DOCUMENT RESUME

ED 088 038 CS 000 969

AUTHOR Martin, Clessen J.; Pantalion, Charles A., Jr.
TITLE Effect of Two Word Deletion Schemes upon the Comprehension of Five Levels of Telegraphic Prose.
PUB DATE Feb 73
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, February 25-March 1, 1973)

EDRS PRICE MF-$0.75 HC-$1.50
DESCRIPTORS *College Students; Computers; *Prose; *Reading Comprehension; *Reading Research; Reading Speed; Student Participation; *Telegraphic Materials

ABSTRACT

This study attempted to increase the reading efficiency of normal readers by developing a method of producing telegraphic prose. The subjects' (undergraduate students in psychology) rank ordered the words in each sentence of a prose passage according to the importance of the words for communicating the main sentence idea. Specific treatment conditions were five subject-generated reductions and five computer-generated random reductions consisting of a 10%, 20%, 30%, 40%, and 50% deletion of words from each sentence in the passage. Variables included comprehension, reading rate, required reading time, and an efficiency measure. For any of the four variables there were no significant differences in comprehension between the traditional version and the reduced versions generated by either the random or subjective deletion schemes at the 10% deletion level. Results also showed that the subject-generated versions of 30% and 40% deletion saved a significant amount of reading time when compared to the traditional version, that there were no significant differences in comprehension between these versions and the traditional version, and that both deleted versions were more efficient than the traditional version. These results supported the feasibility of the telegraphic concept based on a subject deletion scheme. (HOD)

1973 ANNUAL MEETING

OF

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

New Orleans, Louisiana

February 25 - March 1, 1973

EFFECT OF TWO WORD DELETION SCHEMES UPON THE COMPREHENSION OF FIVE LEVELS OF TELEGRAPHIC PROSE

Presented by

Clessen J. Martin

Texas A&M University

Charles A. Pantalion, Jr.

Texas Christian University

EFFECT OF TWO WORD DELETION SCHEMES UPON THE COMPREHENSION OF FIVE LEVELS OF TELEGRAPHIC PROSE

Clessen J. Martin
Texas A&M University

Charles A. Pantalion, Jr.
Texas Christian University

Introduction

The present experiment is part of a programmatic research effort which has as its objective the development of text reduction techniques. This study was an attempt to increase the reading efficiency of normal readers by developing a method of producing telegraphic prose. The telegraphic reduction method used in this study is based on Ss rank ordering the words in each sentence of a prose passage according to the importance of the words for communicating the main sentence idea.

The information obtained from the rank ordering is used to generate telegraphic passages at increasing levels of deletion. Low levels of deletion eliminate words judged unimportant by the Ss, whereas higher levels of deletion eliminate the more important words. The subjectively deleted telegraphic prose was compared to randomly deleted telegraphic prose. The random deletion scheme is used as a base line in order to determine to what extent the subjective deletion scheme is superior to a random elimination of words.

Method

Subjects

The Ss for the study were undergraduate students in Psychology at Texas A&M University. A total of 550 Ss were tested with 50 Ss being randomly assigned to each of the eleven treatment conditions.

Design and Analysis

The basic design for the study was a 1 x 11 analysis of variance to determine if any differences existed between treatment conditions in relation to comprehension, reading rate, required reading time, and efficiency. The

specific treatment conditions were five subject generated reductions consisting of a 10% (S_1), 20% (S_2), 30% (S_3), 40% (S_4), and 50% (S_5) deletion of words from each sentence within the passage, five computer generated random reductions also consisting of a 10% (R_1), 20% (R_2), 30% (R_3), 40% (R_4), and 50% (R_5) deletion of words from each sentence within the passage, and a traditional version (T) of the passage.

The Tukey Test of Multiple Comparisons (Winer, 1962) was used to determine the specific condition or conditions responsible for the significant differences.

A 2 x 5 factorial analysis of variance was performed to determine if any interaction existed between the reduced versions of the passage and the method utilized for reduction relative to comprehension, reading rate, required reading time and efficiency.

Materials

The passages were extracted from an Air Force manual entitled History of the USAF. The manual is a reprint of the Air Training Command pamphlet 190-1. The specific passage, on pages 4-7 and 4-8 of the manual, consists of 67 sentences entitled "Hemisphere Defense." The passage relates the problems confronted by the Army Air Force prior to World War II relative to the defense of the Western Hemisphere. The passage from the manual constituted the traditional version (T).

Procedure

Subjected generated material. The construction of the five treatment conditions was based on a subjectively determined mean rank order for each word within each sentence. Subjects rank ordered each word with respect to its significance to the intended meaning of the sentence. A rank order of one denoted the least significant word, and the rank order equivalent to the total number of words within the sentence denoted the most significant word. The rank order assigned to each word within a sentence was summed across all Ss and divided by the number of Ss to obtain the mean rank order.

A computer program was developed to analyze the subject reduced data and to construct the reduced versions of the sentences that eventually comprised the treatment conditions.

A total of 220 Ss were used to rank order the words within the sentence. Twenty Ss were randomly assigned to examine one of eleven sets of sentences. Ten of the sets consisted of six sentences, and one set consisted of seven sentences. The nature of the problem and instructions for the procedure to be followed were presented to the Ss prior to their participation.

Test material. The comprehension test, was a multiple choice test consisting of 50 items with four alternatives per item. The test was administered immediately after reading the prose passage.

Results

The major analysis involved comparisions among the means for each treatment condition on the following dependent variables: comprehension, reading rate, required reading time, and an efficiency measure e_t which was defined as the number correct on the comprehension test times 100 divided by reading time. The means and standard deviations for each dependent variable are presented in Tables 1, 2, 3, 4.

Five 1 x 11 ANOVAs were used to analyze each of the dependent variables across all treatment conditions. Tukey's Method of Multiple Comparisons was used to determine the nature of the differences among the treatment conditions. These results are presented in Table 5. The .05 level of significance was used in making all comparisons.

The results of the 1 x 11 ANOVA on the comprehension variable yielded an $F=17.97$, $df=10/549$, $p<.01$. As indicated in Table 5, all reduced versions of the passage did not significantly differ from the traditional version with the exception of S_5. With the exception of treatment condition R_1, all randomly reduced versions had significantly lower comprehension means compared to the traditional version.

For the variable of reading rate the results of the 1 x 11 ANOVA were $F=12.29$, $df=10/549$, $p<.01$. Table 5 shows that versions S_1, S_2, R_1, R_2, and R_3 did not differ significantly from the traditional version with respect to reading rate.

The results of the 1 x 11 ANOVA on the variable of reading time yielded an $F=11.69$, $df=10/549$, $p<.01$. Significantly less reading time was required by those Ss reading 40 and 50% reduced passages when compared to the traditional version. Versions S_1, R_1, R_2, did not significantly differ from the

TABLE 1

Mean Comprehension Score and Standard Deviation for Each of the 11 Treatment Conditions

Treatment Conditions	Mean	SD
T	29.30	6.20
S_1	27.64	7.55
S_2	26.38	5.63
S_3	26.78	6.98
S_4	26.74	6.07
S_5	25.02	5.78
R_1	25.78	5.56
R_2	21.56	5.29
R_3	20.70	4.37
R_4	19.60	4.19
R_5	19.86	4.11

TABLE 2

Mean Reading Rate (wpm) and Standard Deviation for Each of the 11 Treatment Conditions

Treatment Conditions	Mean	SD
T	172.60	34.93
S_1	179.99	48.01
S_2	162.56	38.03
S_3	141.24	32.37
S_4	136.37	35.89
S_5	120.49	22.80
R_1	174.42	35.10
R_2	158.80	46.95
R_3	154.45	53.55
R_4	139.69	53.63
R_5	125.42	37.07

TABLE 3

Mean Required Reading Time and Standard Deviation for Each of the 11 Treatment Conditions (in seconds)

Treatment Conditions	Mean	SD
T	575.74	122.36
S_1	505.80	124.31
S_2	495.12	121.58
S_3	499.20	112.77
S_4	441.30	97.97
S_5	398.26	65.90
R_1	508.32	87.55
R_2	512.94	138.09
R_3	473.80	132.84
R_4	443.66	107.62
R_5	398.34	77.97

TABLE 4

Mean e_t and Standard Deviation for Each of the 11 Treatment Conditions

Treatment Conditions	Mean	SD
T	5.27	1.54
S_1	5.78	2.08
S_2	5.63	1.82
S_3	5.80	2.33
S_4	6.35	2.12
S_5	6.38	1.52
R_1	5.22	1.48
R_2	4.51	1.63
R_3	4.68	1.64
R_4	4.71	1.75
R_5	5.22	1.72

TABLE 5

Results of the Significant Differences Between Treatment Means for the Dependent Variables Based Upon Tukey's Method of Multiple Comparisons*

Dependent Variables	Treatment Conditions
Comprehension	$T \quad S_1 \quad S_3 \quad S_4 \quad S_2 \quad R_1 \quad S_5 \quad R_2 \quad R_3 \quad R_5 \quad R_4$
Reading Rate	$S_1 \quad R_1 \quad T \quad S_2 \quad R_2 \quad R_3 \quad S_3 \quad R_4 \quad S_4 \quad R_5 \quad S_5$
Required Reading Time	$S_5 \quad R_5 \quad S_4 \quad R_4 \quad R_3 \quad S_2 \quad S_3 \quad S_1 \quad R_1 \quad R_2 \quad T$
e_t	$S_5 \quad S_4 \quad S_3 \quad S_1 \quad S_2 \quad T \quad R_1 \quad R_5 \quad R_4 \quad R_3 \quad R_2$

*Treatment conditions arranged from highest (left) to lowest (right). Treatment conditions underlined by a common line do not differ significantly; treatment conditions not underlined by a common line do differ significantly ($p<.05$).

traditional version.

For the variable of e_t the results of the 1 x 11 ANOVA yielded an F= 6.33, df=10/549, $p<.01$. Inspection of Table 5 for the e_t variable shows a superiority of subject reduced versions with treatment condition S_5 being most efficient. The subject reduced versions did not differ significantly from the traditional version.

Each of the four dependent variables were analyzed in a 2 x 5 factorial ANOVA in which Deletion Levels and Deletion Schemes were the independent variables. Again, Tukey's Method of Multiple Comparisons (Winer, 1962) was used to determine the exact nature of the treatment differences.

Results of the analysis for comprehension scores indicated that both main effects and the interaction effect were significant. (Deletion Level F = 8.23, df=4/499, $p<.001$; Deletion Scheme F=98.10, df=1/499, $p<.001$; and interaction F=3.05, df=4/499, $p<.05$.) Regardless of the Deletion Scheme, the deletion of 10% was superior and significantly different from the remaining four levels of deletion. Generally, a relative decrease in comprehension occurred with a corresponding increase in percent deletion. Twenty and 30% deletion levels differed significantly from the 50 percent level. No significant differences were obtained between 20, 30, and 40% deletion levels or between 40 and 50% levels. The significant difference existing between the two types of deletion schemes was in favor of the subject deletion scheme. The interaction of Deletion Levels by Deletion Schemes also favored the subject generated treatment conditions.

Analysis of reading rates revealed a significant F-ratio for the Deletion Level main effect, (F=25.22, df=4/499, $p<.001$). Combining the reading rates at each Deletion Level for both Deletion Schemes indicated a significant decrease ($p<.001$) in reading rate at each successively higher level of deletion. No significant difference existed between the two types of Deletion Schemes. The interaction effect was not significant.

Analysis of required reading time indicated that the main effect of Percent Deletion was significant, (F=18.23, df=4/499, $p<.001$). Generally, reading time was less with increasing percentages of deletion. Deletion Levels of 40 and 50% were significantly different from each other and also significantly different from the remaining three levels of deletion. No

significant difference existed between 30 and 20% or between 20 and 10%; however, there was a significant difference between 30 and 10 percent. No significant difference existed for the type of Deletion Scheme, nor was there a significant interaction.

The 2 x 5 ANOVA for e_t indicated a significant Deletion Level main effect, (F=2.40, df=4/499, $p<.05$). Most efficient was 50% deletion followed by 40, 10, 30, and 20 percent. With the exception of 40% deletion, 50% deletion differed significantly from the remaining deletion levels. Fourty percent deletion was not significantly different from 10%, but was significantly different from 30 and 20%. No significant difference was obtained between 10, 30, and 20 percent. A significant F ratio for the Deletion Scheme main effect was obtained and favored the subject scheme, (F=47.12, df=1/499, $p<.001$). The interaction effect was not significant.

Discussion and Conclusions

For any of the four dependent variables, there were no significant differences between the traditional version and the reduced versions generated by either the random or subjective deletion schemes at the 10% deletion level. With the exception of the comprehension dependent variable, however, the 10% subjective deletion version was the superior of the three.

The traditional version was superior and significantly different for the variable of comprehension when compared to a random deletion level of 20% and higher. For reading rate, however, the traditional version was superior and significantly different from versions R_4 and R_5 but was not significantly different from versions R_2 and R_3.

The main effect of the deletion scheme was in favor of the subject generated scheme for three of the four dependent variables. A significant F-ratio was not obtained for reading rate and required reading time. No significant interaction effects were obtained between the two deletion schemes across all percentages of deletion for each dependent variable with the exception of comprehension. The e_t and comprehension variables indicated that the subject generated versions were superior to the random versions at each level of deletion.

The variable of required reading time indicated that highly reduced

versions regardless of deletion scheme were significantly different and superior to the traditional version. When time and comprehension were considered together in the e_t dependent variable, however, the subject generated versions were superior to the traditional version and all of the randomly deleted versions. In addition, each randomly deleted version was inferior to the traditional version.

The results of this research were interpreted as support for the feasibility of the telegraphic concept based on a subject deletion scheme. This conclusion was reached on the basis of three factors. One factor was the indication that the subject generated versions of 30 and 40% deletion saved a significant amount of reading time when compared to the traditional version. The second factor was the lack of any significant differences among versions S_3, S_4, and the traditional version on comprehension. The third factor was that both S_3 and S_4 were more efficient than the traditional version as measured by e_t.

Further research should be directed toward investigating optimum levels of deletion for a given sentence, optimum sample size for generating telegraphic versions of selected passages, and the particular grammatical categories consistently deleted by subjects. The establishment of optimum levels of deletion for each sentence within a passage should enhance comprehension and may possibly increase ease of reading. With respect to sample size, the present study utilized a relatively large number of subjects for deletion to assess the agreement among subjects relative to the rank ordering of words within sentences. However, greater practicality of a subject deletion scheme for developing telegraphic prose could be achieved if the scheme involved only one person. Moreover, the present study did not incorporate an analysis of the grammatical categories consistently deleted from sentences by subjects; however, it was noted that articles and conjunctions were generally the first words to be deleted from a sentence. Consequently, it is felt that research involving the relative importance of the various grammatical categories in relation to communicating the intended meaning of a sentence should also be conducted.

Reference

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.