THE EFFECTS OF WORD FAMILIARITY AND LETTER STRUCTURE
FAMILIARITY ON THE PERCEPTION OF WORDS

Sidney E. Owsowitz

Consultant to The RAND Corporation, Santa Monica, California
ERRATA

P-2820 - "THE EFFECTS OF WORD FAMILIARITY AND LETTER STRUCTURE
FAMILIARITY ON THE PERCEPTION OF WORDS"

Page 5 - 1st paragraph, line 13 - "between a and 10 per million)"
should be changed to: "between 1 and 10 per million)."

Page 9 - 2nd paragraph, line 2 - close parenthesis after "in the
experiment)

Page 13 - Table 2, footnote c - "For all F rates df = 1/9" should
be changed to "For all F ratios df = 1/9.

Page 14 - 1st paragraph, line 14 - "tailed test if of no logical
consequence" - should be changed to: "tailed test is of
no logical consequence".
The Effects of Word Familiarity and Letter Structure
Familiarity on the Perception of Words

Sidney E. Owowitz*
Consultant to The RAND Corporation
Santa Monica, California

Since the work of Cattell (Woodworth and Schlosberg, 1954, p. 101) late in the nineteenth century, familiarity with written material (usually defined in terms of frequency of past experience) has been demonstrated to facilitate visual perception. Cattell as well as Erdmon, and Dodge (Woodworth et al, 1954, p. 101) demonstrated that considerably more material could be read during a short interval if the material were in the form of words rather than unconnected letters (i.e. nonsense material). Postman and Conger (1954), and Howes and Solomon (1951), have both found word frequency, defined by the Thorndike and Lorge (1944) word count, to be significantly correlated (negatively) with visual threshold. It should be noted that word frequency is defined by a sampling technique, and is dependent upon written material. Discussions on this point

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This article is based on a Masters Thesis submitted to the University of Southern California in partial fulfillment of the requirements for the Master of Arts Degree in Psychology. The thesis title is The Effects of Frequency Factors on the Perception of Words. The author would like to acknowledge the valuable encouragement and advice given by Dr. M.S. Mayzner in conducting this study, and the generous advice of Dr. A. Sweetland of The RAND Corporation in preparing this document for publication.
(Gibson, Pick, Osser, and Hammond, 1962, Osgood, 1953) tend to indicate that for visual recognition this is a reasonable operational definition. Howes and Solomon (1950) have interpreted the findings of McGinnies (1949) on perceptual defense in terms of word frequency. Solomon and Postman (1952), Postman and Conger (1954), and Postman and Rosenzweig (1956) have all found built-in or preexisting familiarity with written material (non-words) to facilitate perception.

A study which bears most heavily on this investigation is that of Miller, Bruner and Postman (1954). Using eight letter stimuli which ranged from zero to fourth order approximation of English (as defined by Shannon (1951)), they demonstrated that as the approximation to English increased so too did the number of letters correctly identified during exposures of from 10 to 50 ms. Miller et al noted that if correction is made for redundancy (which increases with order of approximation of English), that the actual information processed by the S was the same for all of the orders of English approximation which they investigated.

Letter structure as a determiner of perceptibility thus seems to be a consequence of information reduction. That this is the case, seems further supported by Sperling (1960). Sperling by sampling the S's perception was able to demonstrate that the S, for a very brief period of time, had more information available than he could report under conditions of immediate recall.

Mayzner and Tresselt (1962; 1963) have recently demonstrated that not only are Ss aware of the frequency of letter combinations in English but they also seem to be aware of the position in the
words where these letter combinations occur. Anagram solution
time seems highly dependent on the position factor.

Mayzner and Tresselt (Personal Communication, 1962), have
compiled a set of tables which allow the taking into account of
position factors in the letter structure of words. In order to
measure position factors it is first necessary to take work length
into account, since the relative positions of the n digrams (pairs
of letters) depends on the total number of digrams in the word. Such
being the case, Mayzner and Tresselt first segregated their sample
of 20,000 words of English text into word length groups ranging
from three to seven letters in length (all others having been
excluded from the sample of 20,000). They tallied for each of the
n-1 position digrams (an n letter word having n-1 position digrams)
the frequency of any of the 676 (26x26) possible digrams. This was
done separately for each of the n letter words.

Two examples may help to illustrate how and what was done.

1. The word "black" is a five letter word made up of the
positioned digrams "bl", "la", "ac", and "ck" in that order. And
thus each time the word "black" occurred in the sample, digram "bl"
as the 1st position digram of five letter words received an additional
tally, "la" as the 2nd position digram of five letter words received
an additional tally, and so on for each of the four digrams.

2. If we examine the digram "ft", we find it never occurred
in the sample as either the first or second digram of a three letter
word. In four letter words it never occurred as the first or second
digram, but it did occur 22 times as the third digram. For five
letter words it never occurred as the first digram, but did occur 54 times as the second digram, 7 times as the third digram, and 3 times as the fourth digram. And so on for six and seven letter words.

The absolute values of digrams are not directly comparable between the words of different lengths since the portion of the samples which constituted each of the five different word lengths vary, as follows: 6,807 three letter words, 5,456 four letter words, 3,422 five letter words, 2,264 six letter words, and 2,051 seven letter words.

According to the findings of Miller et al (1954), that order of approximation to English is positively related to perception of written material, and the findings of others (Howes and Solomon, 1959, and Postman and Conger, 1954) that word familiarity or built in familiarity for nonsense material (Solomon and Postman, 1952), facilitates perception, should familiar words with unfamiliar letter structure, or unfamiliar words with familiar letter structure have lower visual thresholds? Word familiarity indicates the former, while measures analogous to order of approximation to English indicate the latter.

If meaningfulness alters the picture such that order of approximation to English becomes less important when real words are involved, a ranking within a given level of word familiarity according to letter structure familiarity might occur.

This experiment will test the hypothesis that letter structure familiarity (as defined by position digram values) facilitates
the visual perception of words. It will also investigate the inter-
relationship between word familiarity (defined by the Thorndike-
Lorge word count), letter structure familiarity, and letter case,
in their effect on visual threshold.

Somewhat arbitrarily, five letter words were selected to be
studied (they are approximately the average length of English words
in text). The individual and total Position Digram values for all
of the five letter words in Thorndike and Lorge's *Teacher's Word
Book of 30,000 Words* (1944) were determined. All (2,543) of the
five letter words (excepting those which were solely proper nouns)
were culled out along with their "G" (General) value which is the
measure of frequency (several measures being given by Thorndike and
Lorge, and the "G" value was selected as being the most appropriate
for this study). The words were then assigned appropriate position
digram values.* High frequency words (Thorndike - Lorge values (TL)
of between 50 and 100 occurrences per million) and low frequency
words (TL values of between 1 and 10 per million) having uniformly
high or low Position Digram values (PD) were sorted out of the list.
Within each of the four categories of words four words were selected.
Similarity of letter structure in terms of initial and final letters,
of digrams and of position digrams were kept to a minimum. Table 1
contains the set of words selected and their Thorndike - Lorge (word
frequency), and Position Digram (letter structure) values.

The sixteen words in Table 1 were evaluated for letter legibility.
That is, the degree to which the different words were constructed
of letters which differed in their separate legibility. The data

---

*The appendix contains the 2,543 five letter words, along with
these and other measures.
Table 1

The 16 Stimulus Words and Their Position Digram and Thorndike-Lorge Values

<table>
<thead>
<tr>
<th>Word</th>
<th>TL</th>
<th>PD1</th>
<th>PD2</th>
<th>PD3</th>
<th>PD4</th>
<th>PDT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beach</td>
<td>A</td>
<td>41</td>
<td>134</td>
<td>51</td>
<td>126</td>
<td>352</td>
</tr>
<tr>
<td>Cheer</td>
<td>A</td>
<td>40</td>
<td>263</td>
<td>43</td>
<td>253</td>
<td>599</td>
</tr>
<tr>
<td>Trace</td>
<td>A</td>
<td>45</td>
<td>50</td>
<td>51</td>
<td>76</td>
<td>222</td>
</tr>
<tr>
<td>Wound</td>
<td>A</td>
<td>163</td>
<td>222</td>
<td>59</td>
<td>51</td>
<td>495</td>
</tr>
<tr>
<td>Clout</td>
<td>1</td>
<td>43</td>
<td>67</td>
<td>107</td>
<td>92</td>
<td>309</td>
</tr>
<tr>
<td>Corse</td>
<td>1</td>
<td>72</td>
<td>85</td>
<td>77</td>
<td>149</td>
<td>383</td>
</tr>
<tr>
<td>Sheen</td>
<td>2</td>
<td>62</td>
<td>263</td>
<td>43</td>
<td>75</td>
<td>443</td>
</tr>
<tr>
<td>Thong</td>
<td>4</td>
<td>302</td>
<td>78</td>
<td>66</td>
<td>146</td>
<td>592</td>
</tr>
<tr>
<td>Elect</td>
<td>A</td>
<td>1</td>
<td>12</td>
<td>8</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>Knife</td>
<td>A</td>
<td>13</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Occur</td>
<td>A</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Pupil</td>
<td>A</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>Dusky</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>Adobe</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>Snipe</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Curio</td>
<td>1</td>
<td>2</td>
<td>13</td>
<td>2</td>
<td>3</td>
<td>20</td>
</tr>
</tbody>
</table>

TL -- Thorndike-Lorge value for the word. The letter "A" refers to those words having a frequency of between 50 and 100 occurrences per million words of text, a number indicates actual frequency per million words.

PD -- the Position Digram value (frequency) for the digram indicated by the number.

PDT -- position digram values totaled.

of Burtt and Bosch (1923) on Bodoni type was used, since no data was found for the specific type actually used in the study. The evaluation was made for both upper and lower case type. As some of the 16 words contained letters which were not measured for threshold by Burtt and Bosch, the average of all threshold rated letters for a given word was
taken as the words "letter legibility score." The correlation between "letter legibility score" for upper and lower case type was found to be +.73 (rho), which is significant at beyond the .05 level. Analysis of variances were calculated where the variables were Position Digram, and Thorndike and Lorge frequencies. Both analyses (upper and lower case type) were found to be not significant at the .05 level. We thus can ignore the effects of individual letters on the visual threshold of the words selected.

Method

Subjects

Twenty Ss were used in the study. All of the Ss were students at the University of Southern California, and were paid two dollars for participating in the experiment.

Method of Stimulus Presentation

Stimulus slides were projected with a Bell and Howell Explorer Automatic 754 Slide projector. In order to achieve maximum control of the stimulus intensity, which was the variable used to determine threshold, two mechanisms were used. Two pieces of polarized glass were adjusted at an angle (reducing transmitted light) so that the diaphragm of an Alphax shutter could be used for fine control of stimulus intensity.

The stimulus image was projected through the polarized filters and the Alphax shutter (the shutter was set at a calibrated speed of .023 sec.).

The image was projected onto the center of a piece of semi-
opaque (white) glass, in front of which the S was seated at a distance of about four feet.

The S controlled the presentation of each stimulus with a long plunger connected to the shutter. The S could press the plunger at any time after the E had changed to the next slide and so informed the S.

Stimulus Materials

An IBM electric typewriter with a carbon ribbon and standard elite type was used in making the materials photographed. The materials were on 36 mm slides especially prepared so that the projected image was of white letters on a black background. Thus the pre and post exposure conditions are approximately identical with the background during stimulus presentation.

Three sets of stimulus materials were used. One set was the digits "0" through "9" which was used as a familiarization task. The other two sets of slides were the experimental materials, the sixteen previously mentioned words, one set in upper case type and the other in lower case type.

Ambient and Stimulus Lighting

The ambient light conditions were constant, the lighting being more than sufficient for reading and writing. The light was supplied by one fluorescent tube in the room. The room was 8 x 18 feet in size. A small amount of light came through a transom partially opened to a constantly lighted hallway.

All stimulus materials were presented at ten intensities of
brightness. Intensity was controlled by varying the opening of
the diaphragm of the shutter. Each step of setting representing
an equal angular adjustment of the shutter's diaphragm control. Since
a brightness meter was not available, no measure of brightness or
brightness contrast was made. However, the brightness scale was most
likely a positively accelerating one and certainly not linear.

The projector bulb was changed several times, well prior to
rated life expectancy.

The ten settings of intensity were predetermined (using
persons not serving in the experiment to range from well below
to well above threshold. No experimental $S$ identified any of the
words on the first settings, and no $S$ failed to see fewer than 15
of the 16 stimulus words by the 10th setting.

Order of Stimulus Presentation and Subject Assignment to Conditions

To control for order effects, a balanced random order of
presentation was used for each of the ten presentations of the
sets of stimuli.

The subjects were assigned in a balanced random order to one
of two groups -- ten subjects were shown the upper case type and
ten the lower.

Conduct of Measurement

The $S$ was assured as to the non-traumatic and non-deceptive
nature of the study and read instructions for the familiarization
task.

"You will be shown the digits 0 through 9 in a random order
on the screen in front of you. The digits will always appear in
the center of the area marked off with tape. They will appear as
white digits about 1 inch high on a black field. After going through
the ten digits, I will always increase the intensity of the digits.
Initially, you may find it difficult to see them, but if you think
you have seen a digit, record what you think. They will only be
flashed for a brief instant so you will have to look carefully.
Eventually a setting will be reached where you should be able to
see all of the digits quite clearly.

"You will be able to control the presentation of each digit by
pushing the plunger of the cable you see. When I have selected the
slide for each of your presentations, I will say 'ready', and you
may push the plunger when you are ready.

"Before you push the plunger, sit with your back against the
back of the chair and focus on the center of the screen.

"After each presentation, use the recording device I have
provided you. Either record the digit in the square or put a check
by the side of the square if you have seen nothing. Then slide the
cover over what you have just recorded and prepare for the next slide.

"When you are finished recording, push the recording device to
the side of the table. You may keep it off to the side at all times
or move it for recording purposes as you wish.

"Have you any questions?"

The § was then run on the digits, one presentation of each of
the ten digits at setting one, then the ten digits at setting two and
so on through the ten intensity settings.
Upon the completion of the familiarization task (which took 15 to 20 minutes) the S was allowed to rest for five minutes. He was then read the instructions for the presentation of words.

"This will be basically the same task as the one you have just finished.

"Instead of 10 digits you will see 16 five letter words, in upper case (lower case) type.

"You are to record the letters in the five boxes provided. If you are not sure of the word but think you have identified some of the letters, enter those letters in the corresponding boxes. That is, if you have seen some of the initial and/or final letters, and/or middle letters record them in the appropriate boxes. If you see nothing, place a check by the side and prepare for the next slide. Record whatever you see.

"Have you any questions?"

The 16 words were then presented at each of the ten intensity settings.

The Ss took between 40 and 50 minutes to complete this portion of the session.

Results

Recognition Threshold

The number of errors prior to first correct identification of each word within a class of words were totaled and used as the main dependent variable. The procedure here being followed is basically the same as that of Postman and Rosenzweig (1956).
Analysis of Variance of the Main Hypothesis

The analysis of variance (Treatment X Treatment X Subjects) for both upper and lower case type conditions yielded significant interactions between word familiarity (word frequency) and letter structure familiarity (position digram frequency). For the upper case type group $F = 5.995, df = 1/9$. For the lower case type group $F = 5.831, df = 1/9$. Both of these are significant at beyond the .05 level.

Since for both letter case groups the interactions proved significant, main effects could not be directly evaluated. Thus simple effects were determined following the procedure of Lindquist (1953). Table 2 summarizes the simple effects.

It seems reasonable to establish a relatively stringent significance level for this rather crude data, .01 would not seem unduly high. Since in the test of the simple effects only one simple effect would be significant at the .05 level and all others are either less significant or significant at beyond the .005 level, we can conveniently consider only the latter as significant. In all instances where significant, the word familiarity effect was as predicted. On the other hand the letter structure effect was the opposite of the predicted direction, unfamiliar letter structure seeming to facilitate perception (except in the one case significant at the .05 level which was in the predicted direction).

Digrams Generated Prior to Correctly Identifying Words

The predicted effect of letter structure familiarity (position
Table 2

Summary of Simple Effects for Both Upper and Lower Case Type Groups

<table>
<thead>
<tr>
<th>Level and Variable Held Constant</th>
<th>Effect Analyzed</th>
<th>Sums of Squares</th>
<th>(Effect Analyzed)</th>
<th>(Subject X Effect Analyzed)</th>
<th>F</th>
<th>Significancea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower Case Type</td>
<td>Upper Case Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Thorndike-Lorge</td>
<td>Position Digram</td>
<td>33.80</td>
<td>18.05</td>
<td>16.20</td>
<td>23.45</td>
<td>18.778</td>
</tr>
<tr>
<td>Low Thorndike-Lorge</td>
<td>Position Digram</td>
<td>174.05</td>
<td>5.00</td>
<td>47.45</td>
<td>25.00</td>
<td>33.014</td>
</tr>
<tr>
<td>High Position Digram</td>
<td>Thorndike-Lorge</td>
<td>80.00</td>
<td>61.25</td>
<td>16.00</td>
<td>37.25</td>
<td>44.994</td>
</tr>
<tr>
<td>Low Position Digram</td>
<td>Thorndike-Lorge</td>
<td>2.45</td>
<td>1.80</td>
<td>36.05</td>
<td>27.00</td>
<td>.612</td>
</tr>
</tbody>
</table>

aSignificant Thorndike-Lorge (word familiarity) effects showed fewer errors (lower thresholds) for high frequency Thorndike-Lorge words. Significant Position Digram (letter structure familiarity) effects showed fewer errors (lower thresholds) for low Position Digram Words.

bSignificant at .05 level with threshold for high Position Digram words lower than that for low position digram words.

cFor all F rates df = 1/9.
digram), that words with a familiar letter structure would have lower thresholds than those with unfamiliar letter structure was primarily based on the findings of Miller et al. As was stated above they found that with higher orders of approximation to English, Ss correctly identified more letters when tachistoscopically presented. Since the position digram was the measure of English approximation used in this study it was decided to evaluate their effect in terms similar to those of Miller et al.

The proportion of correct (and in their correct position) digrams to incorrect digrams generated prior to the S's first correct identification of the word was chosen as the measure. The total number was not used as it might be correlated with the actual threshold measure for the word. The Wilcoxin Matched Ranked Sign Test (Siegel, 1956) was used. One tailed tests were conducted to test the hypothesis that words having a familiar letter structure (high position digrams) would produce a higher proportion of correct (and in the correct position) digrams to total, than words having unfamiliar letter structure. The one tailed test was also applied to the variable of word familiarity (Thorndike-Lorge word count), although no direction was predicted. Since the one tailed test for word familiarity did not prove significant in either direction, this failure to use a two tailed test if of no logical consequence.

Table 3 contains a summary of the results of these tests. Significance levels are from Dixon and Massy (1957), since Siegel (1956) has no values for one tailed tests below the .025 level.

In all instances words with familiar letter structure (high
Table 3

Wilcoxin Match Ranked Sign Tests -- Proportion of Correct to Incorrect Digrams Generated Prior to First Correct Response to Words

<table>
<thead>
<tr>
<th>Effects</th>
<th>Level and Variable Held Constant</th>
<th>Effect Analyzed</th>
<th>Lower Case Type</th>
<th>Upper Case Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>None</td>
<td>Thorndike-Lorge</td>
<td>22 10 Not sig.</td>
<td>18 9&lt;sup&gt;b&lt;/sup&gt; Not sig.</td>
</tr>
<tr>
<td>Main</td>
<td>None</td>
<td>Position Digram</td>
<td>3 10 .005</td>
<td>0 10 .005</td>
</tr>
<tr>
<td>Simple</td>
<td>High Position Digram</td>
<td>Thorndike-Lorge</td>
<td>22 10 Not sig.</td>
<td>24 10 Not sig.</td>
</tr>
<tr>
<td>Simple</td>
<td>Low Position Digram</td>
<td>Thorndike-Lorge</td>
<td>23 10 Not sig.</td>
<td>17 9&lt;sup&gt;b&lt;/sup&gt; Not sig.</td>
</tr>
<tr>
<td>Simple</td>
<td>High Thorndike-Lorge</td>
<td>Position Digram</td>
<td>9 10 .032</td>
<td>2 9&lt;sup&gt;b&lt;/sup&gt; .006</td>
</tr>
<tr>
<td>Simple</td>
<td>Low Thorndike-Lorge</td>
<td>Position Digram</td>
<td>5 10 .010</td>
<td>0 10 .005</td>
</tr>
</tbody>
</table>

<sup>a</sup>T = smaller of the like-signed ranks.

<sup>b</sup>Though N = 10 for all conditions, three measures had instances of being tied proportions for the effects analyzed.

<sup>c</sup>Significant differences showed higher proportions of correct to total digrams responded to high Position Digram words than low Position Digram words.

Position digrams) show a significantly greater ratio of correct to total position digram responses, while word familiarity (Thorndike-Lorge count) shows no significant effect. The apparent interaction between word familiarity, letter structure familiarity, and letter case found for the main criterion makes it clear that these results are not an artifact of the main criterion.

Discussion

It is apparent that the initial hypothesis, that letter structure
familiarity facilitates the perception of words, is not substantiated, and indeed the reverse is in part indicated.

Examining the simple effects, we find for lower case type, that words with unfamiliar letter structure have lower thresholds both for familiar and unfamiliar words. It should be further noted that for the lower case type the unfamiliar words with unfamiliar letter structure had lower thresholds than familiar words with familiar letter structure. Letter structure here seems more important for threshold than word familiarity.

For upper case type where the geometric pattern of the letters is possibly less distinct, there is no significant difference in the visual thresholds of words with familiar and those with unfamiliar letter structure (unless we are willing to reject the null hypothesis at the .05 level), at either of the two levels of word familiarity.

When we examine the effects of word familiarity, we find the same effects for both upper and lower case type. Familiar words have lower thresholds than unfamiliar words when the letter structure is familiar, but thresholds are not significantly different when the letter structure is unfamiliar.

Table 3 indicates that familiar letter structure facilitates correct digram identification. The lower information content of the string of letters (Miller et al, 1954) and thus its facilitation of short term memory (Sperling, 1960) seem to explain why this happens. But the previous studies have dealt with non-meaningful material, this study deals with real words. The Ss in this study were aware that the letters they saw were parts of words, hence in addition
to recognizing the letters they attempted to tie them together (along with other letters not yet identified) into words.

The early identification of some letters may hamper identifying words. Since there is a finite set of letter patterns and some are more likely than others, both in terms of letter groupings and in terms of the words they form, the viewer is in a position to make an educated guess. If the viewer sees "th_", "the" is easily deduced, while "tho" is a much less likely response. With longer words less frequently encountered the task becomes more difficult. When the letter pattern identified is familiar, the viewer is in a position to start searching his response repertoire. The more familiar the pattern the more likely it is one of a finite set of familiar words having that basic pattern of letters. But if the word is, in fact, not a familiar one, the viewer will (in the situation of this experiment) start making guesses which on successive presentations seem to be correct, but are off by perhaps only one or two letters from the correct response. He is thus in a position to experience reinforcement for incorrect responses.

A second process may also occur with words having familiar letter structure. The viewer may not come up with a word as a response hypothesis but may simply generate some of the letters of the word which to him, although more or less confirmable, simply have one or two letters which don't seem to fit.

If the letter pattern is unusual, early hypothesis are not so easy for the viewer to come by. As has been observed (Miller et al, 1954), less familiar groups of letters are less readily responded
to correctly in the tachistoscopic presentation. Early hypothesis are easily proved false and the observer is not likely to be greatly influenced by them. It is thus not surprising that the one consistent finding (for Ss in both the upper and lower case type groups) for word thresholds, was that word familiarity facilitated perception amongst the words having a familiar letter structure, and had no significant effect upon those words having unfamiliar letter structure. Since more hypothesis testing and confirmation occurs for words with familiar letter structure, the correct identification is more likely for those in that group which are high in the response hierarchy than for those which are low. For words with unfamiliar letter structure much less of this guessing can occur. This is shown quite clearly by Table 4. Here we find that the generation of a familiar word as a response is considerably more likely for those words with familiar letter structure than those with unfamiliar letter structure.

The effect of letter structure familiarity on the perception of words is the same as the information (English approximation) effect found by Miller et al (1954), but as the S reaches the point where he is attempting to connect the letters into a word, the letter structure effect can, by its facilitation of perception of the separate letters, delay the identification of the word, and this particularly being the case if the word is unfamiliar. When the Position Digram structure of the word was very low, it was found (for the words tested) that the S's familiarity with the word had no significant effect.
Table 4

Generation of Incorrect 5-Letter Responses --
Including Incorrect Wordsa

<table>
<thead>
<tr>
<th>Stimulus Category</th>
<th>High Frequency Words</th>
<th>Total 5-Letter Responses</th>
<th>Proportion of High Frequency Words to Total 5-Letter Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Position Digram High Thorndike-Lorge</td>
<td>27</td>
<td>52</td>
<td>.52</td>
</tr>
<tr>
<td>High Position Digram Low Thorndike-Lorge</td>
<td>29</td>
<td>78</td>
<td>.37</td>
</tr>
<tr>
<td>Low Position Digram High Thorndike-Lorge</td>
<td>6</td>
<td>37</td>
<td>.16</td>
</tr>
<tr>
<td>Low Position Digram Low Thorndike-Lorge</td>
<td>9</td>
<td>62</td>
<td>.14</td>
</tr>
</tbody>
</table>

aPrior to first correct identification of a given word.

bHigh frequency words are those with count in excess of 50 per million in Thorndike-Lorge count.

Summary

An experiment was conducted to test the hypothesis that familiarity of letter structure (as opposed to familiarity of the word) would facilitate the perception of the word. The results showed an interaction between letter structure familiarity and word familiarity such that while letter structure familiarity facilitated correct identification of the letters of the word, letter structure familiarity resulted in inhibiting the perception of unfamiliar words. Where the letter structure was less familiar, familiar and unfamiliar words did not differ in threshold.
LIST OF REFERENCES


APPENDIX

The 2543 five letter words which were extracted from the Thorndike-Lorge List (1944) were assigned the 21 values described below:

1-5. For each of the four positioned digrams the frequency value for that digram in that position of a five letter word as found by Mayzner and Tresselt (sample of 3,422 words), and the total of the four values (PD1, PD2, PD3, PD4, and PDT).

6-10. For each of the four positioned digrams the frequency value for that digram in any position of a five letter word as found by Mayzner and Tresselt, and the total of the four values (WL1, WL2, WL3, WL4, WLT).

11-15. For each of the four positioned digrams the frequency value for that digram in any position of words of from three to seven letters in length as determined by Mayzner and Tresselt (sample of 20,000 words), and the total of the four values (TD1, TD2, TD3, TD4, TDT).

16-21. Four each of the five positioned letters the frequency value for that letter in that position of a five letter word as found by Mayzner and Tresselt, and the total of the five values (PL1, PL2, PL3, PL4, PL5, PLT).

The words, their Thorndike Lorge values, and the 21 above described values are given in this appendix. A "T-L" value of "100" in the tables refers to the "AA" Thorndike Lorge words which are those
words which occurred 100 or more times per million in the "G" (General) count of Thorndike Lorge. A "T-L" value of "50" refers to the "A" words in the "G" count which are those words which occurred 50-99 times per million words. Successive values are the actual number per million words except for those numbers prefixed by "S" and those containing only a "T". "S" indicates that the number given is the frequency in 18 million (an approximate figure is given). "T" refers to those words occurring 4 times per 18 million words.