Anagram solving: A function of the frequency of vowel-consonant pattern[1]

RONALD SCHWARTZ, Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital, Chicago, Ill. 60616

An estimate of the frequency of occurrence of the eight possible vowel-consonant patterns in three-letter words was derived from the Lorge magazine count. It was hypothesized that anagram solution time will be shorter when the solution word has a vowel-consonant pattern which occurs frequently, and longer for the solution word whose vowel-consonant pattern is less frequent. The hypothesis was confirmed.

Traditionally, five-letter words have been employed in the experimental investigation of anagram solving. A review of this work can be found in Johnson (1966). The problem has proved to be a difficult one. It seems likely that the variables significant in the solution of five-letter anagrams would be reduced if smaller verbal units were used. The extreme case is the two-letter word. Here, solution involves a simple rearrangement of the two letters. In preliminary exploratory studies, two-letter anagrams were used as stimuli, but it was found that the solution time was often too rapid to be measured reliably. In further exploratory work, it was found that the solution time of three-letter anagrams could be reliably measured, and a series of studies was initiated to determine the variables influencing the solution time of these smaller units.

Evidence indicates that letter sequence habits may involve the treatment of letters as members of the class of vowels and consonants (Underwood & Schulz, 1960, pp. 214-215). This suggests that prior experience with the pattern of vowels and consonants in words becomes incorporated in the verbal repertoire. It follows that, in anagram solving, the particular vowel-consonant (V-C) pattern of the solution word might affect solution time. For three-letter words there are eight possible V-C patterns (CCC, CCV, CVV, CVC, VCC, VCV, VVC, VVV). An estimate has been made, from printed text, of the frequency of occurrence of each of these patterns, and the present study concerns solution time as a function of frequency of these V-C patterns. Specifically, the hypothesis is that anagram solution time will be shorter when the solution word has a V-C pattern which occurs frequently, than for the solution word whose V-C pattern is less frequent.

METHOD

Subjects

Ten male volunteers, students at a nearby college, served as Ss. The experiment was conducted individually.

Stimulus Materials

Each anagram was presented on a white metal board. The letters were black plastic 1½ in. in height. Each letter contained a small magnet by means of which it was held to the board. The three letters occupied a distance of 3-5/8 in. and were evenly spaced.

The experiment was run twice. For the first run, eight anagrams were made from words taken from a list of the 500 most frequently occurring words in English (Thorndike & Lorge, 1944, pp. 267-268). For the second run, eight words were selected from a list of the 500 next most frequently occurring words (Thorndike & Lorge, 1944, pp. 268-270). In all other respects, the two runs were identical. Selection criteria also included the following: (1) V-C pattern. Four words having a CVC pattern (first run, SUN, CUT, CAN, PUT; second run, JOB, RAN, BOX, LIP) and four words having less

frequently occurring patterns (first run, SEA, OUT, DIE, AGE; second run, OIL, TIE, DUE, ICE). (2) Anagrams with single solutions only. Letters could be arranged to spell just one word. No additional solutions were found by Ss. (3) Words with a repeated letter, words having a "Y," and proper nouns were excluded.

An estimate of the frequency of occurrence of each of the eight possible V-C patterns for three-letter words was made by using the Lorge magazine count in the *Teacher's Word Book of 30,000 Words* (Thorndike & Lorge, 1944). This is a count of 4½ million words of magazine text. Every three-letter word with a frequency of at least one occurrence per million words was categorized according to its V-C pattern and its frequency was noted. The frequencies were then summed for each pattern. This frequency count is shown in Table 1. The numbers in this table indicate the frequency with which the V-C pattern occurs per 4½ million words. The frequency is approximate, however, because the very infrequently occurring words, less than once per million words of text, were not included in this count.

Note that Pattern CVC is most frequent, followed by Patterns CCV and VCC. These latter two frequencies do not properly reflect the degree of past verbal experience with these V-C patterns because of the very high frequency of a single word: for CVV, the word THE has a frequency of 236,472 per 4½ million and, for VCC, the word AND has a frequency of 138,672 per 4½ million words. If these two words are eliminated from the count, the CCV pattern frequency becomes 55,121 or 5.5% of the total and the VCC pattern frequency becomes 65,180 or 6.6%. With these two words omitted, the high frequency of the CVC pattern, which is not due to any single word or small number of very frequently occurring words, becomes more apparent. In order to avoid this problem, words having a CCV or VCC pattern were not employed as stimuli.

Design and Procedure

Each S was presented with 16 anagrams, eight in the first run and eight in the second run. Four anagrams had a solution word having a CVC pattern and four had solution words having some other V-C pattern. The order of presentation of the anagrams was random within and between Ss with the restriction that every other anagram had a solution word with a CVC pattern.

Letter order for each anagram was randomly determined with the restriction that all five possible permutations of three letters (not including the word itself) would be represented within the group of eight. Thus, three permutations had to be duplicated for each S. The determination of the duplicate permutations across Ss was also random.

In addition, each of the five-letter orders of an individual word was presented to at least one of the 10 Ss. All Ss were

Table 1
Frequency of Occurrence of Vowel-Consonant (v-c) Patterns

V-C Pattern	Frequency	% of Total
CCC	0	0
CVC	350,132	35.0
CVV	29,319	3.0
VCV	46,989	4.7
VVC	29,156	2.9
VCC	203,852	20.5
CCV	291,593	29.4
VVV	42,588	4.3

Note: Numbers were derived from the Lorge magazine count.

Table 2
Medians and Ranges of Vowel-Consonant Pattern Solution Times*

V-C Pattern	First Run		Second Run	
	Median	Range	Median	Range
CVC	2.0	8.8	1.8	6.3
CVV	8.7	58.6	3.3	9.1
VVC	1.9	12.3	3.6	20.5
VCV	3.5	17.7	2.9	14.8

presented an anagram of, for example, SUN, and at least one S would be presented with a specific permutation, for example, UNS. In this way all possible letter orders of each word were represented, along with all possible letter orders for an individual S.

The S was seated so that his eyes were approximately 2 ft from the metal board which rested on a table. E arranged the letters behind a mask. The S was told that he would be shown three letters which were not in proper order and that he should "mentally" rearrange them, as quickly as possible, until they spelled a meaningful word, and then say the word aloud. He was told that the solutions were common words. The S was instructed to get ready, the mask was lifted, and the response time was recorded. If S could not solve the anagram within 60 sec the response time was recorded as 60 sec. If S gave an incorrect solution and the clock was mistakenly stopped, the response time was not included in the data.

RESULTS

Nonparametric statistics were used because the data were highly skewed and truncated at 60 sec. Table 2 shows the medians and ranges of solution times for each of the four V-C patterns. It can be seen that CVC, which is the most frequently occurring pattern, has the shortest solution time in the second run and the second shortest time in the first run. The very long median solution time for pattern CVV in the first run is largely due to the word SEA, which has unusual properties as discussed below. Only three of all the stimuli presented in both runs could not be solved within 1 min. They were all anagrams of SEA. If the solution times for SEA are omitted, the median CVV solution time in the first run becomes 4.12 which is close to the CVV solution time in the second run.

The data of both runs were subjected to the Mann-Whitney U test for large samples (n > 20; Siegel, 1956, pp. 120-121). Solution times for CVC words were compared with the combined solution times for all of the other V-C patterns. It was found that CVC solution time was significantly shorter (first run, Z = 3.09, p < .002; second run, Z = 3.92, p < .002). Individual comparisons were then made between CVC solution times and the solution times for each of the other V-C patterns. In the first run, all were found to be significant except CVC vs VVC. The other comparisons showed CVC to have a significantly shorter solution time (p < .002 in both cases). In the second run, CVC was shown to have a significantly shorter time than each of the other V-C patterns (p < .002 for each comparison).

Comparisons were also made between all possible pairs of V-C patterns not involving CVC using the Mann-Whitney U test for (9 ⩽ n ⩽ 20; Siegel, 1965, pp. 119-120) and no differences were found to be significant in either run. With the exception of CVC vs VVC in the first run, CVC has proved to have a significantly shorter solution time, while all other V-C pattern comparisons were not significantly different.

Comparisons were also made across runs to test the effects of word differences, since the two runs differed only with respect to the frequency of the words used. The CVC pattern solution time for the first run was compared with the CVC pattern solution time for the second run and no significant difference was found. The same was true for the three other

pattern comparisons. It should be noted that the solution time for the CVV pattern in the first run was shown not to be significantly longer than in the second run. This is probably due to the insensitivity of the ranking procedure where the influence of extreme cases is reduced.

Since each S received four anagrams with a CVC solution word, interspersed with four anagrams with solution words having other V-C patterns, it could be argued that the shorter CVC solution time is a practice effect. While this possibility cannot be disregarded, inspection of the data reveals no systematic reduction in CVC solution time with practice.

DISCUSSION

The present study demonstrated that the V-C pattern of three-letter solution words is a significant variable in anagram solving. This is interpreted to mean that verbal experience is relevant not only to specific letters and to letter sequence habits, i.e., bigram and trigram frequency (Mayzner & Tresselt, 1959, 1962; Dominowski & Duncan, 1964) but also to the vowel and consonant aspect of letters. When presented with an anagram having two consonants and one vowel, the S is first apt to rearrange the letters into the pattern with which he is most familiar. In line with his past experience he is likely to place the vowel in the center and try the two possible permutations by exchanging the position of the consonants. Some of the Ss reported that this was, in fact, how they solved the anagram. The less frequently occurring, and therefore less likely patterns, are tried later.

The range of solution time for the three-letter anagrams we have employed in the present study and other studies from our laboratory is approximately 1 sec to over 5 min. Since there are only five possible permutations of three letters (excluding the anagram itself) it was surprising to find solution times exceeding ½ min or so. If anagram solving is simply a matter of trying the various possible permutations, solution times should be very short for the very common words used. Since this was frequently not the case, it seems reasonable to conclude that the mere permuting of letters is not a sufficient description of the strategy employed by Ss.

When told the solution of an apparently unsolvable anagram, some of our Ss said they had no idea as to why they couldn't solve it. Others reported that they had tried the correct permutation only it didn't "seem like a word." The best example of this is the word SEA. We presented various anagrams of SEA to 36 Ss and 13 could not solve it within 1 min. Many Ss reported that they tried the SEA permutation but failed to recognize it as a word. Yet all were familiar with the word, knew its meaning, and had seen it many times. An aspect of SEA that might be relevant to its inordinate difficulty is its low graph emephoneme correspondence. This is the subject of present investigation in our laboratory.

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.

JOHNSON, D. M. Solution of anagrams. Psychological Bulletin, 1966, 66, 371-384.

MAYZNER, M. S., & TRESSELT, M. E. Anagram solution times: A function of transition probabilities. Journal of Psychology, 1959, 47, 117-125.

MAYZNER, M. S., & TRESSELT, M. E. Anagram solution times: A function of word transition probabilities. Journal of Experimental Psychology, 1962, 63, 510-513.

SIEGEL, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

THORNDIKE, E. L., & LORGE, I. The teacher's word book of 30,000 words. New York: Teachers College, Columbia University Press, 1944.

UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and verbal learning. Philadelphia: Lippincott, 1960.