Journal of Experimental Psychology

VOL. 56, NO. 6 DECEMBER, 1958

VEXIERVERSUCH: THE LOG RELATIONSHIP BETWEEN WORD-FREQUENCY AND RECOGNITION OBTAINED IN THE ABSENCE OF STIMULUS WORDS [1]

ISRAEL GOLDIAMOND AND WILLIAM F. HAWKINS [2]

Southern Illinois University

The relationship between word-frequency and recognition has recently received considerable experimental attention. Words more frequently used in the language have produced higher intelligibility (articulation) scores than those less frequently used (**2**), and similar relations hold for other studies of recognition of spoken words (**9, 12**), as well as for visual recognition of words flashed on a screen (e.g., **10, 16**). The extent to which the frequency-recognition relationship accounts for data from both modalities is the basis for considerable controversy in perceptual defense studies (**8**). This controversy relates to the more basic question of whether the data obtained are to be given a perceptual interpretation or a response interpretation. Stated otherwise, does frequency affect seeing or saying (cf. **13**)? This argument may have relevance for the validity of sensory inferences drawn from articulation and intelligibility scores, as well as from other recognition data.

The frequency with which a word appears in the language can only be estimated; word counts from selected publications (**6, 18**) provide samples which may not be representative of all *S*s (**4, 5**). In the investigation of the frequency-recognition relationship, nonsense syllables present an advantage in that their frequency is more amenable to experimental control, since they have less of a history of prior usage than other words. In a training session, for example, one syllable may be presented 25 times, another 10 times, and so on. Each syllable is then presented in a recognition session. Such experiments report frequency-recognition relationships of the type discussed (**1, 7, 11, 14, 16**).

This study reports such a nonsense syllable experiment. Following a training session of the type described, *S* was told that one of the training words would be flashed subliminally

[1] This research was made possible through grants from the National Science Foundation, to whom the author is greatly indebted for generous support.

[2] Now at the University of Virginia.

at regular intervals on the screen. He was to respond with a word whether he saw one or not; when he was correct, E would so inform him and would initiate a new series. This experiment, however, differs in that *all flashes were blanks.* If the previously reported frequency-recognition relationship is now obtained, this relationship cannot be ascribed to perception.

Recognition involves accuracy, and accuracy is usually regarded as the congruence between S's responses and the stimulus sequence. Hence the perceptual inference. In actuality, however, E seldom co-monitors the signals with S; accuracy is usually "the congruence between S's response and E's score sheet" (**8**, p. 35).

If (*a*) words more frequently used in a training session (where the relation between reinforcement and use is constant) are also the more frequent responses in a recognition session, and if the training session words appear equally on E's score sheet, then (*b*) the frequent words will score more hits than the infrequent words. The frequent words should also produce such congruences earlier in the series. If, as in the ascending Method of Limits, a hit terminates the series, and ascending stimulus energy values are assigned to successive presentations in the series, then (*c*) recognition thresholds of more frequent words should be lower than those for less frequent words.

This study has its origins in the Vexierversuche of classical psychophysics, where blanks were introduced as a check on S's biases (**3**, p. 479). The training session in the current experiment would be considered as one of many possible ways to bias S's responses, that is, to give responses unequal likelihood of appearing in the recognition response sample. Results consistent with the three predictions made have appeared in the recognition literature. In this experiment, the predictions are made without reference to a signal, and since none is presented, perception is precluded from the interpretation. The second and third predictions relate E's *scoring procedures* and *score interpretations*, respectively, to the *response bias* of the first prediction, a function of training procedures familiar to students of learning.

Method

Subjects.—The Ss were 25 male college students, volunteers from a sophomore lecture section of a physical education course.

Stimulus words.—Stimulus words were nonsense syllables of low association value drawn from a list compiled by Hilgard (see **17**). Ten words were divided into five pairs, and systematically varied so that each pair would be presented at each of five frequencies; 13 other words, introduced to lengthen the training series, were presented once and did not appear in the recognition series. The paired words were: MIV, WUX; TUD, ZOF; HIF, MAF; FEP, HAJ; VUK, VOF. The extraneous words were: BEJ, DAJ, FEH, JEX, KEJ, LIJ, PUY, TEV, TOV, VAB, VEF, ZID, ZIM. The frequencies used were 25, 10, 5, 2, 1. This resulted in 99 cards per S.

Training procedure.—Each of the words was typewritten in ¼-in. capital letters in the center of a 3 × 5-in. white card, with the 25-frequency words appearing on 25 cards, and so on. The deck was shuffled separately for each S and was placed face down between E and S. Every 8 sec., E exposed a card in view of S, who then read the letters and the name, thus: "F-E-P, fep." The Ss were told that these were words of a foreign language, with E interested in their pronouncing each word as it appeared. These responses will be called *training responses.*

Interpolated activity.—Following the training period, S was given the January, 1958 issue of *Consumer's Report*, opened to an article on subliminal advertising, a half-page clipping from the *Southern Illinoisan*, a local daily, containing an interview with E on this subject, plus a smaller clipping from a news weekly, and was requested to read these. After 10 min., he was transferred to the perception booth.

Perception procedure.—The booth has controlled and monitored lighting, and S sits 10 ft. from a uniformly illuminated screen, in the center of which, at eye level, is an 11 × 11-in. opal glass screen for rear projection. Instead of white light, a gray mottled section from Rorschach Plate I was flashed at each presentation.[3] Interruption of the light was through a Wollensak solenoid-operated shutter connected to Hunter interval timers. Communication with E was through a high-fidelity two-way system.

The S was told that 2 sec. after the warning bell, one of the foreign words he had pronounced would be flashed into the window, so quickly that he might see only a flash; whether he saw a word or not, he was to respond with one, since the purpose of the experiment was to relate behavior to unseen or barely seen stimuli. Within three presentations, Ss were responding with a word occasioned by the window lighting up. That a presentation was made was discernible; duration was .02 sec., with interval being 7 sec. All temporal relations and bell were automatic.

Score sheet and psychophysical method.—The method mimicked was the ascending Method of Limits; S was told that the same word would be presented until he got it right, that he would be informed when correct, and that another word (or the same word, by random selection) would then be shown. This procedure was used to avert perseverations; the data obtained can be analyzed as though a random presentation method were being mimicked. The E score sheet had 10 columns, each headed by a word. These were balanced so that each word would appear equally as the first or second, the third or fourth, and so on in the 10 series. When S said the word at the top of the column, E stated: "That was the correct word. We shall now try another series," and began to make entries in the next column. All words given were recorded in the appropriate column. A pilot study having produced no "correct" responses beyond Presentation 22, if the word at the head of the column were not given by Presentation 25, E said: "We shall now try another series," and began to make entries in the next column.

It will be noted that by this procedure, all Ss did not give an equal number of responses, nor was there the same number of entries in each column, since giving the word at the head of the column terminated entries for that column, as in the studies using the ascending Method of Limits. In these studies, all entries preceding the correct one are called "prerecognition hypotheses." Responses given in the booth will be called *perceptual responses*.

Three Ss failed to produce a single correct response, and a fourth produced only one. Their records contained runs of the same syllables. Since the same learning procedure was used for all Ss who vary in learning in both training procedures and instructional periods, the performance of these Ss was considered indicative of poor learning procedures, and they were replaced by four other Ss prior to the analysis of the data.

Results

Perceptual-response frequency as a function of training-response frequency.—This relationship is presented in Fig. 1; the logarithmic relationship between the two frequencies usually found in recognition experiments of this type has been obtained here. The degree of association between

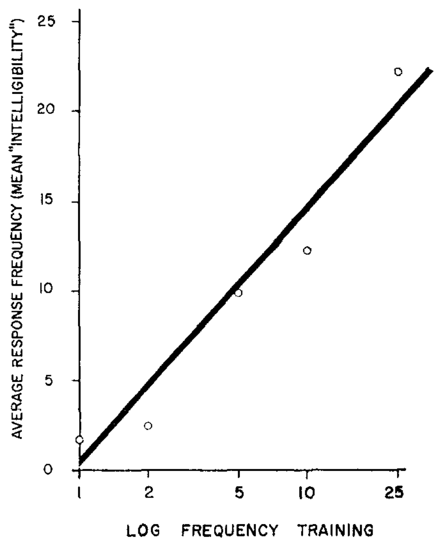

FIG. 1. Mean frequency of perceptual responses as a function of traihing frequency. If treated as ascending Method of Limits, mean frequency of "prerecognition hypotheses" as a function of training frequency. If treated as random series, mean "recognizability" ("intelligibility") as a function of training frequency

[3] This was used because it was available, and produced a random-appearing gray; the basic motivation being that "something" should be presented. This precaution turned out to be unnecessary; Ss in a later experiment reacted as well to pure white light.

obtained order and predicted order, as measured by Kendall's Tau, is 1.00, which has a significance value of $P < .01$ for five ranks. For the training frequencies of the experiment, the average perceptual response frequency is: 25, 22.18; 10, 12.30; 5, 9.88; 2, 2.38; 1, .83 (two perceptual words, 1.04; 13 extraneous words, .80).

These results confirm the prediction made and suggest that had a random method been mimicked, instead of the ascending Method of Limits, *S*'s congruences would have tended to follow the distribution in Fig. 1, and "intelligibility" and "recognition" scores would have followed suit as an artifact of this relationship. If the data are interpreted as relating to the ascending method, Fig. 1 can be considered as depicting "prerecognition hypotheses."

Number of congruences.—When *S* gave the word at the head of a column, that series was terminated, as in the ascending Method of Limits recognition studies, where such congruence defines recognition. The relationship between the percentage of words thus recognized and training frequency is indicated in Fig. 2, in which

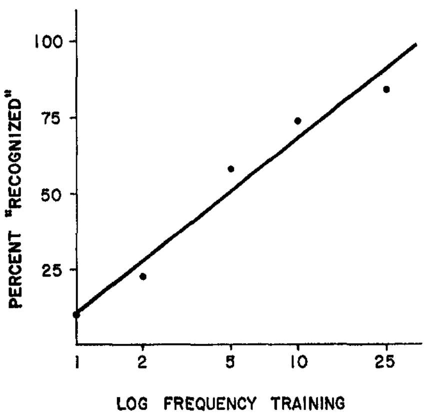

FIG. 2. Percentage of "presented" syllables "recognized" as a function of training frequency. Data treated as ascending Method of Limits.

TABLE 1

NUMBER OF RECOGNITIONS FOR EACH FREQUENCY AS A FUNCTION OF NUMBER OF RECOGNITIONS PER *S*

Words Recognized	*Ss*	Word-Frequency				
		25	10	5	2	1
7	5	10	10	8	4	3
6	5	8	10	7	4	1
5	4	7	8	3	2	0
4	5	8	6	6	0	0
3	6	9	3	5	0	1
Totals	25	42	37	29	10	5
Number unrecognized		8	13	21	40	45
Mean presentation at which recognized word was recognized		4.50	6.78	6.52	6.40	10.80

the logarithmic relationship found in Fig. 1 is continued.

No *S* recognized all words. Table 1 relates the total number of words recognized by each *S* to the number of words recognized by *S* for each frequency. If *S* recognized all 10 words, the recognitions would be evenly distributed among all five frequencies. If he recognized only one, this should tend to be a 25-training frequency word, with a curve in between gradually moving away from the 25-frequency words. The table indicates that recognitions generally distributed themselves in accordance with this expectation.

Average threshold.—With regard to the relationship between thresholds and frequency, to what extent are the results of this experiment similar to results obtained where signals are actually presented? Assuming that signals were presented here would require that stimulus energy be increased with each presentation by some physical value we shall call an erg. It will be noted from Table 1 that many words required energy levels higher than 25 ergs for recognition—they were not recognized by

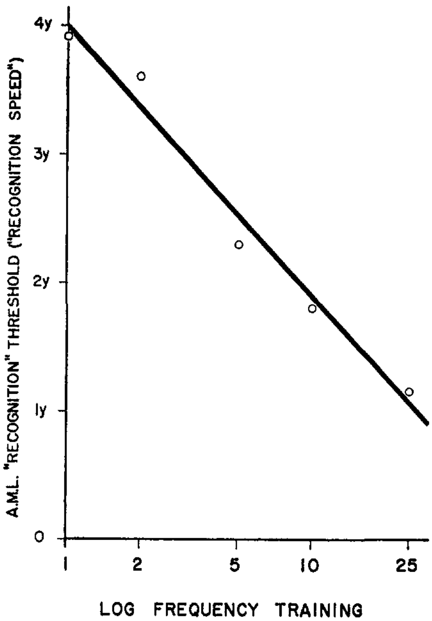

FIG. 3. "Threshold energy" required to "recognize" words as a function of training frequency. Data treated as ascending Method of Limits.

Presentation 25. If we assume that *all* words, regardless of frequency, require an average of 26 ergs for recognition, we can compute an average threshold for each word, based upon number of words unrecognized at Presentation 25 (given values of 26), plus number of words unrecognized at preceding presentations, but previously congruent through the interaction of positive response bias with score sheet entries (given values of the presentation of congruence). Such average thresholds are presented in Fig. 3. The ordinate is expressed in terms of relative stimulus energy; y can stand for ergs, milliseconds, millivolts, and the like. The inverse logarithmic relation between stimulus energy and frequency reported in the literature, and often considered indicative of perceptual effects, has been obtained.

Other results.—The syllables were selected from a list of words of equal association value; that their values were not equal for the population in this study is suggested by marked differences in the frequency of perceptual responses given. Each of the recognition words was presented an average of 8.3 times in the training session; the number of perceptual responses for each was: TUD, 394; HAJ, 382; MIV, 294; WUX, 270; VOF, 230; FEP, 200; MAF, 192; HIF, 153; VUK, 143; ZOJ, 127. The 13 nonrecognition words were presented once in the training session; their perceptual response frequencies range from TEV, 103; PUY, 42; to KEJ, ZUM, LIJ, each 0.

A total of 1259 neologisms were given by all *S*s, constituting .31 of all responses. All but 156 of these shared two letters with the training words. Consideration of these as partial responses (or partial "recognitions") did not alter the logarithmic functions when added to the data. No consistency could be found in their use. Words scoring high on partial response scored both high and low on full response; similar results obtained for other scores. The analysis was complicated by duplications (for example, T-V in TEV and TOV), for which arbitrary corrections were made.

DISCUSSION

The results of this study can be interpreted as challenging a perceptual interpretation of the relationship between word-frequency and recognition-intelligibility, where word-frequency can be placed under laboratory control. Perception was not involved in this study, yet the logarithmic recognition-frequency curves were obtained. If we pretend, as in Fig. 3, that syllables were flashed on the screen, the data make sense when we assume that all syllables are equally perceptible. Stated otherwise, we assume that frequency as a variable does not affect perceptibility. If, continuing with the signal pretense, we interpret the data to indicate that frequency does affect perception—that is, the results indicate differences in perceptibility—we find ourselves in the position of stating that the same data can be adequately explained by having frequency both affect perception and not affect it. On the other hand, stating that frequency does not affect perception, but does affect response bias, eliminates the contradiction as well as explains the data. The lowered stimulus energy at which recognition of high-frequency words occurs has been interpreted to mean that "the more frequently a word is used . . . the more readily it can be seen" (**15**). This interpretation is open to question since the inverse energy-frequency relation from which it is drawn may simply be an artifact of coupling an ascending energy series to sequential progression of a series.

The results of this study may also be relevant for interpretation of articulation tests. Using *S* as referent, congruence between responses and *E*'s score sheet is used to define sensitivity; using the word as referent, such congruence is used to define intelligibility. In the one case, inferences are often made about the person; in the other, about the perceptibility of the stimulus.

Both identification and recognition involve an accuracy indicator which may be heavily influenced by congruences related to response biases and irrelevant to the perceptual issues under study. Although this experiment deals with nonsense syllables, the frequency of which is experimentally varied, the author would consider the data as suggesting, at least, that similar considerations may hold for studies relating other types of frequency to recognition thresholds; implications of the consideration of the role of response bias are drawn for other experiments as well in a recent methodological analysis of perceptual experiments (**8**). Most recently, Howes (**9**) has interpreted parallels between auditory and visual recognition data as relating to response variables, and has assigned 69% of the variance in intelligibility data to word-frequency.

Summary

An experiment was conducted to ascertain the extent to which previously reported relations between word-frequency and recognition thresholds could be obtained in the absence of a perceptual stimulus, and could therefore be explained without resort to perception. Twenty-five *S*s were first given a training session in which they repeated nonsense syllables at different frequencies. They were then told, in a recognition session, that these words would be flashed subliminally; they were to guess until accurate. Accuracy was defined as responding with the word on *E*'s score sheet. Blanks were presented throughout.

A logarithmic relation was found between frequency of training and frequency of response in the recognition session. Since the training words appeared equally on *E*'s score sheet, the more frequent responses led to more hits, this accuracy-frequency relationship also being logarithmic. Since an ascending Method of Limits was mimicked, with accuracy terminating a series, the more frequent words also scored hits earlier, producing lower thresholds.

The similarity between these results, which cannot be ascribed to perception, and the data obtained from similar experiments where a stimulus has been presented, was interpreted as challenging a perceptual interpretation of the word-frequency-recognition relationship, where similar procedures are utilized. Implications of these results for recognition, identification, intelligibility, and articulation studies were discussed, with the suggestion offered that the existence of a contribution to the variance from response bias necessitates caution in inferring sensitivity or intelligibility (perceptibility) from such data.

References

1. Baker, K. E., & Feldman, H. Threshold-luminance for recognition in relation to frequency of prior exposure. *Amer. J. Psychol.*, 1956, **69**, 278–280.
2. Black, J. W. Accompaniments of word intelligibility. *J. Speech Dis.*, 1952, **17**, 409–418.
3. Boring, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts, 1942.

4. DASTON, P. G. Perception of idiosyncratically familiar words. *Percept. mot. Skills*, 1957, **7**, 3–6.
5. DAVIDS, A. Personality dispositions, word frequency, and word association. *J. Pers.*, 1956, **24**, 328–338.
6. DEWEY, G. *Relativ frequency of English speech sounds.* Cambridge: Harvard Univer., 1923.
7. FORREST, D. W. Auditory familiarity as a determinant of visual threshold. *Amer. J. Psychol.*, 1957, **70**, 634–636.
8. GOLDIAMOND, I. Indicators of perception: I. Subliminal perception, subception, unconscious perception: An analysis in terms of psychophysical indicator methodology. *Psychol. Bull.*, 1958, **55**, 373–411.
9. HOWES, D. On the relation between the intelligibility and frequency of occurrence of English words. *J. acoust. Soc. Amer.*, 1957, **29**, 296–305.
10. HOWES, D. H., & SOLOMON, R. L. Visual duration threshold as a function of word-probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
11. KING-ELLISON, P., & JENKINS, J. J. The durational threshold of visual recognition as a function of word-frequency. *Amer. J. Psychol.*, 1954, **67**, 700–703.
12. MASON, H. M., & GARRISON, B. K. Intelligibility of spoken messages: Liked and disliked. *J. abnorm. soc. Psychol.*, 1951, **46**, 100–103.
13. NEISSER, U. An experimental distinction between perceptual process and verbal response. *J. exp. Psychol.*, 1954, **47**, 399–402.
14. POSTMAN, L., & ROSENZWEIG, M. R. Practice and transfer in visual and auditory recognition of verbal stimuli. *Amer. J. Psychol.*, 1956, **69**, 209–226.
15. ROSENZWEIG, M. R., & POSTMAN, L. Frequency of usage and the perception of words. *Science*, 1958, **127**, 263–266.
16. SOLOMON, R. L., & POSTMAN, L. Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 1952, **43**, 195–201.
17. STEVENS, S. S. (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951.
18. THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Columbia Univer., 1944.

(Received March 14, 1958)