Journal of Experimental Psychology
1973, Vol. 100, No. 2, 349–353

PERCEPTION OF LETTERS, WORDS, AND NONWORDS [1]

DOMINIC W. MASSARO [2]

University of Wisconsin

Letter vs. word and word vs. nonword identification were compared, with redundancy adequately controlled. The processing time of the test stimulus was varied to provide a number of levels of correct performance. The first experiment showed that letters were recognized better when presented alone than when embedded in words. In the second experiment, the identification of letters in words did not differ from the identification of letters in nonwords. These results conflict with earlier findings that have shown that a letter is better identified in a word than in a nonword or presented alone. The differences in the experimental procedures indicate either that redundancy may not have been adequately controlled in the earlier studies or that some other process besides a perceptual one might account for the results.

Recent experiments have demonstrated that a letter is more easily identified when it is presented in a word than when it is presented alone or in a sequence of unrelated letters (Reicher, 1969; Wheeler, 1970). These results were considered important because it was argued that redundancy could not be responsible for the effect. Redundancy or the knowledge of the constraints which characterize letter sequences in words can enhance identification by simply delimiting the set of possible response alternatives. Since Reicher's experimental procedure supposedly eliminates the contribution of redundancy, any word advantage is, therefore, taken as a pure perceptual, rather than decision, effect.

In Reicher's (1969) paradigm, *S*s are presented with either a letter, a word, or a nonword. Immediately after the stimulus flash, a visual noise mask is presented which covers the former stimulus position. Two letter alternatives, one of which appeared in the original stimulus and one of which is incorrect, are presented either 1 or 2 sec. before or after the test stimulus. To eliminate the operations of redundancy on word trials, both letter alternatives form a common word with the other letters of the word. The results from this paradigm indicate that *S*s are more likely to choose the correct alternative when the stimulus is a word or spelling pattern than when the stimulus is a letter or a nonword (Aderman & Smith, 1971; Reicher, 1969; Wheeler, 1970).

Thompson and Massaro (1973) replicated these findings by showing that identification of letters embedded in words was 9% superior to identification of single letters. However, a second manipulation in that experiment had no effect. The incorrect test alternative could be visually similar or dissimilar to the test letter. In contrast to expectations, identification given a distinctive incorrect test alternative did not exceed performance with a similar incorrect test alternative. From these results, Thompson and Massaro reasoned that *S*s are not able to incorporate the knowledge of the test alternatives into their perception of the test stimulus. That is to say, in the Reicher (1969) paradigm, when the 2 test alternatives are presented immediately before or after the test stimulus, *S* appears to arrive at a decision about the identity of the test stimulus without considering the 2 test alternatives.

Given that *S*s arrive at a decision about the test stimulus before the alternatives are considered, redundancy is still available and can facilitate performance when letters are presented in words. During

[1] This research was supported in part by U.S. Public Health Service Grant MH-19399-02. The author wishes to thank Jeffrey J. Moore for writing the computer program, John Hammerly for testing the *S*s and performing the data analysis, and David A. Grant for a reading of the manuscript.

[2] Requests for reprints should be sent to Dominic W. Massaro, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

word presentations, redundancy would serve to reduce the number of alternatives and, therefore, *S* would need less visual information for correct letter identification (see Thompson and Massaro, 1973, for a complete description of how redundancy operates in the Reicher, 1969, paradigm). A second experiment was therefore designed with a better control for redundancy. In this study, *S* was given a fixed number of response alternatives (4) before the experimental session. Given sufficient practice with this fixed set of alternatives, *S* should be able to incorporate the knowledge of the alternatives into the perceptual process eliminating the effects of redundancy. The results eliminated the word advantage effect: letters presented alone were actually better recognized than letters presented in words. A large effect of visual confusability was also shown for both letters in words and letters presented alone. Since both of these results were opposite those found in the Reicher paradigm, Thompson and Massaro concluded that it cannot be assumed that presenting the response alternatives immediately before or after the stimulus trial eliminates the operation of redundancy in word recognition. Redundancy, therefore, can account for previous findings that letters are recognized better in words than when presented alone or in nonwords (Aderman & Smith, 1971; Reicher, 1969; Wheeler, 1970).

However, the results of Thompson and Massaro (1973) conflict with the results of Smith and Haviland (1972). In that study, *S*s were presented with 3-letter words or consonant trigrams that were experimentally equated for redundancy. The *S*s were completely informed about the nature of the redundancy for both the words and the consonant trigrams. The *S*s were then tested using Reicher's (1969) forced-choice procedure. In 2 experiments, the results indicated that a letter was better identified when embedded in words than in consonant trigrams. Therefore, in contrast to Thompson and Massaro's results, the word advantage occurred even though *S*s were practiced with the possible alternatives before the experimental session.

The present study deals with 2 primary methodological differences between Smith and Haviland's (1972) study and Thompson and Massaro's (1973) experiment. Smith and Haviland presented blocks of trials under the word and trigram conditions whereas Thompson and Massaro presented letters or words randomly from trial to trial. This difference could be a critical one since Aderman and Smith (1971) have shown that expectancy affects the size of the word advantage effect. Therefore, the first experiment replicates the Thompson and Massaro study while controlling for expectancy by presenting the letter and word conditions on different days. The second experiment provides a word–nonword comparison, since Smith and Haviland compared word vs. nonword recognition whereas Thompson and Massaro compared word vs. letter recognition. In both experiments, the amount of processing time is varied to provide a more complete test by measuring performance at a number of levels of accuracy.

Experiment I

Method

Subjects. Two male and 2 female *S*s from an introductory psychology course at the University of Wisconsin volunteered for the present study.

Procedure. The test stimuli in the present experiment consisted of 4 letters or 4 words. The test letters were P, R, C, and G; the test words were APE, ARE, ACE, and AGE. The test stimulus duration was 1 msec. for all *S*s. The test stimulus was either presented alone or followed by a masking stimulus after a variable blank interval. The duration of the blank interval was either 10, 20, 40, 70, 110, 160, or 240 msec. All 8 processing conditions were equally likely to occur on any trial. The masking stimulus was presented over the 1 or 3 letters in the letter and word conditions, respectively. The masking stimulus consisted of 2 msec. each of the characters 1, H, and # presented twice in that order, totaling 12 msec. in duration.

Four *S*s could be tested simultaneously in separate sound-attenuated rooms. All experimental events were controlled by a PDP 8/L computer. The letters were presented over visual displays made of light-emitting diodes (Mansanto Model MDA 111). These displays are described in Nealis, Engelke, and Massaro (1973). The displays were viewed from a distance of about 55 cm. The height

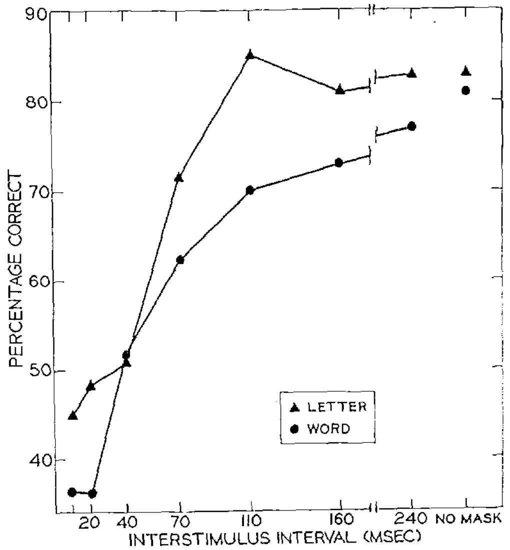

FIG. 1. Percentage of correct identifications of a letter presented alone or embedded in a word as a function of the duration of the interstimulus interval.

and width of the characters subtended visual angles of 50′ and 38′, respectively. The horizontal visual angle of the 3-letter displays was 3°20′. The *S* made his response by pressing 1 of 4 push buttons mounted on a panel in front of him.

Each trial began with presentation of the test stimulus. The *S* then had a 2-sec. response interval followed by presentation of the correct alternatives for 500 msec. The intertrial interval was 1.5 sec. On the first day, *S*s were given one session of 300 trials under the letter condition and a session of 300 trials under the word condition. The letter and word conditions were then alternated on the next 4 days. Two of the *S*s were given the letter condition on the second day; 2 were given the word condition. Two sessions of 300 trials each were given per day. Within a session, the 4 test alternatives and the 8 processing conditions were presented in a completely random order with equal probability.

Results

Table 1 presents the results of letter and word identification for each of the 4 *S*s. All *S*s identified the letter more often when it was presented alone than when it was embedded in a word, F (1, 3) = 16, $p < .05$. Figure 1 shows that identification performance improved significantly with increases in the blank interval between the test and masking stimulus, F (7, 21) = 89, $p < .001$. This result occurred for both the letter and word conditions.

The size of the letter advantage varied unsystematically across the 8 processing intervals with a slight reversal at 40 msec.; the interaction between test stimulus condition and processing interval was significant, F (7, 21) = 3.35, $p < .025$. This interaction provides some measure of the unreliability of the results in this paradigm. Each of the 4 *S*s contributed about 150 observations to each data point. Apparently this is not enough. In contrast, the main effect of letter vs. word has 1,200 observations per data point from each *S*, which provides a much more reliable estimate of performance. Most of the letter–word studies measure performance at only one accuracy level with 20–250 observations per data point per *S*. These studies seem highly susceptible to a Type I error, especially considering the number of investigators who have failed to find a word advantage and have not reported it.

The results of the present experiment replicate the results of Thompson and Massaro (1973). A letter is better recognized when presented alone than when embedded in a word, even though the expectancy of the *S* is confirmed on each trial. This result substantiates the argument of Thompson and Massaro that word advantage effect in the Reicher (1969) paradigm is a result of the opera-

TABLE 1

PERCENTAGE OF CORRECT IDENTIFICATIONS FOR EACH *S* AS A FUNCTION OF THE LETTER AND WORD CONDITIONS IN EXPERIMENT I AND THE WORD AND NONWORD CONDITIONS IN EXPERIMENT II

Experiment I			Experiment II		
S	Letter	Word	*S*	Word	Nonword
1	68	64	5	63	70
2	58	49	6	62	73
3	73	67	7	71	67
4	75	63	8	61	60
Average	69	61	Average	64	68

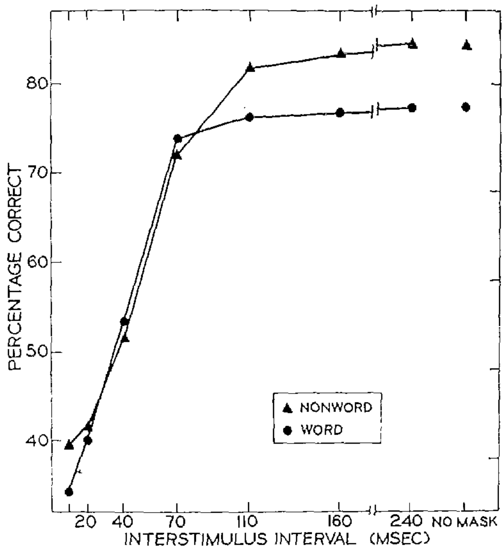

FIG. 2. Percentage of correct identifications of a letter embedded in a word or a nonword as a function of the duration of the interstimulus interval.

tion of redundancy. When redundacy is eliminated, as in the present experiment, the word advantage is eliminated.

Thompson and Massaro (1973) hypothesized that lateral masking was responsible for the letter advantage over words. That is, the presence of the outside letters seemed to interfere with perception of the middle letter in the word condition. Otherwise, since redundancy was controlled, there should have been no difference between the letter and word conditions. Accordingly, identification of a letter embedded in a word should not differ from identification of a letter in a nonword when the amount of lateral masking is controlled. To test this prediction, a word–nonword comparison was provided in the second experiment.

Experiment II

Method

Subjects. Four undergraduates volunteered for the present study as a class requirement for introductory psychology.

Procedure. Experiment I was replicated exactly except that the alternatives VPH, VRH, VGH, and VCH were employed instead of the letter alternatives P, R, G, and C. The letters V and H were chosen as outside letters since their inside features are similar to the features of the letters A and E used in the word condition. All other procedural details were exactly the same as in Experiment I.

Results

Table 1 shows that recognition of words did not differ significantly from nonwords, $F\ (1,\ 3) < 1$. Figure 2 shows that recognition performance improved with increases in the blank interval in both the word and nonword conditions, $F\ (7,\ 21) = 132$, $p < .001$. This main effect of processing time provided a test of the word vs. nonword conditions at a wide range of overall performance (35–80%). The word–nonword conditions did not differ significantly from each other at any of the 8 processing intervals; the interaction of word vs. nonword and processing time was nonsignificant, $F\ (7,\ 21) = 1.72$, *ns*. Given the large effect of the processing time variable, the failure to find a difference between the word–nonword conditions cannot be attributed to a weak experimental test.

Discussion

The present results support the findings of Thompson and Massaro (1973), which show that the word advantage is eliminated when redundancy is properly controlled. This hypothesis is consistent with all of the literature except the study by Smith and Haviland (1972). It is possible that another process instead of a perceptual process accounts for the Smith and Haviland finding. Although Smith and Haviland's *S*s knew the possible response alternatives before each session, they were not cued with respect to the position of the letter to be reported until after the test presentation. Furthermore, their *S*s had to look at a choice card that was presented outside of the tachistoscope. Accordingly, it is possible that there was differential forgetting during this time for the word and consonant trigram presentations. Assume that *S* recognizes the 3 test letters equally on word and nonword trials. To remember these, however, he must rehearse them until he decodes the probe card outside the tachistoscope. In the word case the *S* can rehearse a word during this interval, whereas he must remember 3 separate letters in the nonword condition. The phonemic similarity of the letters was also fairly high, which would make rehearsal of the separate letters more difficult. The smallest forgetting difference between the 2 conditions could account for all of the word advantage effect since it was so small (4%–7%).

The recognition of a string of letters involves a readout of the visual features of the letters available in a given eye fixation. In short tachistoscopic presentations, the number of visual features that are extracted is limited and recognition is not completely accurate. However, redundancy or the rules of English orthography can operate at the recognition stage of processing to improve performance when the letter strings are spelling patterns or words.

Consider a particular trial in which the word WORD is presented. Assume that the *S* has some visual information about each letter. The visual features that were extracted reduce the letter alternatives to V or W, C or O, R or P, and D or B in the 4 positions, respectively, giving 16 possible letter sequences. However, if the letters must spell a word, WORD is the only valid alternative and it is recognized as such. If the letter D is presented alone or in a nonword, the recognition process cannot use the spelling rules of English to reduce the number of alternatives for the tested letter. When the partial visual information limits the alternatives to D or B, *S*'s recognition will be correct only half the time. On half the trials, he sees a D; on the other half, he sees a B.

According to this model, redundancy operates at the recognition stage of information processing, not at the response selection stage as implicitly assumed in Reicher's (1969) paradigm. Perception of the letters occurs before the probe letters are considered at test. This theoretical analysis poses a particular problem for studies of word vs. nonword identification that attempt to equate the items for psychological redundancy. How long must *S* practice with a set of new rules to utilize them in perception? In this light, the differences between the findings of Smith and Haviland (1972) and the present study (Experiment II) can be explained. Smith and Haviland's *S*s alternated between the word and nonword conditions every 16 or 64 trials. Since the word and nonword conditions contained conflicting rules, it would not be surprising if the redundancy rules in the word condition were better utilized since they have been employed for a lifetime of reading. In the present study, the word and nonword conditions were better equated for the utilization of redundancy since the word and nonword conditions were alternated from day to day with 600 trials per day, allowing *S* more time with the redundancy rules in each condition. This analysis predicts that the word advantage found by Smith and Haviland should be eliminated if a substantial number of trials are given together in the nonword condition so that *S*s can effectively employ the redundancy rules.

In summary, the present results support the hypothesis that the letter is the basic perceptual unit in letter, nonword, and word identification. Each letter is uniquely represented by a property list of visual features in long-term memory. Redundancy can operate at the recognition stage of information processing. A string of letters can be correctly identified given partial visual information, if the letters conform to definite spelling rules that are well learned and utilized by the reader.

REFERENCES

ADERMAN, D., & SMITH, E. E. Expectancy as a determinant of functional units in perceptual recognition. *Cognitive Psychology*, 1971, 2, 117–129.

NEALIS, P. M., ENGELKE, R. M., & MASSARO, D. W. A description and evaluation of light emitting diode displays for generation of visual stimuli. *Behavior Research Methods and Instrumentation*, 1973, **5**, 37–40.

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

SMITH, E. E., & HAVILAND, S. E. Why words are perceived more accurately than nonwords: Inference versus unitization. *Journal of Experimental Psychology*, 1972, **92**, 59–64.

THOMPSON, M. C., & MASSARO, D. W. Visual information and redundancy in reading. *Journal of Experimental Psychology*, 1973, **98**, 49–54.

WHEELER, D. D. Processes in word recognition. *Cognitive Psychology*, 1970, 1, 59–85.

(Received November 16, 1972)