the mounting funds available for research in the biosciences from both government and nongovernment sources; the growing number of investigators receiving research awards, especially the many new names that appear each year; and finally, the shift in the pattern of order of support of subject categories evidenced in government-supported research as contrasted with the relative stability of the pattern in nongovernment-supported research.

References and Notes

1. S. L. Deignan and E. Miller, *Science* **115**, 321 (1952).
2. In this report the term *grant* is used to mean an amount of money approved for the support of a project for the period of 1 year and refers to both grants and contracts.

Verbal Habits and the Visual Recognition of Words

Leo Postman and Beverly Conger

Department of Psychology, University of California, Berkeley

There is considerable experimental evidence that the threshold of visual recognition for words varies with the familiarity of the stimuli. Significant correlations have been found between the recognition thresholds for words and the frequencies with which these words occur in written English (*1–4*). The question remains open, however, whether it is the frequency of past visual exposure to the stimuli *per se* or frequency of past usage of the words that is the essential variable. The following experiments are concerned with this problem.

Experiment I. Visual recognition thresholds were determined for 27 three-letter English words. The words were chosen so as to represent (i) different frequencies of usage as English words, and (ii) visual stimulus patterns of different frequencies of occurrence. Estimates of frequency of usage are provided by the Thorndike-Lorge word counts (*5*). These counts are based on large samples of written English but do not, of course, measure the actual frequency of usage in spoken English. For our purposes, they provide estimates of the frequencies with which we respond to different combinations of letters as *word units*.

A given three-letter combination may not only be a word in its own right but also form part of a variety of other words. Thus, the three-letter sequence *fin* not only is a meaningful word but also forms part of other words, such as *finger, define, finish.* Such three-letter sequences, regardless of whether or not they form meaningful English words, are designated as *trigrams.* The relative frequencies of trigrams in written English may be used to estimate the frequencies with which they function as visual stimuli in reading. Such a trigram count has been published by Pratt (*6*).

For three-letter sequences, the frequency of usage as words and the frequency of occurrence as trigrams are, to a considerable degree, independent. For a sample of 356 words that appear both in Pratt's trigram count and in the Thorndike-Lorge word counts, the correlation between the two measures is .30. For our experimental sample of 27 words, the correlation is .20.

The stimulus words and their frequency values as words and as trigrams appear in Table 1. The word counts refer to frequencies of occurrence in samples of 4½ million words; the trigram count is based on relative frequencies in a sample of 20,000 words. Since the ranges of frequencies are wide, it is appropriate to scale them logarithmically (*7*). When the two frequency scales are divided into high, medium, and low values, the stimulus words sample all possible combinations of values of the two variables.

The words were presented for recognition by means of a slide projector. The order of presentation was random. The speed of exposure, controlled by a photographic shutter, was held constant at 0.01 sec. Variations in the brightness of the flash were used to determine the threshold of recognition. Starting at a fixed low intensity, the flash intensity was increased in 1-v steps on each successive trial. Fourteen such exposures were given for each word; this was a sufficient number of exposures to insure recognition by virtually all subjects. The number of exposures required for recognition was used as the measure of the threshold. Twenty-one students at the University of California served as subjects.

To make the recognition scores of different subjects comparable, all threshold measures were converted into standard scores. Scatter plots of the thresholds as a function of the two frequency variables are presented in Fig. 1. The higher the frequency of word usage, the lower the recognition threshold tends to be. The product-moment correlation is −.39, which is significantly different from zero at the 5-percent level of confidence.

As Fig. 1 shows, there is no relationship between recognition thresholds and the trigram frequency of the words. The product-moment correlation is .09, which is not significantly different from zero, and in

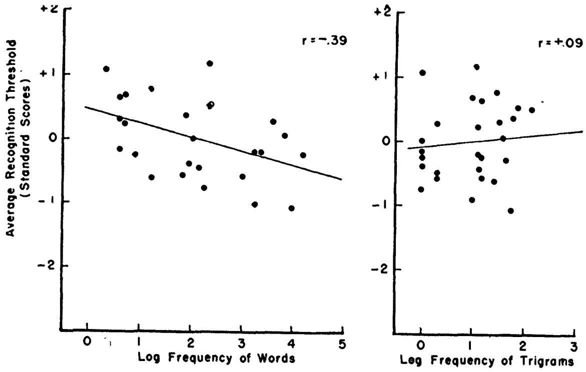

Fig. 1. Scatter diagrams showing recognition thresholds as a function of log frequency of word usage and log frequency of trigrams.

the wrong direction at that. The picture does not change when partial correlations are computed. The correlation between frequency of word usage and recognition thresholds, holding trigram frequency constant, is raised from −.39 to −.43, a value significant at the 3-percent level of confidence. When the frequency of word usage is held constant, the correlation between trigram frequency and recognition thresholds is .19, which is not significantly different from zero.

Our results point to the importance of *verbal habits* as determiners of the speed of recognition of letter sequences. When a word is presented rapidly and at low illumination, only a part of the stimulus pattern is likely to be discriminated. On the basis of such incomplete cues, the subject may then attempt to reconstruct the stimulus word. The stimulus fragments that have been discriminated may form part of several different words. Which of these words will be elicited by a given stimulus fragment depends, we assume, on the relative frequencies with which the alternative verbal responses have been given in the past. If the stimulus is a high-frequency word, a correct response in the presence of incomplete cues is highly probable.

This analysis is supported by an examination of the subjects' errors prior to correct recognition. Figure 2 compares the frequencies of usage of incorrect response words (*8*) with those of the stimulus words. The log frequency of each incorrect response word was expressed as a deviation from the log frequency of the stimulus word. The stimulus words were then divided into three groups, namely, words of relatively high, medium, and low frequency of usage. For each of these groups, the average discrepancy between the log frequency of the incorrect response words and the log frequency of the stimulus words is plotted in Fig. 2. In the case of low-frequency stimulus words, the subjects respond with words that are more frequent than the stimuli. The same relationship exists, but to a lesser degree, for the stimulus words of medium frequency. Incorrect responses to high-frequency words are actually lower in frequency of usage than the stimulus words. The general picture presented by Fig. 2 is the familiar one of regression toward the mean. We are, however, not dealing with a mere statistical artifact. As Fig. 2 shows, the positive discrepancy between response words and low-frequency stimulus words is considerably greater than the negative discrepancy between response words and high-frequency stimulus words. This fact, in conjunction with the positive discrepancy for medium-frequency stimulus words, supports the view that subjects respond with relatively frequent words in the presence of incomplete stimulus cues.

The results of this experiment agree with the hypothesis that frequency of response to stimulus units, rather than sheer frequency of visual exposure, is a significant determinant of speed of recognition. Since all the stimuli were words, the situation was maximally favorable to the determination of recognition responses by verbal habits. The effects of relative exposure frequency may, therefore, have been masked by the verbal response habits. It seemed advisable to repeat the experiment, using three-letter nonsense syllables that varied in frequency of occurrence as trigrams in written English. Such a series of stimuli should provide maximum opportunity for the effects of exposure frequency to manifest themselves independently of verbal habits.

TABLE 1. Stimuli used in the experiments.

Experiment I			Experiment II	
Word	Log frequency of word usage*	Log frequency of trigram	Syllable	Log frequency of trigram
can	3.83	1.60	nat	1.60
end	3.22	1.64	ist	1.61
out	3.98	1.73	ain	1.78
arm	3.37	1.11	alt	1.20
him	4.18	1.18	hir	1.23
set	3.25	1.00	dem	1.00
bed	3.00	0.30	dah	0.30
job	3.00	.30	jac	.30
sir	3.59	.30	gir	.30
ate	2.34	2.14	ati	2.11
ore	1.88	1.80	ove	1.81
wit	2.37	1.88	rom	1.81
ash	2.13	1.15	ang	1.23
kin	1.82	1.18	hil	1.20
ton	2.35	1.08	tif	1.08
jam	2.03	0.0	kam	0.0
shy	1.95	.0	smo	.0
rug	2.24	.0	rup	.0
ire	0.60	1.52	ile	1.52
fin	1.18	1.45	fer	1.45
pub	1.20	1.42	pos	1.42
elf	0.70	1.11	emb	1.11
lop	.70	1.00	rov	1.00
cam	.60	1.18	ced	1.18
cob	.90	0.0	cib	0.0
ilk	.30	.0	isk	.0
pun	.60	.0	tus	.0

* Average of Lorge Magazine and Semantic counts in Thorndike and Lorge (5).

Experiment II. Recognition thresholds were determined for 27 three-letter nonsense syllables. The stimuli and their frequency values as trigrams in written English appear in Table 1. The distribution of trigram frequencies used in Experiment I was duplicated as closely as possible. The experimental procedure was in all respects the same as in Experiment I. Twenty-three students served as subjects.

The results again fail to show any relationship between recognition thresholds and trigram frequency. The correlation coefficient is .10, which is almost identical with the coefficient of .09 obtained in Experiment I (*9*). The subjects' errors prior to correct recognition also fail to show any systematic tendency to respond in terms of relatively frequent letter sequences. Figure 3 compares the frequencies of incorrect three-letter responses with those of the stimulus trigram. The frequencies of the response trigrams are in general low. Incorrect responses to high-frequency trigrams have considerably lower frequency values than the stimuli. The same relationship holds, but to a lesser degree, for stimulus trigrams of medium frequency. For low-frequency trigrams, there is virtually no difference in frequency between stimuli and incorrect responses. Comparison of Figs. 2 and 3 brings out clearly the bias toward relatively frequent responses in the case of words but not in the case of meaningless three-letter sequences.

The frequency values of the trigrams refer primarily to their occurrence as parts of longer words. Recognition was, therefore, tested under conditions different from those under which the three-letter sequences are ordinarily encountered. There is some evidence that familiar letter groupings are recognized faster than unfamiliar ones when presented as parts of long sequences of letters (*10*). In such cases, familiar word responses may generalize to nonsense sequences that approximate to various degrees the structure of English words. Our stimuli, on the other hand, were designed to test the effects of the sheer frequency of past exposure independently of associative context.

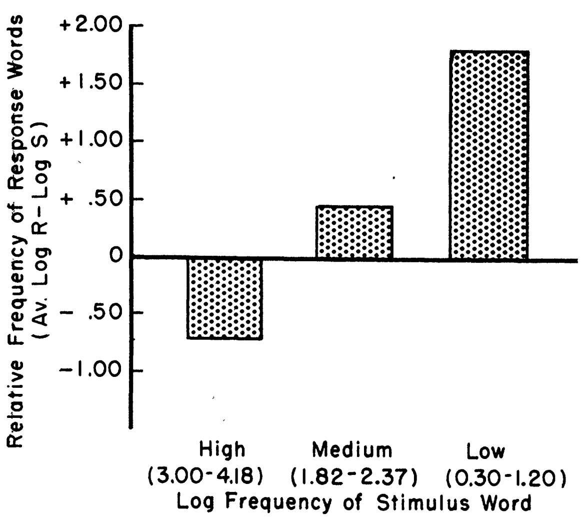

Fig. 2. Relative frequency of incorrect word responses given to stimulus words of high, medium, and low frequency.

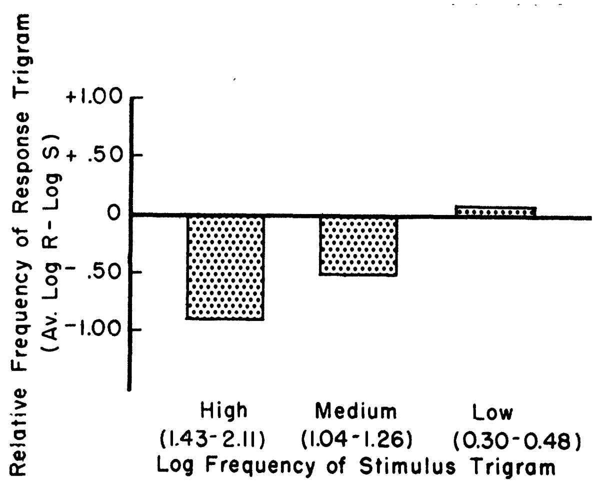

Fig. 3. Relative frequency of incorrect three-letter responses given to stimulus trigrams of high, medium, and low frequency.

We conclude that the speed of recognition for letter sequences varies significantly with the strength of the verbal habits associated with such stimuli. There are no demonstrable effects of sheer frequency of exposure.

References and Notes

1. D. H. Howes and R. L. Solomon, *J. Exptl. Psychol.* **41**, 401 (1951).
2. E. McGinnies, P. B. Comer, and O. L. Lacey, *J. Exptl. Psychol.* **44**, 65 (1952).
3. L. Postman and B. M. Schneider, *Psychol. Rev.* **58**, 271 (1951).
4. R. L. Solomon and D. H. Howes, *Psychol. Rev.* **58**, 256 (1951).
5. E. Thorndike and I. Lorge, *The Teachers' Word Book of 30,000 Words* (Bureau of Publications, Teachers' College, Columbia University, New York, 1944).
6. F. Pratt, *Secret and Urgent* (Bobbs-Merrill, Indianapolis, 1939), p. 264.
7. An inverse linear relationship between log frequency of usage and recognition thresholds was reported by Howes and Solomon (*1*).
8. Only incorrect word responses are considered here. There were, of course, other errors, including three-letter nonsense sequences.
9. The scatter plot is not presented here since it is substantially like that shown in Fig. 1.
10. G. A. Miller, J. S. Bruner, and L. Postman, *J. Gen. Psychol.* **50**, 129 (1954).