Storage and retrieval of low-frequency words*

CARL P. DUNCAN

Northwestern University, Evanston, Illinois 60201

Retrieval of words as a function of their language frequency was studied by having Ss attempt to recognize the words, recall the words after one presentation, or produce (think of) the words from their initial bigrams. It was found that one reason many low-frequency words could not be thought of (often necessary in anagram and other problem-solving tasks) was because they were not stored by S, as measured by failure to recognize them as words. Those low-frequency words that were stored were more difficult to retrieve than high-frequency words, both in production and in recall. High-frequency words did not exhibit failure of storage. but showed considerable difficulty in retrieval, both in recall and in production.

Studies of anagram solving have repeatedly found that solution words of low printed language frequency (from Thorndike & Lorge, 1944) are harder to solve than high-frequency words (e.g.. Mayzner & Tresselt, 1958; Dominowski & Duncan, 1964). Mayzner and Tresselt suggested that Ss use the anagram or parts of it as stimuli to think of words. They also assumed. as have Underwood and Schulz (1960) and others, that Ss emit words in the order of frequency of experience with words. Thus, low-frequency words would be thought of less readily.

Duncan (1966, 1970) suggested that the frequency effect should occur if. generalizing from Mayzner and Tresselt. Ss were given only a few letters rather than a complete anagram and asked to think of words containing those letters. In several experiments, the frequency effect was found to be very strong, probably more so than in anagram studies. However, in considering (in the 1970 study) why low-frequency words were so poorly retrieved. the question was raised as to which low-frequency words are actually stored. The distinction between what is stored vs what can then be retrieved from the store is now quite sharply drawn in memory research (e.g.. Norman, 1970). In experiments on memory, retrieval is usually from a list recently provided to S by E. But in studies of anagram solving or thinking of a word, a recent store is rarely provided: S must retrieve words assumed to be in the long-term store. In the case of very low-frequency words (often used in thinking research). some may be very weakly stored (low habit strength because infrequently encountered), others may never have been encountered at all.

The present study is an attempt to measure both storage and retrieval. with special reference to low-frequency words. It seems possible that there are some low-frequency words existing in the long-term

store which for various reasons are hard to retrieve, and other words of equivalent frequency which are not in the store at all. If so, then both storage and retrieval are implicated in any research involving S's attempts to emit low-frequency words from the presumed long-term store.

METHOD

Stimuli

The basic procedure was to use different methods of retrieval on the same pool of words. The pool consisted of all five-letter words from Thorndike and Lorge (1944) beginning with the bigrams ba. be. or br. There were 104 such words (out of a total of 200 five-letter words beginning with b in Thorndike and Lorge). The pool ranged in frequency from words below 1 per million to words above 100 per million.

Retrieval Procedures

Production

The Ss were given one of the bigrams and told to write down all the five-letter English words that were not names or capitalized words beginning with that bigram. A sheet with more than enough blanks was provided. The Ss could take as much time as they wished. There were 90 Ss. 30 for each of the three bigrams.

Recognition

The Ss were given all 200 five-letter words beginning with b from Thorndike-Lorge. A random order of the 200 words was prepared so that neither alphabetical nor order-of-frequency sequences would occur. The words were printed in two columns of 25 words each on sheets. The four sheets were stapled into booklets. Four different orders of arranging the sheets into booklets were used. The Ss were instructed to put a check mark beside every word that they thought they had ever seen in print, whether or not they knew the meaning of the word. There were 90 Ss. The number of Ss checking each of the 104 words needed for the present experiment was the score.

Recall

The words beginning with one of the three bigrams were printed in three different random orders (no obvious alphabetical or frequency sequences) on memory drum tape. The same was done for the words beginning with each of the other
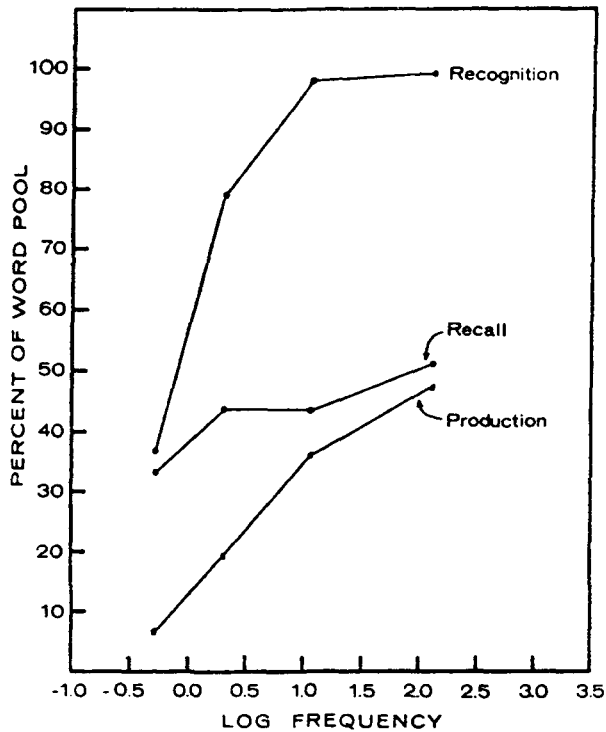
Fig. 1. Mean percent of word pool retrieved as a function of log word frequency.

two bigrams. Each of these three lists was presented to a different group of 30 Ss (10 Ss per order) for one trial, with instructions to learn. Rate of presentation was 2 sec/word. The Ss were told in advance that all of the words were five letters in length and that all began with (ba, be, or br). Immediately after presentation of the list, S was given pencil and paper and told to write down as many of the words as he could in any order.

Subjects

The 270 Ss were students from introductory psychology classes at Northwestern University. Students in these classes serve in experiments as part of the course requirement.

RESULTS

The purpose of this study is to examine retrieval as a function of word frequency. First, it is necessary to set up categories of the frequency dimension. The method decided upon is essentially the same as used previously (Duncan, 1966). The total pool of 104 words was divided into categories containing approximately equal numbers of words. Four categories were used. The frequency values per million words and the number of words in each category were: < 1 frequency (25 words), 1-4 (26), 5-29 (28), and 30-537 (25). Words in Thorndike and Lorge with A or AA frequencies were assigned a numerical frequency estimated from the Lorge count. It will be noticed that since the frequency range spanned within a category becomes larger as frequency gets higher, a nearly log scale of frequency is obtained. Thus, also following previous practice (Duncan. 1970), the results will be reported as a function of log frequency.

The results are shown in Fig. 1. The graph shows the mean percent of the words in each of the four frequency categories that were retrieved under each of the three retrieval methods. The points on each curve are plotted above the midpoint of each log category. (It may be noticed that the four points on each curve are not quite equally spaced. This results from categorizing to get categories which are as equal as possible in number of words rather than categorizing in equal-sized log steps. The basic findings would not be changed if equal log steps were used, but the number of words in the highest frequency level would be rather few. Because of the extreme J-shaped distribution that exists between number of words and the printed language frequency of the words in word pools of almost any size, it is difficult to achieve categories of equal log frequency steps that also contain equal numbers of words.)

As Fig. 1 indicates, the overall effect of the word frequency variable was such that performance was poorer on low-frequency words. The difference in retrieval between high- and low-frequency words was especially great in recognition and in production, less so in recall. At the lowest frequency level (words less than 1 per million frequency), Ss recognized, as words, less than 40% of the words, whereas at the highest level (words at or above 30 per million), recognition was essentially perfect. A great difference in retrieval as a function of frequency was also found in production, rising from well below 10% at the lowest frequency to over 40% at the highest frequency. Recall increased only from about 35% to near 50% over the whole range of frequency studied.

In two of the retrieval methods, recall and production, Ss were told to write down the words in any order. It is possible, then, to analyze order of listing as a function of word frequency. This is shown in Fig. 2. In this figure, cumulative percent of the word pool retrieved is plotted as a function of the first 20 positions of listing (thereby including most of the words retrieved). The total pool of 104 words was divided into two categories, 50 words above the median frequency (5 per million) and 54 words at or below the median. Separate curves are plotted in Fig. 2 for these two frequency categories. For both methods of retrieval, higher-frequency (above median) words were emitted in greater numbers than lower-frequency words from the earliest positions of listing, indicating operation of the Underwood-Schulz (1960) spew law. In the case of production, the tendency to emit high- rather than low-frequency words increased in later positions; low-frequency words were not thought of even after most of the high-frequency words that Ss could think of had been emitted.

Presenting the words once only on the memory drum, then asking for their free recall, had a strong effect on the retrievability, via recall, of lower-frequency words. In the earlier listing positions, these words were emitted almost as often as above-median words. Only in later

positions do the two recall curves begin to diverge in favor of high-frequency words. The closeness of the two recall curves in early positions can perhaps be accounted for by the results with production. The above-median production curve is the highest of all until Position 12, after two-thirds or more of all words that will be emitted under either method have been emitted. This very strong operation of spew was almost, though not quite, eliminated in early positions by presenting the words on the drum, then measuring recall. It appears that lower-frequency words, which were so greatly strengthened by one drum presentation, may have interfered with the recall of higher-frequency words. The result is, as shown in Fig. 1 and as Hall (1954) has also found, that high-frequency words are better recalled than low-frequency words, but the difference is not great.

Fig. 2. Cumulative percent of word pool retrieved as a function of ordinal position of listing.

DISCUSSION

The present study stemmed originally from work in problem solving. Mayzner and Tresselt (1958) had suggested that anagrams of high-frequency words are more easily solved than anagrams of low-frequency words because high-frequency words are thought of first. Underwood and Schulz (1960) called the tendency to emit high-frequency verbal units first the spew principle and extended it to verbal units other than words. Duncan (1966, 1970) then suggested that spew should operate perhaps even more clearly when Ss try to think of a word with only a few letters of the words present than when all of the letters are present (as is the case in anagram solving). Although quantitative comparisons are difficult, it did appear that the strong bias in favor of high-frequency words in anagram studies was even stronger in the thinking-of-a-word studies. In any case, if very low-frequency words were used, such as those less than 1 per million from the Thorndike-Lorge count, performance either in anagram solving or in thinking-of-a-word tasks (even when Ss were given several clues to the word) was very poor. College students could not retrieve, at least in problem-solving types of tasks, many of the extremely low-frequency words listed in Thorndike and Lorge.

These findings raised the question of whether the difficulty in retrieving low-frequency words might be in part a failure or an absence of storage. The results of the present study suggest that this is the case. The Ss indicated that they did not recognize (did not remember ever having seen), on the average, more than 60% of the words below 1 per million frequency and 20% of the words from 1 to 4 per million frequency. It seems clear that in problem-solving tasks, S would not think of a good many low-frequency solution words simply because they are not in his long-term store. Furthermore, even if S rearranged an anagram to produce other letter combinations, he would not
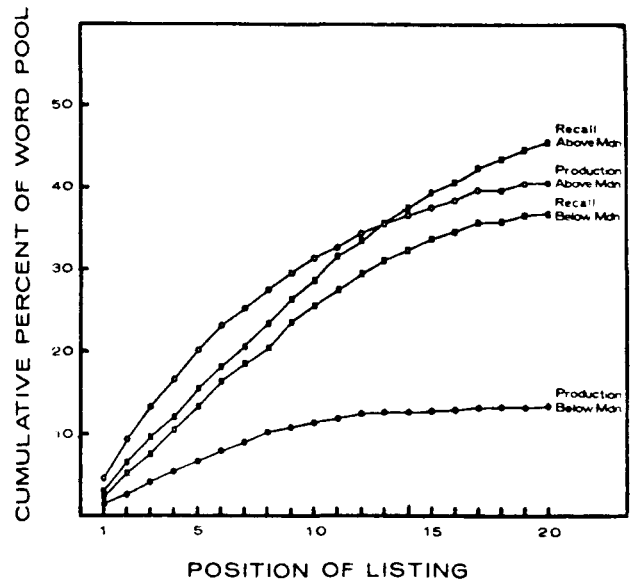
recognize some combinations that are actually words but of low frequency.

Even those low-frequency words that are stored may fail to be retrieved (thought of) in trying to solve anagrams or to think of words from clues. In Fig. 1, the curve for production shows that only about 5% of words at less than 1 per million frequency (lowest point on the curve) were thought of. But this 5% retrieval is based on all of the words at that frequency level, some stored, some not. Since S can only think of words that are stored, it is of interest to compute percentages for production based only on words stored, as indexed by the production curve. This was done by computing the percentage that the production value (from Fig. 1) is of the recognition value. At the lowest frequency level, this new percentage for production was 18%, i.e., this is the percentage of words thought of out of the pool of words actually stored. At the 1-4 per million level, the new percentage for production was 25% (and essentially the same as shown in Fig. 1 at the two highest frequency levels). Even these new production values indicate that those low-frequency words that are in the store are difficult to retrieve. Thus, many low-frequency five-letter words listed in Thorndike and Lorge have not been stored by college student Ss, and those low-frequency words that are stored have a low probability of being retrieved in thinking-type problems.

With high-frequency words, there was little or no failure of storage: almost all five-letter words above, say, 10 per million frequency were in the store. However, as the anagram and thinking-of-a-word studies found, failure to retrieve high-frequency words was not uncommon. In the present study, more than 50% of even the highest-frequency words were not thought of, nor recalled, perhaps because of interference from other words that are similar on one or more verbal dimensions.

In sum, many very low-frequency words are not in the store (not recognized). Those that are in the store but have not recently been presented are extremely poorly retrieved (production). Very high-frequency words are recognized with near certainty and, although thought of much more easily than low-frequency words, still show many failures of retrieval via production. Recent presentation of words varying in frequency greatly improves retrieval in recall of low-, but not necessarily of high-, frequency words.

REFERENCES

Dominowski, R. L., & Duncan, C. P. Anagram solving as a function of bigram frequency. Journal of Verbal Learning & Verbal Behavior, 1964, 3, 321-325.

Duncan, C. P. Effect of word frequency on thinking of a word. Journal of Verbal Learning & Verbal Behavior, 1966, 5, 434-440.

Duncan, C. P. Thinking of a word under different retrieval constraints. Journal of Verbal Learning & Verbal Behavior, 1970, 9, 356-361.

Hall, J. F. Learning as a function of word frequency. American Journal of Psychology, 1954, 67, 138-140.

Mayzner, M. S., & Tresselt, M. E. Anagram solution times: A function of letter order and word frequency. Journal of Experimental Psychology, 1958, 56, 376-379.

Norman, D. A. (Ed.) Models of human memory. New York: Academic Press, 1970.

Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

Underwood, B. J., & Schulz, R. W. Meaningfulness and verbal learning. Philadelphia: Lippincott, 1960.