Journal of Experimental Psychology
1974, Vol. 102, No. 6, 963–968

WORD RECOGNITION AND MORPHEMIC STRUCTURE [1]

GRAHAM A. MURRELL [2] AND JOHN MORTON

Medical Research Council, Applied Psychology Unit, Cambridge, England

Sixteen *S*s learned a word list and were then tested on recognition of words presented tachistoscopically. When a test word was identical to one of the learned words, recognition was facilitated in the usual way. Pretraining with a word that was a different derivative of the same root morpheme as a test word gave significant facilitation of recognition of that test word. However, learning a word with comparable visual–acoustic similarity but no morphemic relation to a test word gave only slight, nonsignificant facilitation of recognition. It was concluded that the process of recognizing a word involves assigning it to a linguistic unit with specific semantic associations, i.e., a morpheme. Analysis of error responses suggested that, to some extent, root and suffix morphemes are recognized independently, and that suffixes themselves may be subject to frequency effects. The principles underlying the phenomena studied may be most appropriately characterized as "the morpheme-frequency effect."

Words of higher frequency of occurrence in the language are more easily recognized than low-frequency words. The nature of the mechanism leading to this phenomenon is a subject of current debate (Broadbent, 1967, 1971; Catlin, 1969; Morton, 1968; Nakatani, 1970; Treisman, 1971). All current theories, however, are agreed that the recognition system is biased toward high-frequency words, although that bias is seen as operating in different ways. In the past, it was debated whether this bias was a function of the differential frequency of the perceptual experience of words (Howes, 1954; Neisser, 1954) or the frequency of production of words (Daston, 1957), with corresponding definition of the unit of recognition as perception or response based.

A way of reconciling these positions was put forward by Morton (1964, 1969), who proposed the logogen model, in which the central units, logogens, were affected not only by visual and acoustic information, but also by semantic information. The latter is implicated in speech production and also in the recognition of connected material. Any logogen receiving more than a threshold amount of information "fires," making the corresponding word available as a response. As a result of firing, the threshold of the logogen is lowered and very gradually rises to a level slightly below its previous value. In this way, experience with words, both recognition and production, leads to permanent changes in threshold, biasing future responses.

The short-term change in threshold was included in the model to account for the finding that when a word is recognized, the same word will be recognized more easily if presented again within a short time.

The latter effect was utilized in the present experiment to create a short-term, experimentally controlled word-frequency effect. The technique has been used previously by Solomon and Postman (1952), who used nonsense words so there would be no "a priori" word-frequency effect. As will become clear, it was necessary to use familiar real words in the present experiment, but the uncontrolled a priori frequency effect could only mask the results and will not affect the conclusions.

Neisser (1954) found that pretraining with a word facilitated recognition of the same word but not of a homonym, e.g., after priming with "phrase," recognition of this word was easier than normal, but recognition of "frays" was not affected.

[1] This work was carried out at the Psychological Laboratory, University of Cambridge, while the first author was a final year undergraduate. The assistance of those who acted as *S*s is gratefully acknowledged.

[2] Requests for reprints should be sent to Graham A. Murrell, Medical Research Council, Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.

TABLE 1
EXPERIMENTAL WORDS

(1)	(2)	(3)
Car	Cars	Card
Bored	Boring	Born
Sings	Singer	Singe
Pains	Pained	Paint
Lives	Lively	Livery
Sees	Seen	Seed
Man	Manned	Manner
Hanging	Hang	Hangar
Reader	Reading	Ready
Placed	Places	Placate
Walled	Wall	Wallet
Topped	Topping	Topple
Basis	Basic	Basin
New	Newer	Newt
Mile	Miles	Mild
Numbing	Numb	Numbers

His interpretation was that pretraining "does not merely raise the probabilities of certain verbal responses, but rather facilitates the perception of specific visual patterns [Neisser, 1954, p. 402]." Following this, Ross, Yarczower, and Williams (1956) attempted to demonstrate a monotonic relation between degree of facilitation and level of visual similarity between pretraining word and test word. They found no such relation, although they confirmed Neisser's main result. This, and other considerations, casts doubt upon Neisser's interpretation, and particularly his assumption that "phrase" and "frays" may be regarded as the same response (even assuming they are phonetically identical).

The interaction of semantic and sensory information necessary to account for the facilitative effects of a context on word perception (Morton, 1969; Rubenstein & Pollack, 1963; Tulving, Mandler, & Baumal, 1964) also necessitates a refinement of the definition of the unit of recognition. Thus, since the contexts appropriate to the two homographs "chop" are totally different, they might be expected to have functionally separate representations in the recognition system.

The question of interest in the present study is whether the unit concerned in recognition facilitation effects is a word as a visual–phonetic pattern, or a form bound to a specific semantic unit, i.e., a morpheme. The conclusions will bear on the nature of the response units made available by logogens.

The experiment investigated the effect on word recognition of pretraining with a word containing the same root morpheme, compared with pretraining with a word containing the same visual–phonetic pattern as its "root" when this is not actually the same morpheme. The effects of pretraining with the identical word and zero pretraining, as control conditions, were also tested.

METHOD

Subjects. Sixteen male undergraduates volunteered as *S*s, and were divided into two groups of 8. None were psychology students. Each completed the procedure once only, with no prior information about the experiment apart from the specific instructions (see below).

Materials. The apparatus was a Cambridge tachistoscope with an auxiliary timing device to permit control of flash duration in 5-msec. steps between 10 msec. and 140 msec. A neutral density filter (×16) was fitted over the viewing aperture to reduce stimulus contrast. The stimulus words were typed in lowercase on plain white cards, aligned so as to appear in the space between two short horizontal lines drawn on the "background" card upon which a word was superimposed in the stimulus flash. The *S* was provided with the tachistoscope trigger button on a cable, to initiate the flashes.

The experimental words are shown in Table 1. The three words in each row of the table have an initial sequence of letters in common. Group 1 and 2 words differ only in the inflection of their common root morpheme, but the same letter sequence in the Group 3 word is not the same morpheme, and the word bears no semantic relation to the other two. However, viewed purely as visual–acoustic patterns, the degrees of similarity between the words in different columns are comparable (e.g., the mean suffix length of Group 3 is intermediate between Groups 1 and 2, the difference between the means of Groups 1 and 2 being less than half of one letter).

Twenty other words were used in initial practice on tachistoscopic recognition. These were selected for low similarity with experimental words.

There were four sets of 12 words for pretraining. Each set comprised 4 words from each of Groups 1, 2, and 3, typed on a card in random order. Each *S* received a recognition test on the words of either Group 1 or Group 2 of Table 1. Thus, of the 16 words each *S* was tested on, 4 were subject to each of the four pretraining conditions, which were as follows: (*a*) pretraining with the same word (Condition S), (*b*) pretraining with a different inflection of the same root morpheme (M), (*c*) pretraining with an unrelated word containing the same visual–acoustic form but different terminal letters (V), and (*d*) zero pretraining, no similar word having been presented (O).

Over the eight combinations of pretraining and testing word sets, every test word was subjected to every pretraining condition. Thus from each group of eight *S*s, recognition scores for each word under each condition were obtained.

Procedure. The outline of the procedure is best conveyed by quoting the instructions read to *S*s.

> The experiment is in three phases. First, I am going to test your ability to recognize words that are presented to you in a very brief flash. Secondly, I will give you a short list of words to learn. Thirdly, I will test your recognition of a further series of words on brief exposure, some of these words being among those you have learned. Please place your face against the rubber visor on this apparatus, and view the white card. When I say 'now,' focus between the lines and when you're ready press the button once to flash up the word. Then, immediately say aloud and distinctly what you think the word was. Guess if you're not sure, and if you have no idea what it was, say 'no.' [practice phase]
>
> For the next phase, there are 12 words printed on this card in a random order. You have three minutes to learn them and go over them in your head. You will need to keep them in your mind through the third phase. Begin now. [pretraining]
>
> Time is up, and it is time for the third phase, but try to remember those words. Some of the following series of words will be among the words you have learned, some will be similar to those you have learned, and others will be quite different words. The procedure is the same as before. When I say 'now,' focus between the lines . . . . [testing]

In the practice period, about 100 word exposures were given, of varying durations, to familiarize *S* with the task. The shortest duration at which words were correctly recognized at the end of the practice was taken as the individual *S*'s baseline for testing.

The set of words for testing was presented in a random order at baseline flash duration, then in the same order again at 5 msec. longer duration, and so on for nine serial presentations. Baselines varied from 20 msec. to 60 msec., and each *S* received exposures increasing in duration to 40 msec. above his baseline.

All responses in the test were recorded, but no knowledge of results of any kind was given.

Results and Discussion

There were 39.8% correct responses, 48.1% error responses and 12.1% "no" responses. Only 7.0% of the error responses were not real words. Of these, nearly three-quarters were from four of the *S*s, and five *S*s gave no such responses. In general, error responses were similar in word length and gross shape to their eliciting test words.

TABLE 2

Recognition Performance

Scoring measure	Pretraining condition			
	S	M	V	O
Mean recognition score for words	86.8	63.3	51.1	46.8
Mean recognitions per 9 presentations	4.35	3.72	3.25	3.02
Percent recognition of root morphemes	49.9	46.4	36.8	36.6
Percent recognition of suffix morphemes	62.2	54.2	51.6	51.0

Note. See text for explanation of pretraining conditions.

It was decided not to use a "threshold" measure for comparison of scores, since it would depend heavily upon the arbitrary criterion set by *E* and would neglect much of the available data. Instead, a measure of performance was devised which gave a value to every correct response according to exposure difficulty, normalized across *S*s.

First, each correct response was given a value equal to the reciprocal of the proportion of correct responses at that exposure over all *S*s. This weighting gives higher values to more difficult discriminations. Thus weighted, correct responses were summed over the nine exposures to give a score for each word by each *S*. The scores of each *S* were further weighted by the multiplier necessary to correct his total score to 100. This procedure compensated for individual differences.

The mean recognition scores under the four conditions are shown in the first row of Table 2. The untreated data are summarized in the second row of Table 2 as the mean number of wholly correct recognitions per nine presentations under each condition. No further reference will be made to the latter figures.

A three-factor analysis of variance was carried out on the recognition scores (model from Winer, 1971, p. 380). The factors were conditions, words, and replications (i.e., the two sets of eight *S*s duplicating each result). The pretraining conditions differentially affected word recognition, F (3, 93) = 12.65, $p < .001$.[3] The result of principal interest is that although the difference between scores under Conditions V and O fell well short of reliability ($F < 1$), Condition M did facilitate word recognition compared with O, F (1, 93) = 5.38, $p < .025$. However, the effect of Condition M was rather less than that of S, F (1, 93) = 10.79, $p < .01$. The Replications × Conditions interaction was not significant ($F < 1$), indicating that the differences between conditions were consistent between the two groups of *S*s. Similarly, the Words × Conditions interaction was not significant, F (93, 93) = 1.32, indicating that the conditions had the same effects over all the test words used.

There is some indication of effects of phonological similarity, which may be worthy of study in their own right. In six of the rows of words in Table 1, the "root" of Word 3 differed phonologically from those of Words 1 and 2, e.g., reader, reading, ready. A separate analysis of the recognition of the 12 test words from these rows indicated that under Condition V the phonological dissimilarity tended to counterfacilitate recognition. However, the analysis of variance of recognition scores was repeated with these 12 words excluded, and even when the pretraining and test word roots had the same phonology under Condition V, there was no significant facilitation compared with O, F (1, 57) = 1.20.

The results confirm the effectiveness of the pretraining technique for creating a short-term, experimentally controlled word-frequency effect. This effect does not occur when the pretraining word resembles the test word in visual–acoustic features only. There is reliable facilitation only when there is morphemic identity in addition to visual–acoustic similarity.

Considered in terms of the logogen model, this means that the logogen whose threshold is lowered by pretraining with a word corresponds not to that word, nor even to the visual–acoustic pattern it contains as root, but strictly to the morpheme from which the word is derived. Thus, the word recognition process, even in the absence of contextual or semantic stimulus information, involves assignment of the stimulus to a particular morpheme with its specific semantic associations. The morpheme, and only that one morpheme, to which the decision process assigns the word, has its logogen threshold lowered. Thus, when a pretraining word belongs to a different morpheme from a similar test word (as in Condition V), no facilitations would be expected. This interpretation would also account for the negative results of Ross et al. (1956).

The conclusion that the unit involved in facilitation effects is the morpheme was further illustrated by the percentage of responses containing the same root morpheme as the eliciting test word. These results are shown in the third row of Table 2. Bearing in mind that the recognition of whole words was considerably better under Condition S than under M, it is significant that the recognition of root morphemes under these two conditions did not differ reliably. It may be concluded that the latter result was due to the equal facilitation of root morphemes in these two pretraining conditions. Furthermore, recognition of root morphemes was not reliably different under Conditions V and O. This would be predicted, since under Condition V the root morpheme of the test word did not occur in pretraining and thus could not have been facilitated.

Even though the root *letter sequences* in pretraining were the same in Conditions M and V, with suffixes differing from the test words in both cases, the recognition of root *morphemes* was better under M ($t = 2.38$, $p < .025$). The facilitative effects of V pretraining were apparent in the intrusion

[3] In order to avoid the "language-as-fixed-effect fallacy" pointed out by Clark (1973), the appropriate quasi-F ratio for the conditions main effect was also calculated: F' (3, 17) = 8.03, $p < .01$. This provides statistical support for generalizing the effect to any comparable sets of words.

errors containing the root morpheme of the pretraining word. Such errors amounted to 8.7% of the responses under V, a figure comparable with the difference in root morpheme recognition between M and V. This is evidence that response biases of similar magnitude were created under these two conditions.

There appears to be some validity in considering experimental words in terms of root and suffix. The results suggest that, to some extent, free morphemes and suffixes are recognized independently. This is consistent with the linguistic analysis of words into morphemes, and the view that, the response units made available by logogens are morphemes.

It appears that the difference in word recognition scores between Conditions S and M was due to the pretraining with the test word suffixes under S, which did not occur in the other conditions. The scores for recognition of suffix morphemes are shown in Table 2. The difference in suffix recognition between S and O ($t = 2.56$, $p < .025$) demonstrates suffix facilitation. The differences in suffix recognition between M, V, and O were not reliable. The difference between S and M was only poorly reliable ($t = 1.59$, $p = .07$), probably because suffix recognition under M was inflated by the constraint imposed on response suffixes if the pretrained root was correctly produced. Thus, if the test word were "newer," then the likelihood of producing the correct suffix "-er" is zero if the root has been reported as "newt." If "new" has been seen correctly, the likelihood of guessing "-er" is appreciable.

Of the errors which were real words given in response to test words having the zero morpheme (i.e., no suffix), 213 (93.0%) had no suffix. These responses were not counted in the suffix recognition scores of Table 2, which are therefore an underestimate.

It is clear that there was substantially better recognition of suffixes than of roots. It may simply be that the suffixes were more distinctive visual patterns. The suffix "-ing" was correctly recognized on 81.5% of the trials in which it occurred in the test word, including 35.0% of the trials on which the first part of the response was wrong. For the suffix "-ed," there was recognition on 65.1% of its occurrences, including 28.3% error responses. Correct recognition of "-er" was 53.7%, including 15.3% errors.

It may be that "-ing" is more distinctive than "-ed." However, it is an attractive proposition that the suffixes themselves are subject to a frequency effect. If this were so, the relative proportions of trials on which the root and suffix of a test word were recognized would reflect the relative familiarity of the two morphemes to the subject.

This experiment gives grounds for considering a well-known phenomenon in new terms, as "the morpheme-frequency effect."

REFERENCES

BROADBENT, D. E. Word-frequency effect and response bias. *Psychological Review*, 1967, **74**, 1–15.

BROADBENT, D. E. *Decision and stress.* New York: Academic Press, 1971.

CATLIN, J. On the word frequency effect. *Psychological Review*, 1969, **76**, 504–506.

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

DASTON, P. Perception of idiosyncratically familiar words. *Perceptual and Motor Skills*, 1957, **7**, 3–6.

HOWES, D. H. On the interpretation of word frequency as a variable affecting speed of recognition. *Journal of Experimental Psychology*, 1954, **48**, 106–112.

MORTON, J. A preliminary functional model for language behavior. *International Audiology*, 1964, 3, 216–225. (Reprinted in R. C. Oldfield & J. C. Marshall, Eds., *Language.* London: Penguin Books, 1968).

MORTON, J. A retest of the response bias explanation of the word frequency threshold effect. *British Journal of Mathematical and Statistical Psychology*, 1968, **21**, 21–33.

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

NAKATANI, L. H. Comments on Broadbent's response bias model for stimulus recognition. *Psychological Review*, 1970, **77**, 574–576.

NEISSER, U. An experimental distinction between perceptual process and verbal response. *Journal of Experimental Psychology*, 1954, **47**, 399–402.

ROSS, S., YARCZOWER, M., & WILLIAMS, G. M. Recognitive thresholds for words as a function of set and similarity. *American Journal of Psychology*, 1956, **69**, 82–86.

RUBENSTEIN, H., & POLLACK, I. Word predictability and intelligibility. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 147–158.

SOLOMON, R. L., & POSTMAN, L. Frequency of usage as a determinant of recognition threshold for words. *Journal of Experimental Psychology*, 1952, **43**, 195–201.

TREISMAN, M. On the word frequency effect: Comments on the papers by J. Catlin and L. H. Nakatani. *Psychological Review*, 1971, **78**, 420–425.

TULVING, E., MANDLER, G., & BAUMAL, R. Interaction of two sources of information in tachistoscopic word recognition. *Canadian Journal of Psychology*, 1964, **18**, 62–71.

WINER, B. J. *Statistical principles in experimental design*. (2nd ed.) New York: McGraw-Hill, 1971.

(Received July 12, 1973)