2)
FISH-BOOSTER
ROOSTER-WISH
CROOK-RING
KING-BROOK
CORN-DIME
LIME-HORN
VEST-FLIRT
SHIRT-TEST
RIB-BLAND
GLAND-FIB
HALL-WENT
TENT-FALL

Yes-No (4)
MOTH-WORSE
HORSE-BOTH
BATHER-LAUGHTER
DAUGHTER-RATHER
BROTH-PINT
MINT-SLOTH
HOOD-LOSE
HOSE-MOOD
HEAD-RAVEL
NAVEL-BEAD
HUT-WARN
BARN-PUT

No-Yes (2)
ROAD-PIG
WIG-TOAD
TOP-WITCH
HITCH-COP
REEF-WAFER
SAFER-BEEF
SONNET-TIE
DIE-BONNET
PECK-BRAIN
TRAIN-NECK
DODGE-FORT
SORT-LODGE

Appendix

Group 1: Pairs	Group 2: Controls	Group 2: Pairs	Group 1: Controls
No-Yes (3)	No-Yes (4)	No-Yes (3)	No-Yes (4)
m-GASP-WASP	LOW-WASP	YEAR-BEAR	PLAN-BEAR
LOW-COW	GASP-COW	PLAN-SWAN	YEAR-SWAN
HONK-MONK	COVER-MONK	BLOWN-CLOWN	SOUTH-CLOWN
COVER-MOVER	HONK-MOVER	SOUTH-YOUTH	BLOWN-YOUTH
FREAK-STEAK*	TOUGH-STEAK	DEAR-PEAR	LINGER-PEAR
TOUGH-DOUGH	FREAK-DOUGH	LINGER-GINGER	DEAR-GINGER
HOE-SHOE*	DROVE-SHOE	SOWN-GOWN	SOOT-GOWN
DROVE-GLOVE*	HOE-GLOVE	SCOT-BOOT	SOWN-BOOT
LOOT-FOOT	BULLET-FOOT	HEARD-BEARD	COUTH-BEARD
BULLET-GULLET	LOOT-GULLET	COUTH-MOUTH	HEARD-MOUTH
HARD-WARD	HAVE-WARD	LOWER-TOWER	ROUSE-TOWER
HAVE-CAVE	HARD-CAVE	ROUSE-HOUSE	LOWER-HOUSE
No-No (1)	No-No (2)	No-No (1)	No-No (2)
m-GUILT-BUILT(An)**	GRACE-BUILT(An)	m-PRICK-TRICK(Sh)	GLASS-TRICK(Sh)
m-GRACE-TRACE(Cl)	GUILT-TRACE(Cl)	GLASS-CLASS(Cl)	PRICK-CLASS(Cl)
m-PITCH-DITCH(An)	MIGHT-DITCH(An)	m-MADE-WADE(Bpt)	RUNG-WADE(Bpt)
m-MIGHT-TIGHT(Fd)	PITCH-TIGHT(Fd)	m-RUNG-SUNG(Sh)	MADE-SUNG(Sh)
DENT-RENT(Pr)	NUMBER-RENT(Pr)	PRIDE-BRIDE(Fd)	TAN-BRIDE(Fd)
NUMBER-LUMBER(An)	DENT-LUMBER(An)	TAN-BAN(An)	PRIDE-BAN(An)
m-BRUISE-CRUISE(Sh)	FILE-CRUISE(Sh)	TENSE-SENSE(Sh)	WEAK-SENSE(Sh)
m-FILE-TILE(Pr)	BRUISE-TILE(Pr)	WEAK-PEAK(An)	TENSE-PEAK(AN)
m-CANDLE-HANDLE(Fd)	SET-HANDLE(Fd)	m-MUCH-SUCH(Bpt)	BARGE-SUCH(Bpt)
m-SET-WET(Pr)	CANDLE-WET(Pr)	m-BARGE-LARGE(Fd)	MUCH-LARGE(Fd)
DUNE-TUNE(Pr)	POISE-TUNE(Pr)	m-SOFT-LOFT(Pr)	NUMB-LOFT(Pr)
m-POISE-NOISE(Cl)	DUNE-NOISE(Cl)	m-NUMB-DUMB(Sh)	SOFT-DUMB(Sh)
m-MARK-DARK(Fd)	TILT-DARK(Fd)	DEW-NEW(Pr)	FAIL-NEW(Pr)
m-TILT-WILT(Bpt)	MARK-WILT(Bpt)	m-FAIL-SAIL(Cl)	DEW-SAIL(Cl)
m-BRIBE-TRIBE(Bpt)	BLAME-TRIBE(Bpt)	PAPER-TAPER(An)	VAULT-TAPER(An)
m-BLAME-FLAME(Sh)	BRIBE-FLAME(Sh)	m-VAULT-FAULT(Bpt)	PAPER-FAULT(Bpt)
m-HILL-WILL(Cl)	BORN-WILL(Cl)	m-YIELD-FIELD(Bpt)	SENT-FIELD(Bpt)
m-BORN-WORN(Sh)	HILL-WORN(Sh)	m-SENT-WENT(Pr)	YIELD-WENT(Pr)
m-VAST-PAST(Fd)	NATIONAL-PAST(Fd)	SHEAR-SPEAR(Fd)*	DANCE-SPEAR(Fd)
m-NATIONAL-RATIONAL(Bpt)	VAST-RATIONAL(Bpt)	DANCE-LANCE(An)	SHEAR-LANCE(An)
m-FOND-POND(Cl)	FENCE-POND(Cl)	TABLE-FABLE(Pr)	COIL-FABLE(Pr)
m-FENCE-HENCE(Bpt)	FOND-HENCE(Bpt)	m-COIL-BOIL(Cl)	TABLE-BOIL(Cl)
m-TRIM-GRIM(Sh)	POINT-GRIM(Sh)	PRAY-TRAY(Cl)	FLAT-TRAY(Cl)
m-POINT-JOINT(An)	TRIM-JOINT(An)	FLAT-SLAT(Fd)	PRAY-SLAT(Fd)
No-No (3)	No-No (4)	No-No (3)	No-No (4)
m-DULL-PULL(Fd)	CATCH-PULL(Fd)	HINDER-BINDER(An)	FEAR-BINDER(An)
m-CATCH-WATCH(An)	DULL-WATCH(An)	FEAR-WEAR(Sh)	HINDER-WEAR(Sh)
TREAD-PLEAD(An)*	BAKED-PLEAD(An)	DOOR-BOOR(Bpt)	POUR-BOOR(Bpt)
m-BAKED-NAKED(Sh)	TREAD-NAKED(Sh)	POUR-HOUR(Pr)	DOOR-HOUR(Pr)
POST-COST(Pr)	BROUGHT-COST(Pr)	SAD-WAD(Cl)	TONE-WAD(Cl)
BROUGHT-DROUGHT(Cl)	POST-DROUGHT(Cl)	TONE-NONE(Fd)	SAD-NONE(Fd)
m-CASH-WASH(Pr)	PATIO-WASH(Pr)	m-SOUR-FOUR(Bpt)	MOWN-FOUR(Bpt)
m-PATIO-RATIO(An)	CASH-RATIO(An)	MOWN-TOWN(Pr)	SOUR-TOWN(Pr)
GUSH-PUSH(Cl)	COUCH-PUSH(Cl)	CEILING-VEILING(An)	DEAF-VEILING(An)
m-COUCH-TOUCH(Pr)	GUSH-TOUCH(Pr)	DEAF-LEAF(Bpt)	CEILING-LEAF(Bpt)
m-FOUL-SOUL(Sh)	HUSH-SOUL(Sh)	FAR-WAR(Cl)	ROVER-WAR(Cl)
m-HUSH-BUSH(Cl)	FOUL-BUSH(Cl)	ROVER-LOVER(Fd)	FAR-LOVER(Fd)
GROVE-PROVE(Pr)	GROWN-PROVE(Pr)	m-FEW-SEW(Pr)	READ-SEW(Pr)
m-GROWN-CROWN(Fd)	GROVE-CROWN(Fd)	READ-DEAD(Sh)	FEW-DEAD(Sh)

Appendix

Group 1: Pairs No-No (3)	Group 2: Controls No-No (4)	Group 2: Pairs No-No (3)	Group 1: Controls No-No (4)
m-PAID-SAID(Bpt)	ROUGH-SAID(Bpt)	METER-DETER(Cl)	HEATH-DETER(Cl)
ROUGH-BOUGH(An)	PAID-BOUGH(An)	HEATH-DEATH(An)	METER-DEATH(An)
COMB-BOMB(Bpt)	CART-BOMB(Bpt)	POSTER-FOSTER(Fd)	NEVER-FOSTER(Fd)
CART-WART(Fd)	COMB-WART(Fd)	m-NEVER-FEVER(An)	POSTER-FEVER(An)
m-JURY-BURY(Bpt)	LOST-BURY(Bpt)	m-NATURE-MATURE(Cl)	NASTY-MATURE(Cl)
m-LOST-MOST(Sh)	JURY-MOST(Sh)	m-NASTY-HASTY(Fd)	NATURE-HASTY(Fd)
WILY-LILY(Bpt)	SHOW-LILY(Bpt)	BOWL-HOWL(Pr)	LATER-HOWL(Pr)
SHOW-CHOW(Cl)	WILY-CHOW(Cl)	LATER-WATER(Sh)	BOWL-WATER(Sh)
FREIGHT-SLEIGHT(Fd)*	HONOR-SLEIGHT(Fd)	SLOWER-FLOWER(Sh)	TEEN-FLOWER(Sh)
m-HONOR-DONOR(Sh)	FREIGHT-DONOR(Sh)	TEEN-BEEN(Bpt)	SLOWER-BEEN(Bpt)

*Graphemically similar pairs in which the difference in spelling between the two words is not limited to the first letter position.

mGraphemically similar pairs used by Meyer et al. (1974).

**The categories which were assigned to the N-N pairs are abbreviated as follows:

- Animal (An)
- Food (Fd)
- Person (Pr)
- Shelter (Sh)
- Bodypart (Bpt)
- Clothing (Cl)