Functions of graphemic and phonemic codes in visual word-recognition*

DAVID E. MEYER

Bell Laboratories, Murray Hill, New Jersey 07974

ROGER W. SCHVANEVELDT

State University of New York, Stony Brook, New York 11790

and

MARGARET G. RUDDY

Bell Laboratories, Murray Hill, New Jersey 07974

Previous investigators have argued that printed words are recognized directly from visual representations and/or phonological representations obtained through phonemic recoding. The present research tested these hypotheses by manipulating graphemic and phonemic relations within various pairs of letter strings. Ss in two experiments classified the pairs as words or nonwords. Reaction times and error rates were relatively small for word pairs (e.g., BRIBE-TRIBE) that were both graphemically and phonemically similar. Graphemic similarity alone inhibited performance on other word pairs (e.g., COUCH-TOUCH). These and other results suggest that phonological representations play a significant role in visual word recognition and that there is a dependence between successive phonemic-encoding operations. An encoding-bias model is proposed to explain the data.

Three basic theories have been proposed to describe psychological processes that occur in recognizing printed words. The *graphemic-encoding hypothesis* supposes that a printed word is recognized directly from a visual representation that is used to locate stored information about the meaning of the word (Bower, 1970; Kolers, 1970). In contrast, the *phonemic-encoding hypothesis* presumes that recognition involves converting a visual representation of the word into a phonological representation and that the latter code provides access to lexical memory (Rubenstein, Lewis, & Rubenstein, 1971). These two theories are integrated in the *dual-encoding hypothesis* (Baron, 1973; LaBerge, 1972). According to this model, lexical memory can be accessed through both visual and phonological representations of a printed word, and retrieval processes based on the two types of code may occur in parallel. The three theories are summarized in Fig. 1. Here the dashed and solid lines connect those processes associated with the graphemic- and phonemic-encoding hypotheses, respectively, while the dual-encoding hypothesis is represented by all of the operations in combination.[1]

Despite the plausibility of these alternative views, the evidence supporting each of them is somewhat limited.

In the next two sections, we survey available results concerning the graphemic- and phonemic-encoding hypotheses. The validity of the dual-encoding hypothesis will be considered later. (For other reviews of the literature, see Baron, 1973; Conrad, 1972; Gough, 1972).

EVIDENCE FOR THE GRAPHEMIC-ENCODING HYPOTHESIS

The principal evidence for the graphemic-encoding hypothesis comes from experiments whose results seem incompatible with phonemic encoding and, consequently, imply direct visual word recognition. One major finding involves "speed reading" (Carver, 1972; McLaughlin, 1969). It has been claimed that skilled readers can process text at over 20 words/sec (1200 words/min) with a high degree of comprehension. On the other hand, the maximum rate of covert speech appears to be no more than about 10 words/sec (600 words/min), which suggests that the processing rate is
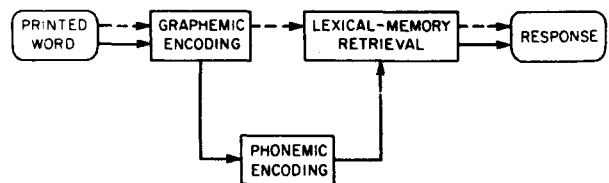

Fig. 1. Possible stages in visual word recognition. Dashed and solid lines indicate the sequence of operations according to the graphemic- and phonemic-encoding hypotheses, respectively. The dual-encoding hypothesis is represented by all of the operations in combination.

substantially less in subvocalization than in speed reading (Landauer, 1962; Sperling, 1967). Some investigators have therefore concluded that people must be able to recognize printed words without converting them into phonological representations (e.g., Bower, 1970; Kolers, 1970).

However, such an interpretation of speed reading is open to question. It ignores the fact that prose material is highly redundant (Shannon, 1951). To gain a good understanding of what is written in most text passages, a person has to recognize only a fraction of the words. When reading at a nominal rate of 1200 words/min, he may actually process just a few hundred items, yet display high comprehension (cf. Pierce & Karlin, 1957). This would make the net reading rate agree more closely with maximum rates of subvocalization.[2]

Even if the actual rate was more than 1200 words/min, the phonemic-encoding hypothesis would not be untenable. Rather than occurring as covert speech itself, phonological representations could be created in other forms, e.g., implicit "motor programs" (Hochberg, 1970; Liberman, 1957). Whereas such programs might take hundreds of milliseconds to execute, some results of tachistoscopic experiments suggest that they may be constructed in tens of milliseconds (Sperling, 1967). Given the latter possibility, the phonemic-encoding hypothesis would be entirely consistent with speed reading, eliminating the need to postulate a mechanism for direct visual word recognition. As Gough (1972) has observed elsewhere, it is also conceivable that phonological representations involve even more abstract components, such as the "systematic phonemes" of Chomsky and Halle (1968), which might permit even higher reading rates.

Some other experiments that help to support the graphemic-encoding hypothesis are questionable for different reasons. For example, Bower (1970) instructed a group of bilingual Ss to translate passages of Greek text into English. In one condition, the passages consisted of Greek words spelled in their usual fashion. In a second condition, the passages were modified by replacing some of the items with phonemically equivalent Greek pseudowords. The alterations were analogous to changing an English text by replacing words like PHONOGRAPH with pseudowords like FONOGRAF. Ss took about twice as long to translate the modified passages. Evaluating his data, Bower (1970, p. 144) asserted: "The [alterations] would not have done this had the normal processing procedure involved a loop into auditory or articulatory space, because the [alterations] produced no change in the auditory or articulatory correlates of the passage." Hence, he inferred that direct visual word recognition does occur in processing ordinary text.

Yet there are other ways to account for Bower's results. Graphemic differences between words and pseudowords may affect operations before phonemic encoding. Perhaps reading involves a visual "preprocessing" stage that is influenced by graphemic structure, e.g., bigram and trigram frequencies (Neisser, 1967). Graphemic structure may also influence operations that convert a letter string into a phonological representation (Brown, 1970). Either of these possibilities could explain Bower's findings without precluding phonemic encoding. Effects of other visual factors, such as rotating or inverting the letters of a text (Kolers, 1972), might be explained in these terms too.

A further test of the graphemic-encoding hypothesis has been developed by Baron (1973). In one of his experiments, Ss had to decide whether or not various printed phrases "looked meaningful." Ss in a second experiment judged whether the printed phrases "sounded meaningful." Three kinds of stimuli were included in both studies: (a) visually and phonemically congruent phrases like MY NEW CAR and I KNEW HIM, whose correct answers were always "yes;" (b) visually anomalous but phonemically congruent phrases like MY KNEW CAR and I NEW HIM, whose correct answers were "no" and "yes" in the two experiments, respectively; and (c) visually and phonemically anomalous phrases like OUR NO CAR and NEW I CAN'T, whose correct answers were always "no."

Baron obtained two major results. In judging that a phrase sounded meaningful, Ss took less time to classify stimuli of Type a than of Type b; i.e., when the phrases were phonemically congruent, reaction time (RT) was shorter if they were also visually congruent. But for judgments that a phrase did not look meaningful, stimuli of Types b and c were classified equally fast; i.e., when the phrases were visually anomalous, RT did not depend on whether they were phonemically anomalous too. From these results, Baron argued that the meaning of a word can be accessed directly from its visual representation without phonemic encoding, although such encoding may be used sometimes.

While Baron's experiments are more persuasive than some other studies, a few doubts might still be raised about his interpretation. For example, consider the comparison between phrases like MY KNEW CAR and OUR NO CAR, where both are visually anomalous but only the latter is phonemically anomalous. Although Ss were equally fast in judging that the two types of phrase did not look meaningful, Baron reported significantly more errors on the former type, which is phonemically congruent. Thus, there may have been a speed-accuracy tradeoff in his experiment. If the two types of phrase had been equated for errors, then the phonemically congruent stimuli might have produced longer RTs, which would be consistent with the phonemic-encoding rather than the graphemic-encoding hypothesis (cf. Baron, 1973).

Another problem for Baron's experiments involves the comparison between phrases such as MY NEW CAR and MY KNEW CAR, where both are phonemically congruent but only the former is visually congruent.

While these two types of phrase have identical pronunciations, they differ in terms of the frequency with which they occur as units in ordinary text; a phrase like MY KNEW CAR never appears in writing, whereas one like MY NEW CAR is not especially uncommon. Without implying direct visual word recognition, such a difference could explain why Ss took less time to decide that the visually congruent phrases sounded meaningful. In particular, as mentioned above, there may be a visual preprocessing stage before phonemic encoding. Thus, if a phrase is sometimes treated as a unit during that stage and if the operation is influenced by the visual familiarity of the unit, then one might expect a more frequent phrase to be processed faster even when its subsequent classification depends solely on phonemic encoding.

EVIDENCE FOR THE PHONEMIC-ENCODING HYPOTHESIS

Evidence for the phonemic-encoding hypothesis is not a great deal more compelling. For example, many people are aware of an inner voice speaking as they silently read (Brown, 1970). This introspective phenomenon might be a subjective correlate of concurrent grapheme-phoneme transformations and, indeed, electromyographic recordings have confirmed the presence of substantial speech-muscle activity during silent reading (Aarons, 1971; Edfeldt, 1960; Locke, 1971). However, such activity does not necessarily imply that phonemic encoding is a precursor to recognizing printed words. As part of reading, covert speech could serve other functions than word recognition. Covert speech may help to maintain individual words in short-term memory after recognition, so that they can be integrated to determine the meaning of a sentence (Conrad, 1972; Gough, 1972). This would be consistent with a substantial body of research suggesting that, when visually presented items must be recalled, they are stored phonemically (Conrad, 1964; Hintzman, 1967; Sperling & Speelman, 1970). Requirements of short-term memory could also explain other results that seem to support the importance of covert speech in recognizing various kinds of printed symbols (e.g., Klapp, 1970).

To test the phonemic-encoding hypothesis more directly, Rubenstein et al (1971) have used an approach that attempts to avoid the involvement of short-term memory. On each trial in their method, a string of letters is presented, and the S must judge whether it is an English word or nonword. This *lexical-decision task* does not require Ss to remember the stimulus items after presentation. But as we shall discuss in more detail later, performance of the task does presumably depend on operations that mediate the recognition of printed words in various situations, i.e., graphemic and/or phonemic encoding followed by accessing of lexical memory.

Rubenstein et al found that, in the lexical-decision task, RT seemed to depend on phonemic properties of the letter strings. Negative ("no") responses were fastest for unpronounceable nonwords (e.g., BRAKV), slower for pronounceable nonwords (e.g., BLEAN), and slowest for nonwords (e.g., BRUME) that were homophonic with English words (e.g., BROOM). Phonemic properties of the stimuli also affected positive ("yes") responses. Reaction time was longer for words (e.g., MAID) that were homophonic with other English words (e.g., MADE) than for nonhomophones (e.g., BATH). Similar results have been obtained by other investigators (Snodgrass & Jarvella, 1972; Stanners, Forbach, & Headley, 1971).

According to Rubenstein et al, their findings indicate the occurrence of phonemic encoding. They proposed that lexical (word-nonword) decisions are made by forming and comparing a phonological representation of a letter string with a subset of phonological representations stored in lexical memory. In this model, a positive response occurs if (a) the comparison process yields a match, and then (b) the string is found to have the correct spelling. The model presumes that the comparison process is omitted whenever illegalities are detected during the initial phonemic-encoding stage. This would explain why negative responses are faster for unpronounceable than pronounceable nonwords. The model also accounts for the RT data concerning homophonic words and nonwords. Negative responses should be slowest for homophonic nonwords, since they match at least one phonological representation in lexical memory and, therefore, have to undergo the additional spelling check before being rejected. Similarly, because homophonic words have two or more spellings that potentially must be checked, they should not be judged as quickly as nonhomophones.

There is, however, some reason to doubt the interpretation provided by Rubenstein et al. In the experiments summarized above, graphemic properties of the letter strings may have been confounded with phonemic properties. For example, it is possible that the unpronounceable nonwords looked least like English, whereas the homophonic nonwords looked most like English. Thus, despite an apparent consistency with phonemic encoding, the results of Rubenstein et al might also be explained in terms of the graphemic-encoding hypothesis.

To illustrate this point in more detail, let us define the "graphemic similarity" of a letter string to English as the number of word types from which the string differs in at most one letter position; e.g., the graphemic similarity of BLEAN to English would be at least one, since it differs from the word CLEAN in only the first letter position; the similarity of MAID would be at least two, since this string is itself a word and also differs from the word PAID by only one letter (cf. Landauer & Streeter, 1973). Now suppose that in the lexical-decision task of Rubenstein et al the following was true: (a) the graphemic similarity of the nonwords to English was

positively correlated with their pronounceability; (b) each decision about a nonword involved comparing the visual representation of the string with some subset of visual representations in lexical memory; and (c) the comparison rate was slower or the number of comparisons was greater for nonwords that had a higher graphemic similarity to English. Then negative responses would have been slower for pronounceable than unpronounceable nonwords, because more time would have been required to discover that the pronounceable nonwords did not match any entry in memory. This is exactly what Rubenstein et al found.

An analogous argument may apply to their results for homophonic words and nonwords. One might even speculate that the graphemic structure of a homophonic word is usually more similar than a nonhomophone to other English words. For example, a homophone like DEAR is graphemically similar to at least one other word, namely its homophonic mate DEER, whereas that is not necessarily true for nonhomophones. This could explain why homophones are judged less rapidly, since they would be visually confusable with more entries in lexical memory.

Confoundings between graphemic and phonemic properties of stimuli might also account for other findings that appear to support phonemic encoding (e.g., Corcoran, 1966, 1967; Eriksen, Pollack, & Montague, 1970).

EXPERIMENT I

The present research was designed to further assess the roles of graphemic- and phonemic-encoding in visual word recognition using a procedure that overcomes some of the problems encountered by previous investigators. As in the experiments of Rubenstein et al, we relied on a lexical-decision task, requiring Ss to judge whether various strings of letters were words or nonwords. However, our procedure differed from theirs in several respects. On each trial, the S was presented with two letter strings instead of just one. The stimulus included a pair of words, a pair of nonwords, or a word and a nonword. In the study that we shall consider first, the strings were displayed simultaneously, and the S had to respond "yes" if both of them were English words; otherwise, he responded "no." The procedure was similar to one used by Meyer and Schvaneveldt (1971) for studying how word recognition is influenced by semantic variables.

With this *simultaneous-presentation method*, we tried to separate the effects of graphemic and phonemic factors on recognition. The critical variables were the graphemic and phonemic relations within the pairs of words. We constructed four types of word pairs. Type 1 involved rhyming words that differed only in their initial letter, thereby being both graphemically and phonemically similar (e.g., BRIBE–TRIBE and FENCE-HENCE). Type 2 was formed by randomly interchanging first members of these pairs. Words in the

resulting control pairs (e.g., BRIBE-HENCE and FENCE-TRIBE) were graphemically and phonemically dissimilar. Type 3 consisted of words that were graphemically similar but phonemically dissimilar because they did not rhyme (e.g., COUCH-TOUCH and FREAK-BREAK). Finally, Type 4 was a control for Type 3 created by randomly interchanging those items to produce more pairs of dissimilar words (e.g., COUCH-BREAK and FREAK-TOUCH).

These four types of words together permit the following test of the graphemic-encoding hypothesis vs the phonemic- and dual-encoding hypotheses, thereby allowing us to determine whether or not phonological representations play a significant role in the recognition process:

(a) Suppose that the graphemic-encoding hypothesis is correct, i.e., that printed words are recognized solely from their visual representations. Then the RT of lexical decisions could depend on the graphemic relation between words, but it should not depend on their phonemic relation. In particular, consider pairs of graphemically and phonemically similar words like BRIBE-TRIBE and FENCE-HENCE (Type 1), as opposed to their controls BRIBE-HENCE and FENCE-TRIBE (Type 2). There might be a significant difference in the RTs for these two types of stimuli, because their graphemic relation differs. For example, graphemic similarity might facilitate recognition, so that RT would be longer for the control pairs. But whatever the RT difference, it should be approximately equal to the corresponding difference for graphemically similar pairs like COUCH-TOUCH and FREAK-BREAK (Type 3) vs their controls COUCH-BREAK and FREAK-TOUCH (Type 4). It is irrelevant that both of the latter types of pairs are phonemically dissimilar, whereas one of the former two types is phonemically similar. According to the graphemic-encoding hypothesis, only visual properties of the stimuli ought to influence recognition. Thus, letting T_i denote the mean RT for word pairs of Type i, the following equation should hold under the graphemic-encoding hypothesis:

$$T_2 - T_1 = T_4 - T_3 \qquad (1)$$

(b) In contrast, suppose that either the phonemic-encoding or dual-encoding hypothesis is correct, i.e., that printed words are recognized at least sometimes through phonological representations. Then one might expect that the RT of lexical decisions would be influenced by the phonemic relation between words. The difference between RTs for the pairs of graphemically similar words and their controls should depend on whether or not they are phonemically similar. Thus, Eq. 1 would not hold. For example, recognition might be facilitated by phonemic similarity in addition to graphemic similarity, so that the following inequality would be obtained instead:

$$T_2 - T_1 > T_4 - T_3 \qquad (2)$$

Method

Subjects

The Ss were six high school students and six technical employees of Bell Laboratories.

Apparatus

The experiment was controlled by a digital computer with a millisecond timer connected to a display oscilloscope (CRT), random-noise generator, and response panel with finger keys for the right and left hands.

Procedure

Ss were seated in a darkened room facing the CRT and wearing a pair of headphones. They were run individually in a single 1-h session that included a short instruction period, two practice blocks of 36 trials each, and six test blocks of 36 trials each. At the start of each trial, the S heard a soft burst of white noise that served as a warning signal and stayed on throughout the trial. Five hundred milliseconds after the noise began, a small fixation point appeared on the CRT and remained visible during a 500-msec foreperiod, which ended with simultaneous presentation of two letter strings. The strings were displayed horizontally, one above the other in capital letters. The first and second members of a pair were always the top and bottom strings, respectively. Together they subtended horizontal and vertical visual angles of about 2 deg and 1 deg, respectively, with the top string centered where the fixation point had been and separated from the bottom string by approximately 0.2 deg. The S had to judge whether or not both strings were English words. He pressed a "yes" key with his right index finger to indicate a positive decision, otherwise pressing a "no" key with his left index finger for a negative decision. He was instructed to respond as quickly and accurately as possible. Reaction time was measured from onset of the letter strings to the response, which terminated the stimulus display. The screen remained blank for 2 sec before the next trial. If the S made an error, then this interval was extended to 4 sec, during which the word INCORRECT was displayed for the first 2 sec. After the trial block, the S was informed about his mean RT, total number of correct responses, and total number of errors for the block. There was a rest period of approximately 2 min between blocks.

The high school students were paid an initial sum of $1.25 for participating in the experiment. In addition, both these and the other Ss were given a bonus for responding quickly and accurately. They received 1 cent for every correct response, lost 2 cents for every error, and were also charged 1 cent for each .1 sec in their mean RT on each block.

Stimuli and Design

Each S classified 32 word-word (WW) pairs, 16 word-nonword (WN) pairs, 16 nonword-word (NW) pairs, and 8 nonword-nonword (NN) pairs in the initial two practice blocks. None of these pairs was seen later in the experiment. During the subsequent six test blocks, 96 WW pairs, 48 WN pairs, 48 NW pairs, and 24 NN pairs were classified. These stimuli were drawn from a larger pool of pairs chosen to counterbalance various factors across Ss. Table 1 summarizes the types of stimuli involved.

Among the test stimuli, the following WW pairs were included: 48 pairs of graphemically and phonemically similar words (Type 1, e.g., BRIBE-TRIBE and FENCE-HENCE); 48 control pairs of dissimilar words (Type 2, e.g., BRIBE-HENCE and FENCE-TRIBE), obtained by randomly interchanging first members of the Type 1 pairs; 48 pairs of graphemically similar words that were phonemically dissimilar (Type 3, e.g., COUCH-TOUCH and FREAK-BREAK); and 48 control pairs (Type 4, e.g., COUCH-BREAK and FREAK-TOUCH), obtained from pairs of Type 3. These words, as listed in the Appendix, were sampled from the frequency norms of Kúcera and Francis (1967). They ranged in frequency from less than 1 to more than 1000 words per million and from three to eight letters in length. The median frequency was approximately 30 per million, and the median length was approximately four letters. 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. Members of the first and second types were equated approximately in average length and frequency with those of the third and fourth types.

In addition to the WW pairs, we used a variety of pairs involving nonwords that followed the general rules of English orthography and phonology. So that graphemic and phonemic relations between letter strings would not be confounded with the correct response ("yes" or "no"), half of the pairs including nonwords (i.e., WN, NW, and NN) were graphically similar. Each of the nonwords in the graphemically similar pairs could be pronounced to rhyme with its mate. However, at least some of the nonwords had more than one possible pronunciation, so that the phonemic relations involving those letter strings were partially indeterminate. This fact is indicated by the question marks in Table 1.

The following pairs with nonwords were included among the test stimuli: 96 WN pairs whose strings were graphemically similar (e.g., RUMOR-FUMOR and HEDGE-PEDGE), and 96 WN control pairs (e.g., RUMOR-PEDGE and HEDGE-FUMOR), which were obtained by randomly interchanging first members of the graphemically similar pairs; 96 NW pairs that were graphemically similar (e.g., SOIST-MOIST and FRUNK-DRUNK) and 96 NW control pairs (e.g., SOIST-DRUNK and FRUNK-MOIST); 48 NN pairs that were graphemically similar (e.g., DEACE-MEACE and CULSE-GULSE) and 48 NN control pairs (e.g., DEACE-GULSE and CULSE-MEACE). Half of the NN pairs were obtained by reversing the order of the nonwords in the other half. Similarly, the NW pairs were constructed by reversing the order of members in the WN pairs. Words in these pairs were again sampled from the norms of Kúcera and Francis (1967). They were matched approximately in average frequency and length with members of the WW pairs but were not actually included among the WW pairs. Nonwords were constructed from common words by replacing their initial consonants with other consonants. The resulting strings of letters were equated approximately in average length with the words.

An incomplete-block design was used to assign stimuli to the Ss. Over the six test blocks, an S was shown half of the stimuli from each type of WW pair and one-fourth of the stimuli from each type of WN, NW, and NN pair. The stimulus assignment was balanced across Ss so that no S saw the same letter string more than once. For example, each WW pair of Type 1 was presented to six Ss, and the corresponding Type 2 control pair was presented to the remaining six Ss. Each graphemically similar WN pair was assigned to three Ss, while the corresponding control pair was assigned to another three. If an S had to classify a particular WN pair, then he was not presented the NW pair obtained by reversing it. During each test block, the stimuli were displayed with the relative frequencies indicated by Table 1. Thus, approximately 44% of the trials required positive responses. The order of presenting the pairs was random within blocks.

Data Analysis

The principal data are the mean RTs of correct responses and percentage of errors for the various stimulus types classified during the test blocks. Because each S did not classify all of the stimulus pairs, we obtained statistical tests between the means

Table 1
Mean RTs of Correct Responses and Percentage of Errors for Each Stimulus Type in Experiment 1*

Type of Stimulus Pair	Graphemic Relation	Phonemic Relation	Examples	Correct Response	Relative Frequency	Mean RT (msec)	Percent Errors
Word-Word (1)	Similar	Similar	BRIBE-TRIBE FENCE-HENCE	Yes	.111	935	5.2
Word-Word (2)	Dissimilar	Dissimilar	BRIBE-HENCE FENCE-TRIBE	Yes	.111	956	6.9
Word-Word (3)	Similar	Dissimilar	COUCH-TOUCH FREAK-BREAK	Yes	.111	1015	11.5
Word-Word (4)	Dissimilar	Dissimilar	COUCH-BREAK FREAK-TOUCH	Yes	.111	928	3.5
Word-Nonword	Similar	?	RUMOR-FUMOR HEDGE-PEDGE	No	.111	1023	7.6
Word-Nonword	Dissimilar	Dissimilar	RUMOR-PEDGE HEDGE-FUMOR	No	.111	1066	6.9
Nonword-Word	Similar	?	SOIST-MOIST FRUNK-DRUNK	No	.111	922	2.8
Nonword-Word	Dissimilar	Dissimilar	SOIST-DRUNK FRUNK-MOIST	No	.111	913	3.1
Nonword-Nonword	Similar	?	DEACE-MEACE CULSE-GULSE	No	.056	830	0.0
Nonword-Nonword	Dissimilar	Dissimilar	DEACE-GULSE CULSE-MEACE	No	.056	850	0.0

*Question marks indicate that nonwords in some of the pairs had more than one possible phonological representation and may have been either phonemically similar or dissimilar to their mates.

by first randomly dividing the Ss into one collection of subsets and the strings of letters into another orthogonal collection of subsets. Within each subset of Ss, all of the members together had judged all of the stimulus pairs; within each subset of letter strings, the same strings had been used to construct corresponding test and control pairs of stimuli. These groupings yielded a relatively sensitive "crossed" design in which the subsets of Ss and letter strings served as sampling units. Next the data were averaged within the cells of this design, and the results were submitted to an analysis of variance (Winer, 1962, pp. 199-207; Clark, 1973) in which both the Ss and strings of letters were treated as "random effects," so that the reliability of our results could be computed over both of these sampling domains.[3] Before the analysis, an arc sine transformation was performed on the error rates.

Results

The mean RTs of correct responses and percentage of errors for each type of stimulus are shown at the right of Table 1.

Effects of Graphemic and Phonemic Relations

Positive responses to the graphemically and phonemically similar WW pairs (Type 1) were faster than responses to the corresponding control stimuli (Type 2), but not reliably so. The difference was 21 ± 17 msec, where the second number indicates the standard deviation of the first $[F(3,16) = 1.2, p > .10]$.[4] The difference in error rates for these two types of pairs was not reliable either (F < 1.0).

In contrast, significantly slower positive responses

occurred for the graphemically similar WW pairs that were phonemically dissimilar (Type 3) than for the corresponding control stimuli (Type 4). The difference was 87 ± 20 msec $[F(1,13) = 11.5, p < .01]$. This effect does not appear to have been caused by a speed-accuracy tradeoff. The error rate was also significantly higher for the Type 3 pairs than for the Type 4 controls $[F(1,15) = 6.7, p < .05]$.

There was a significant effect of varying phonemic similarity in addition to graphemic similarity for the WW pairs. The difference between mean RTs for the Type 3 and 4 pairs was 108 ± 23 msec greater than the difference between mean RTs for the Type 1 and 2 pairs $[F(1,16) = 11.0, p < .01]$. Thus, Eq. 1 did not hold. An analogous interaction appeared in the error rates $[F(1,10) = 3.9, p < .10]$.

For negative responses to WN, NW, and NN pairs, varying the graphemic relation between strings of letters produced no reliable main effects or interactions; $p > .10$ in all cases. Here the largest main effect on RT was 43 ± 32 msec in magnitude, which occurred for the graphemically similar WN pairs vs their controls $[F(1,2) = 1.6]$.

Effects of Display Position

However, the speed and accuracy of negative responses did depend on the location of nonwords in the stimulus display. Reaction time was 127 ± 26 msec longer for WN than NW pairs $[F(1,4) = 20.9, p < .02]$. The error rate was also significantly higher for the

former kind of pair $[F(1,7) = 14.4, p < .01]$.

Finally, negative responses were reliably faster for the NN than NW pairs. The difference was 78 ± 19 msec $[F(1,7) = 10.0, p < .02]$. An analogous effect occurred in the error rates, which were less for the NN pairs $[F(1,3) = 7.5, p < .10]$.

Conclusions

Our most important finding is that performance depended on the phonemic relation within pairs of words, not just the graphemic relation. Contrary to what might have been expected, phonemic similarity together with graphemic similarity facilitated recognition slightly (Type 1 vs Type 2), but performance was actually inhibited by graphemic similarity alone (Type 3 vs Type 4). This result would not have happened if Ss always recognized the words directly from their visual representations (see Eq. 1). Thus, the data are inconsistent with the graphemic-encoding hypothesis and indicate the involvement of phonological representations in visual word recognition, as proposed in the phonemic-encoding and dual-encoding hypotheses.

Although they are of secondary interest here, the results from negative responses suggest a framework for further interpreting the effects of graphemic and phonemic relations. As in some previous studies (e.g., Meyer & Schvaneveldt, 1971), the RTs of negative responses are compatible with the view that processing of the stimuli was serial and self-terminating. It appears that operations usually began on the top string of letters in the stimulus display. When the top string was classified as a nonword, processing stopped and a negative response was made. However, processing was shifted to the bottom string if the top one was classified as a word. When the bottom string was also classified as a word, a positive response occurred; otherwise, the S made a negative response. This would explain why RT was longer for WN than NW and NN pairs, since the latter stimuli would require fewer operations. Reversals in the order of the stages or a partial overlap between them could account for the somewhat faster responses to NN than NW pairs (Meyer & Schvaneveldt, 1971; Schvaneveldt & Meyer, 1973).

Within this serial model, our data indicate that graphemic and phonemic relations between words may induce a correlation between successive operations. In particular, it is possible that phonemic encoding of the second word in a pair depends on graphemic and phonemic similarities with the first word. Such a dependence is reminiscent of effects that semantic similarity has on successive operations in a variety of word recognition tasks (e.g., Meyer & Schvaneveldt, 1971; Meyer, 1973; Schaeffer & Wallace, 1970). But before considering a more detailed explanation for the present results, we shall report one further study.

EXPERIMENT II

To reinforce the interpretation of Experiment I, a procedure was used in which two strings of letters were presented successively rather than simultaneously on every trial. In this *successive-presentation method*, the S had to classify each string separately as a word or nonword, and RT was measured for the individual strings (see Meyer, Schvaneveldt, & Ruddy, 1972). The method provided better control over the order in which Ss examined the stimuli. They were forced to process both of the letter strings in WN, NW, and NN pairs as well as WW pairs, whereas Experiment I sometimes allowed self-termination after the first string had been classified. It was therefore possible to obtain a more precise estimate of the time required to recognize each string.

If our interpretation of Experiment I is correct, then RT measured separately for the second word in a pair should depend on both its graphemic and phonemic relation to the first word. Thus, the difference between mean RTs for graphemically similar words and their controls in Experiment II should vary with phonemic similarity.

Method

Subjects

The Ss were six high school students and six technical employees of Bell Laboratories. They were sampled from the same populations as in Experiment I, but none of them had participated in that study.

Procedure

The same apparatus, stimuli, and basic design were used as in Experiment I. Only the method of stimulus presentation and response differed. After the foreperiod of each trial, the first letter string of a pair was displayed on the CRT, centered where the fixation point had been. The S had to judge whether or not it was a word. He pressed a "yes" key with his left index finger to indicate a positive decision, otherwise pressing a "no" key with his left middle finger for a negative decision. As soon as the response occurred, the first string of letters was removed, and the second letter string of the pair appeared 0.2 deg below it. The S again had to judge whether or not the string was a word. Another set of response keys was used for this decision. The S pressed a "yes" key with his right index finger to indicate a positive decision or a "no" key with his right middle finger for a negative decision. This removed the second letter string. Ss were instructed to respond as quickly and accurately as possible to each string and were provided with feedback as in Experiment I. The two sets of keys were used in an attempt to minimize interactions between the successive motor responses. Reaction time was measured for each decision from the onset of the letter string to the keypress. Ss were paid on the basis of a scheme similar to that in Experiment I.

Results

Responses to the first string of letters in a pair tended to be somewhat slower but more accurate than responses to the second string. In general, there were no systematic

Table 2
Correct Response, Mean RT, and Percentage of Errors for the Second Letter String of Each Stimulus Type in Experiment II*

Type of Stimulus Pair	Graphemic Relation	Phonemic Relation	Examples	Correct Response	Relative Frequency	Mean RT (msec)	Percent Errors
Word-Word (1)	Similar	Similar	BRIBE-TRIBE FENCE-HENCE	Yes	.111	589	1.4
Word-Word (2)	Dissimilar	Dissimilar	BRIBE-HENCE FENCE-TRIBE	Yes	.111	605	0.8
Word-Word (3)	Similar	Dissimilar	COUCH-TOUCH FREAK-BREAK	Yes	.111	633	3.9
Word-Word (4)	Dissimilar	Dissimilar	COUCH-BREAK FREAK-TOUCH	Yes	.111	599	2.8
Word-Nonword	Similar	?	RUMOR-FUMOR HEDGE-PEDGE	No	.111	701	3.8
Word-Nonword	Dissimilar	Dissimilar	RUMOR-PEDGE HEDGE-FUMOR	No	.111	716	3.1
Nonword-Word	Similar	?	SOIST-MOIST FRUNK-DRUNK	Yes	.111	601	2.8
Nonword-Word	Dissimilar	Dissimilar	SOIST-DRUNK FRUNK-MOIST	Yes	.111	565	1.8
Nonword-Nonword	Similar	?	DEACE-MEACE CULSE-GULSE	No	.056	779	5.6
Nonword-Nonword	Dissimilar	Dissimilar	DEACE-GULSE CULSE-MEACE	No	.056	795	6.3

*Question marks indicate that nonwords in some of the pairs had more than one possible phonological representation and may have been either phonemically similar or dissimilar to their mates.

effects of graphemic or phonemic relations on responses to the first string, since such relations were not apparent to the S until the second string had been presented.

The principal data are the mean RTs of correct responses and percentage of errors for the second letter string in pairs of each stimulus type presented during the test blocks. These are shown at the right of Table 2. Observations have been excluded from trials on which the response to the first string was incorrect. The data were analyzed in the same way as Experiment I.

Effects of Graphemic and Phonemic Relations

Positive responses to graphemically and phonemically similar words (Type 1) were again faster than responses to the corresponding controls (Type 2), although not reliably so. The difference was 16 ± 9 msec $[F(2,15) = 2.0, p > .10]$. The difference in error rates was also unreliable.

However, significantly slower positive responses occurred for the graphemically similar words that were phonemically dissimilar (Type 3) than for the corresponding control stimuli (Type 4). The difference was 34 ± 11 msec $[F(1,15) = 5.7, p < .05]$. A difference in error rates was also obtained but was not significant $(F < 1.0)$.

As in Experiment I, there was a reliable effect of varying phonemic similarity in addition to graphemic similarity for the WW pairs. The difference between mean RTs for Types 3 and 4 was 50 ± 7 msec greater

than the difference for Types 1 and 2 $[F(1,16) = 11.4, p < .01]$. A corresponding difference in the error rates was unreliable $(F < 1.0)$.

The only other significant main effect of graphemic and phonemic relations occurred for the NW pairs. Here responses to the words were 36 ± 2 msec slower when preceded by graphemically similar nonwords $[F(1,13) = 18.9, p < .01]$.

Effects of Lexical Status

Within the various kinds of stimulus pairs, words were usually classified more quickly than nonwords. For example, when the first letter string of a pair was a nonword, the response to the second string was 204 ± 20 msec slower on the average if it was also a nonword than if it was a word $[F(1,5) = 55.4, p < .01]$. Although the error rates for words and nonwords did not differ significantly in this case $(p > .10)$, errors tended to be more frequent on nonwords.

Conclusions

The results of Experiment II are consistent with those of Experiment I and substantiate the role of phonemic encoding in the process of visual word recognition. In particular, the difference between mean RTs for the graphemically similar words and their controls again depended on the phonemic relation involved. Since the letter strings were presented successively and the dependence occurred in RT measured separately for the

second string, this supports our conjecture that graphemic and phonemic properties of preceding words may influence phonemic encoding of subsequent words.

DISCUSSION

Various mechanisms could account for the results of both Experiments I and II. However, some of these are rather complicated, and we shall not try to consider all of them here. Instead, let us describe one simple possibility and show how it is consistent with most if not all of our findings.

Encoding-Bias Model

This theory depends on two facts: (a) processing of the letter strings was basically serial in both Experiments I and II, and (b) the grapheme-phoneme correspondence rules of English are not strictly deterministic, so that a string of letters does not always have a unique pronunciation. These points lead to the model in Fig. 2, which is an extension of the one proposed by Rubenstein et al (1971).

The model presumes that the recognition process for a pair of letter strings includes a series of stages. The sequence of operations for Experiment I is indicated by the solid and dashed lines in the figure. During the initial encoding stage, grapheme-phoneme correspondence rules are applied to form a phonological representation of the first letter string in the stimulus. Next a lexical (word-nonword) decision is made by accessing memory to determine whether or not the representation has been stored there previously. If it is not found in lexical memory and if the string has more than one possible representation, then the encoding and decision operations are repeated as shown by the dashed lines. This 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 first letter string is found in memory, then processing terminates, and a negative response is made. But if one of the representations is discovered in memory and if the item passes a subsequent spelling check, then a positive outcome occurs and processing continues to the second string of letters. At this point, the second string is encoded phonemically, and a decision about it is made by again accessing lexical memory. These operations are repeated until either a positive outcome is reached or all possible phonological representations of the second string have been checked. If the string is classified as a word, then a positive response is made; otherwise, the response is negative.

To deal with Experiment II, the preceding operations would be modified so that both strings of letters are processed serially regardless of whether the first string is a word or nonword. Separate responses would then be made to the two strings, based on their lexical status, and the response to the first string would be executed
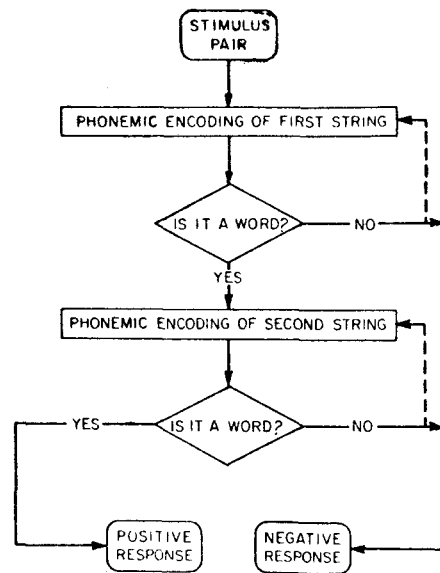

Fig. 2. An encoding-bias model to explain the effects of graphemic and phonemic relations on visual word recognition.

before the second string is encoded.

Assuming that this model is valid, our results can be explained by biases present during phonemic encoding of the second string in a pair. When the second string ends with the same letters as the first string, there may be a tendency to apply the most recently used grapheme-phoneme correspondence rules to encode the second string. Such an encoding bias would inhibit performance on pairs of words that are graphemically similar but phonemically dissimilar. For example, suppose that a word like FREAK has just been processed and that the string of letters B-R-E-A-K is encountered next. If it is encoded to rhyme with FREAK, then the resulting phonological representation will sound like BREEK, and it will not be found in lexical memory. To avoid an error, the S would have to recode the second string in the correct phonological representation and check it during another pass through memory. The repetition would produce relatively long RTs such as were observed in both Experiments I and II for graphemically similar words that are phonemically dissimilar. Likewise, an occasional failure to do the recoding would cause more incorrect responses, as also happened for those items.

A similar argument may account for why responses to words were slower in Experiment II when preceded by graphemically similar nonwords. During the processing of nonwords (and words), the most common grapheme-phoneme correspondence rules are probably applied first in the absence of biasing context. Subsequently, more unusual rules will be tried before a negative decision is reached. But when a word is preceded by a graphemically similar nonword, the word is more apt to be encoded initially in an unusual and incorrect phonological representation similar to the one

last tried for the nonword. This would make further encoding and memory retrieval necessary before the word is correctly classified, thereby increasing RT.[5]

In contrast, the encoding bias would either facilitate or not affect performance on pairs of words that are both graphemically and phonemically similar. If the second word has two or more possible phonological representations and if the correct representation is similar to the first word, then the bias should shorten RT because less encoding and memory retrieval would be done on the average. However, if the second word has only one possible representation, one might expect relatively little effect, since there would be no irrelevant pronunciations to check in the first place. This agrees with the results from both Experiments I and II, where there was a small although not statistically significant facilitation for rhyming words and where some of those words had only one possible pronunciation according to the rules of English.

The *encoding-bias model* is also consistent with some of our other results for nonwords. Once the phonological representation of a word has been found in memory, other possible encodings of it do not have to be checked. But for a nonword, every representation must be considered and rejected before a final negative decision, regardless of graphemic and phonemic relations to preceding letter strings. This could at least partly explain why mean RTs were longer for nonwords than words. It would also lead to relatively unreliable effects of graphemic and phonemic relations on decisions about nonwords, as were observed in both experiments.

Further Tests of the Model

The encoding-bias model has several properties that might be tested in future research. First, it 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 a 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.[6]

The model might also be applied to recognizing pairs of graphemically dissimilar words that are phonemically similar, e.g., GLUE-BLEW, which were omitted from the present studies. It does not necessarily predict that phonemic similarity alone would facilitate performance on such pairs. In fact, differences in the graphemic structures of these words could bias an S to form dissimilar phonological representations of them. This would inhibit recognition compared to performance on items that are both graphemically and phonemically dissimilar, just as recognition is inhibited for graphemically similar word pairs that are phonemically dissimilar (Types 3 and 4).

Finally, the model is relevant to recognizing individual words with ambiguous and unambiguous spellings (cf. Rubenstein et al, 1971). If a word is presented without biasing context and all other factors are constant, then the time required to recognize the word should vary directly with the number of potential pronunciations it has. Recognition would be fastest on the average when there is only one conceivable pronunciation, since no time would have to be spent accessing lexical memory for irrelevant phonological representations of the word. However, it may not be easy to obtain a strong test of this last property. The test requires measuring how many possible pronunciations a letter string has subjectively. This number could be difficult to determine, since the pronunciations may not be limited strictly by the objective grapheme-phoneme correspondence rules of English. Moreover, relatively long RTs would occur only for ambiguous words whose correct phonological representations are not the most likely ones. Thus, the phenomenon could be hidden by effects of other factors on which the words differ.

Implications for a Theory of Visual Word Recognition

Of course, our results do not prove that it is impossible to recognize printed words directly from their visual representations. Visual information is certainly sufficient for recognizing some nonverbal objects in the real world. Under various circumstances, people may also comprehend words directly from their visual representations. For example, this could be true of individuals who read nonalphabetic writing such as Chinese (Kolers, 1970, 1972). Similarly, direct visual recognition may play a role in processing English words that do not follow any of the usual grapheme-phoneme correspondence rules.

However, our results indicate that the graphemic-encoding hypothesis is not completely satisfactory for describing processes whereby printed words are recognized. Instead, it appears that visual word recognition is mediated at least part of the time through phonological representations. This implies that either the phonemic-encoding or dual-encoding hypothesis may be appropriate, with the encoding-bias model perhaps characterizing successive phonemic-encoding operations.

If the dual-encoding hypothesis is correct, then certain temporal relations must hold between graphemic and phonemic operations. Encoding and retrieval processes involving the phonological representation of a printed word must be fast enough that they are sometimes finished before the parallel visual operations. When a person's criterion is liberal, locating either the visual or phonological representation of a word in lexical memory may be sufficient to recognize it. Reaction time in this case would be determined by whichever of the two processes is faster. On other occasions when the criterion is conservative, recognition may not be completed until both kinds of representation have been found. Then RT would be determined by whichever of the two processes is slower.

In conclusion, it should be stressed further that the codes which help in recognizing printed words may depend on the type of task involved (cf. Baron, 1973). There are various "word recognition" paradigms that differ in the extent and kind of processing demanded. In tachistoscopic studies (e.g., Morton, 1969), the S must usually identify (write or pronounce) a presented string of letters, but he does not necessarily have to determine whether it is stored in lexical memory. A second type of task, such as the present one, calls for the S to verify whether a letter string is in memory but does not logically necessitate retrieving its specific meaning (Meyer & Ellis, 1970). Still other paradigms require a person to access detailed semantic information to categorize a word, phrase, or sentence (e.g., Baron, 1973; Collins & Quillian, 1969; Meyer, 1973). Such tasks also differ in the amount of context available to the S. There is more context when sentences are being presented than when a single word is displayed in isolation. Each of these factors could influence the relative importance of visual vs phonological representations. For example, if an S is provided with substantial context and forced to retrieve the specific meanings of printed words, then performance might depend more heavily on visual representations than it did in Experiments I and II, where there was less emphasis on meaning and relatively little context. An interesting problem for future research will be to determine under what circumstances the visual or phonological representations of printed words contribute most to their recognition.

APPENDIX

These are the pairs of words used in both Experiments I and II.

Type 1

Graphemically and phonemically similar pairs: PITCH-DITCH, LOAD-TOAD, MIGHT-TIGHT, POINT-JOINT, PLEA-FLEA, TRIM-GRIM, YELLOW-FELLOW, NATIONAL-RATIONAL, CRANK-FRANK, POISE-NOISE, VAST-PAST, DEAL-SEAL, GRACE-TRACE, GUILT-BUILT, TILT-WILT, CLOT-BLOT, MARK-DARK, BORN-WORN, HOUSE-MOUSE, SET-WET,

PORCH-TORCH, CANDLE-HANDLE, ZERO-HERO, RICE-VICE, TICKLE-PICKLE, BRUISE-CRUISE, FILE-TILE, SOFT-LOFT, NUMB-DUMB, COIL-BOIL, CREAM-DREAM, BLAME-FLAME, BRIBE-TRIBE, FENCE-HENCE, MINK-PINK, RUNG-SUNG, RAN-MAN, HILL-WILL, MUCH-SUCH, MADE-WADE, DISH-WISH, SENT-WENT, YIELD-FIELD, BARGE-LARGE, FAIL-SAIL, FOND-POND, VAULT-FAULT, PRICK-TRICK.

Type 2

Graphemically and phonemically dissimilar pairs: LOAD-DITCH, PITCH-TOAD, POINT-TIGHT, MIGHT-JOINT, TRIM-FLEA, PLEA-GRIM, NATIONAL-FELLOW, YELLOW-RATIONAL, POISE-FRANK, CRANK-NOISE, DEAL-PAST, VAST-SEAL, GUILT-TRACE, GRACE-BUILT, CLOT-WILT, TILT-BLOT, BORN-DARK, MARK-WORN, SET-MOUSE, HOUSE-WET, CANDLE-TORCH, PORCH-HANDLE, RICE-HERO, ZERO-VICE, BRUISE-PICKLE, TICKLE-CRUISE, SOFT-TILE, FILE-LOFT, COIL-DUMB, NUMB-BOIL, BLAME-DREAM, CREAM-FLAME, FENCE-TRIBE, BRIBE-HENCE, RUNG-PINK, MINK-SUNG, HILL-MAN, RAN-WILL, MADE-SUCH, MUCH-WADE, SENT-WISH, DISH-WENT, BARGE-FIELD, YIELD-LARGE, FOND-SAIL, FAIL-POND, PRICK-FAULT, VAULT-TRICK.

Type 3

Graphemically similar but phonemically dissimilar pairs: LEMON-DEMON, BLOW-PLOW, HOME-SOME, HAVE-CAVE, CATCH-WATCH, HORSE-WORSE, BOMB-TOMB, DULL-PULL, ROUGH-DOUGH, JURY-BURY, PAID-SAID, LOST-MOST, LOWER-TOWER, GIVES-WIVES, DEAR-WEAR, YOUTH-SOUTH, PUT-NUT, GROWN-CROWN, MINT-PINT, DOLL-TOLL, FOUL-SOUL, HUSH-BUSH, CROW-BROW, GASP-WASP, DROVE-PROVE, CASH-WASH, DIVER-LIVER, BAKED-NAKED, BEARD-HEARD, BONE-GONE, HANGER-RANGER, FLOWN-CLOWN, NOSE-LOSE, PATIO-RATIO, GOLF-WOLF, FOOD-HOOD, NATURE-MATURE, DAUGHTER-LAUGHTER, NASTY-HASTY, CLOVE-GLOVE, FEW-SEW, HONOR-DONOR, SOUR-FOUR, GATHER-FATHER, BOTH-MOTH, NEVER-FEVER, FREAK-BREAK, COUCH-TOUCH.

Type 4

Graphemically and phonemically dissimilar pairs: BLOW-DEMON, LEMON-PLOW, HAVE-SOME, HOME-CAVE, HORSE-WATCH, CATCH-WORSE, DULL-TOMB, BOMB-PULL, JURY-DOUGH, ROUGH-BURY, LOST-SAID, PAID-MOST, GIVES-TOWER, LOWER-WIVES, YOUTH-WEAR, DEAR-SOUTH, GROWN-NUT, PUT-CROWN, DOLL-PINT, MINT-TOLL, HUSH-SOUL, FOUL-BUSH, GASP-BROW, CROW-WASP, CASH-PROVE, DROVE-WASH, BAKED-LIVER, DIVER-NAKED, BONE-HEARD, BEARD-GONE, FLOWN-RANGER, HANGER-CLOWN, PATIO-LOSE, NOSE-RATIO, FOOD-WOLF, GOLF-HOOD, DAUGHTER-MATURE, NATURE-LAUGHTER, CLOVE-HASTY, NASTY-GLOVE, HONOR-SEW, FEW-DONOR, GATHER-FOUR, SOUR-FATHER, NEVER-MOTH, BOTH-FEVER, COUCH-BREAK, FREAK-TOUCH.

REFERENCES

Aarons, L. Subvocalization: Aural and EMG feedback in reading. Perceptual & Motor Skills, 1971, 33, 271-306.

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

Bower, T. G. R. Reading by eye. In H. Levin and J. P. Williams

(Eds.), *Basic studies on reading.* New York: Basic Books, 1970.

Brown, R. Psychology and reading. In H. Levin and J. P. Williams (Eds.), *Basic studies on reading.* New York: Basic Books, 1970.

Carver, R. P. Speed readers don't read; they skim. Psychology Today, 1972, 6, 22-30.

Chomsky, N., & Halle, M. *The sound pattern of English.* New York: Harper & Row, 1968.

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

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. Journal of Verbal Learning & Verbal Behavior, 1969, 8, 240-247.

Conrad, R. Acoustical confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-83.

Conrad, R. Speech and reading. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Mass: MIT Press, 1972.

Corcoran, D. W. J. An acoustic factor in letter cancellation. Nature, 1966, 210, 658.

Corcoran, D. W. J. An acoustic factor in proof reading. Nature, 1967, 214, 851-852.

Edfeldt, A. W. *Silent speech and silent reading.* Chicago: University of Chicago Press, 1960.

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

Gough, P. B. One second of reading. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Mass: MIT Press, 1972.

Hintzman, D. L. Articulatory coding in short-term memory. Journal of Verbal Learning & Verbal Behavior, 1967, 6, 312-316.

Hochberg, J. Components of literacy: Speculations and exploratory research. In H. Levin and J. P. Williams (Eds.), *Basic studies on reading.* New York: Basic Books, 1970.

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

Kolers, P. A. Three stages of reading. In H. Levin and J. P. Williams (Eds.), *Basic studies on reading.* New York: Basic Books, 1970.

Kolers, P. A. Experiments in reading. Scientific American, 1972, 227, 84-91.

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

LaBerge, D. Beyond auditory coding. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Mass: MIT Press, 1972.

Landauer, T. K. Rate of implicit speech. Perceptual & Motor Skills, 1962, 15, 646.

Landauer, T. K., & Streeter, L. A. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. Journal of Verbal Learning & Verbal Behavior, 1973, 12, 119-131.

Liberman, A. M. Some results of research on speech perception. Journal of the Acoustical Society of America, 1957, 29, 117-123.

Locke, J. L. Phonemic processing in silent reading. Perceptual & Motor Skills, 1971, 32, 905-906.

McLaughlin, G. H. Reading at "impossible" speeds. Journal of Reading, 1969, 12, 449 ff.

Meyer, D. E. Correlated operations in searching stored semantic categories. Journal of Experimental Psychology, 1973, 99, 124-133.

Meyer, D. E., & Ellis, G. B. Parallel processes in word recognition. Paper presented at the meeting of the

Psychonomic Society, San Antonio, November 1970.

Meyer, D. E., & Schvaneveldt, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 1971, 90, 227-234.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Activation of lexical memory. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1972.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Loci of contextual effects on visual word recognition. In P. M. A. Rabbitt and S. Dornic (Eds.), *Attention and performance V.* London: Academic Press, in press.

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

Neisser, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

Pierce, J. R., & Karlin, J. E. Reading rates and the information rate of a human channel. Bell System Technical Journal, 1957, 36, 497-516.

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

Ruddy, M. G., Meyer, D. E., & Schvaneveldt, R. W. Context effects on phonemic encoding in visual word recognition. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1973.

Schaeffer, B., & Wallace, R. J. The comparison of word meanings. Journal of Experimental Psychology, 1970, 86, 144-152.

Schvaneveldt, R. W., & Meyer, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and performance IV.* New York: Academic Press, 1973.

Shannon, C. A. Prediction and entropy of printed English. Bell System Technical Journal, 1951, 30, 50-64.

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

Sperling, G. Successive approximations to a model for short-term memory. Acta Psychologica, 1967, 27, 285-292.

Sperling, G., & Speelman, R. G. Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

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

Webster's New Collegiate Dictionary. Springfield, Mass: G. & C. Merriam, 1961.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

NOTES

1. In addition to the three main hypotheses, others may be needed to deal with special situations that we shall not consider here. For example, Conrad (1972) conducted a series of studies on deaf Ss who had never heard speech. He found some evidence that these people recognized printed letters directly from their visual representations. Other results indicated that phonemic encoding may have been involved too. However, it appeared as if at least some of the Ss processed the letters by encoding them in the finger movements of standard sign language. This would require a fourth theory of recognition.

2. It is difficult to measure the redundancy of English text precisely. However, one estimate suggests that the figure may be over 50% (Shannon, 1951). Thus, a reader might need to process no more than half of the material in an ordinary passage to

make sense of it. This could bring a nominal rate of 1200 words/min down to an actual rate of around 600 words/min, which would seem to be within the limits of covert speech.

3. In the analysis, we calculated "quasi F ratios" by comparing linear combinations of the mean squares (MSs) for treatments, treatments by Ss, treatments by stimulus pairs, and treatments by Ss by stimulus pairs (see Winer, 1962). The reliability of the main effects of treatments was tested in two ways: (a) The MS of the main effect was divided by the sum of the MSs of the second-order interactions less the third-order interaction, and (b) the sum of the MSs of the main effect and third-order interaction was divided by the sum of the MSs of the two second-order interactions. Since the statistical properties of the latter ratio are more regular, we report it, estimating degrees of freedom as outlined in Winer (1962). However, the standard deviations presented above were derived from the sum of the MSs of the second-order interactions less the third-order interaction, which is the appropriate statistic in this case. Degrees of freedom vary from one test to the next because they depend on the MSs involved.

4. Here the larger number of degrees of freedom in the numerator results from properties of the procedure for computing quasi F ratios (Winer, 1962).

5. Although the results for NW pairs in Experiment I failed to reveal an inhibitory effect of graphemic similarity, they are not inconsistent with the basic argument or with the inhibitory effect produced by NW pairs in Experiment II. This is because processing in Experiment I may have terminated immediately after the first nonword, so that graphemic and phonemic relations with the second string would have been irrelevant then. Such self-termination could also explain why there were no effects of graphemic or phonemic relations for NN pairs in Experiment I.

6. An informal examination of the stimuli in Experiments I and II suggests that, among the Type 3 and 4 word pairs, a majority of the second members may have involved less common pronunciations than the corresponding first members. According to the encoding-bias model, this means that if the order of the members were reversed in a subsequent experiment, then the difference between RTs for the Type 3 and 4 pairs could be reduced.