Comprehension of Paragraphs Reconstructed from Scrambled Discourse Using Feedback

Philip Langer

Verne Keenan

Janice Culler

Technical Report No. 88-1, January, 1988

Abstract

Subjects reconstructed paragraphs from scrambled passages. There were two passages (240 and 400 words) and three conditions (read-only, reconstruct with no feedback, and reconstruct with feedback). The dependent measures were recall of idea units, recognition of original sentences, and concordance with the original paragraph clustering. ANOVA yielded: (1) higher recall for idea units for the shorter passage and reconstruct-no feedback condition, (2) higher recognition for original sentences for the reconstruct-no feedback condition, and (3) greater concordance for the longer passage and reconstruct-feedback condition. Generally speaking, feedback seemed to intrude into the reconstruction process, and increased concordance did not lead to improved comprehension.

This paper is one in a series of experiments concerned with the synthesis of meaning from text in an instructional context. Specifically, scrambled discourse is reconstructed with feedback, and comprehension is assessed by approximation to the original (concordance). Our research has not dealt with the mechanics of processing discourse but with the instructional problems associated with feedback and processing text. Traditionally, these areas have not been related in any systematic manner.

The extraction of meaning from language has concerned scholars in various disciplines (Britton & Black, 1985; Just & Carpenter, 1987; van Dijk & Kintsch, 1983). Textual properties that have been investigated include content and structure (Thorndyke, 1977) and referential coherence (Kintsch, 1974). We chose initially to investigate sentence order because text reconstruction from randomly ordered discourse is a process and so amenable to instructional influences. The instructional assistance provided is feedback, which is generally considered an effective instructional adjunct under most circumstances (Langer, 1983). The feedback provided was used to confirm sentence placement during reconstruction. The focus of the research lies in both the importance of passage structure as well as the effectiveness of feedback.

Sequential order of constituents is a fundamental property of any language. Words in sentences must be ordered syntactically to prevent misunderstanding. Undoubtedly the constraints of sequence are somewhat looser at the level of sentences in paragraphs or paragraphs in passages. The first structural element of discourse we investigated was at the level of sentence sequence. We consistently assumed that there are optimal sequences for any text and that comprehension and retrieval for this structure are better than for alternative sequences (Langer, Keenan, & Medosch-Schonbeck, 1986). However, we consistently found that concordance with the original sequence did not improve comprehension.

In this study we decided to investigate this issue further using a second structural element, i.e., the paragraph. Specifically, we wanted to analyze the effects of feedback on sentence clustering at the paragraph level. Structural features of language that are involved in paragraph sequence include argument overlap for coherence (Kintsch & van Dijk, 1978) and proximity of pronouns to their referents. Prose passages may vary on any of these, and the ordering of components may be much more important in some passages than others.

In addition, some writers have recognized that the problems associated with determining the effect of text organization on comprehension must include those skills and knowledge bases possessed by the reader. Frederiksen (1977) acknowledged there is a constant interplay between text and reader-processing systems in semantic memory. Similarly Kintsch and Vipond (1979) recognized that readability is a function of text characteristics as well as the knowledge structure of the reader. Meyer and Freedle (1984) also suggested that the schema available to the reader assists processing discourse. This interaction has led researchers to somewhat contrary conclusions. While Reder (1980) argues for a significant relationship between text organization and comprehension, Irwin (1982) pointed out that over-all intersentential coherence does not seem to assist memory significantly.

For purposes of instruction, feedback has been used to guide the learner along conceptual pathways deemed most critical to achievement (Langer, 1983). For discourse processing the objective is to encourage the reader to accept the logic of the discourse as initially determined by the writer. However, there are several problems associated with any simplistic acceptance of the efficacy of feedback. First, while feedback is almost always associated with improved learning (Berliner & Rosenshine, 1977), evidence from research does not support the premise that there is a simple monotonic relationship between levels of assistance and achievement (Getsie, Langer, & Glass, 1985). Second, as we have noted, models of discourse processing have remained fairly isolated from instructional concerns, including feedback.

Since our reconstruction paradigm utilizes scrambled text, the feedback provided assisted the reader in ordering of sentences based on some developing schemata of comprehension. Historically retrieval has been proved less effective under conditions of scrambled text (Kulhavy, Schmid, & Walker, 1977; Frase, 1969; Thorndyke, 1977). With the exception of a study by Schultz and DiVesta (1972) in which feedback was inadvertently provided, the readers of the scrambled prose received no assistance. In our studies we have used published prose rather than the artificial content usually provided (Bransford & Franks, 1971).

Based on previous data from our laboratory it appears that feedback generally assists recognition rather than recall, although under the reconstruction process one might presume that recall would benefit more. In some instances, when the passage was short and the content familiar, subjects simply reading the scrambled material did about as well on retrieval as subjects reconstructing the text assisted by feedback (Langer, Keenan, & Medosch-Schonbeck, 1986). If nothing else Clark's (1973) warning about treating content as a fixed variable was justified.

Our later studies however led to the premise that the successful processing of scrambled discourse may be significantly related to the episodic-semantic dimension (Tulving, 1983; Langer, Keenan, & Culler, 1987; Langer, Keenan, & Culler, 1988). It appears that within the parameters of our reconstruction paradigm, retrieval is superior for text assisted by semantically-based schemata as compared to episodic. Tulving still perceives the accessibility of episodic memory as a function of spatial-temporal references, while semantic content is more abstract. Hintzman's (1978) characterization of semantic memory as generic may not be too far-fetched.

In this study of reconstruction of scrambled discourse at the paragraph level, we utilized two passages which we classified as semantic in content. The two texts used in this investigation were taken from a study by Young (1984) who had adapted several short essays from published periodicals. Each passage contained five paragraphs and both were moderately difficult. The differences in length, topic, and structure were intended to provide generality of results (Clark, 1973). The Computer Crime passage consists of 19 sentences and is 400 words in length, focusing on the invasion of major, supposedly secure databases by a group of youngsters. The Chesapeake Bay passage is 13 sentences long and 240 words in length, discussing pollution problems and proposals for their resolution. Our assumption was that the topics (computer crime and pollution) were generally familiar to the subjects, but few if any had read the specific articles. This provided the basis for categorizing the content as semantic rather than episodic. The basic research question dealt with the effects of feedback on sentence clustering at the paragraph level. Our assumption was that the looser sequential constraints might yield some differential results.

Method

Sixty volunteers from the Introductory Psychology pool were randomly assigned to the experimental conditions. Each text was typed on cards, one sentence per card, and separately randomized decks were prepared for each passage. A slotted board was provided so the reconstruct subjects could place cards in slots as they were read, and reread as new cards were placed. Two randomly ordered versions were provided for each passage. Tokens were provided the feedback subjects to use for feedback requests. Appendices A and B list the original order of sentences by paragraphs for the Chesapeake and Computer Crime passages respectively.

Three conditions were used with each of the two texts creating six experimental cells. The reconstruct-feedback condition required the subject to sort the scrambled sentences into five paragraphs. Subjects had available 25 tokens to request feedback, for a total of 25 individual requests. In the reconstruct-no feedback condition subjects sorted the sentences into five paragraphs but could not request assistance. For the read-only group the subjects read the sentences aloud from the assigned deck, and then read them aloud again, this time from the alternatively ordered deck. The decks were counterbalanced for the other two conditions.

For sorting, the subjects were given a card deck and told that it consisted of 13 (or 18) sentences which comprised 5 paragraphs. The slots in the board were color coded into 5 zones of 7 contiguous slots each. The task was to put the sentences into the zones which would represent the paragraphs in the passage as originally written. For feedback, a subject could ask if a sentence belonged in the same paragraph with some other indicated sentence, and the experimenter would answer 'yes' or 'no,' as was appropriate. A token was given up for each request.

Following completion of a practice task and the appropriate experimental task, the subject was asked to complete a written recall of the experimental passage, and then to complete a forced-choice recognition test consisting of the experimental sentences paired with paraphrases. Finally, the subject was asked for a phrase telling what the passage was about, and also asked to order the paragraphs as they might be sequenced in the original.

Analysis

The principal dependent measures were recall of idea units, recognition of original sentences, and placement agreement with original sentences in paragraphs. Recall scores were the proportion of idea units obtained to the total number of idea units in each text. Initially the passages were broken down into idea units following Bovair and Kieras (1981). The Chesapeake Bay passage had 44 idea units, while Computer Crime had 66. The ratio of differences in number of idea units between the two passages (.67) is almost identical to the ratio of differences in number of sentences (.68). Hence the two passages have a similar conceptual density, differing in length. Both, as we have noted, were considered semantic in terms of content.

The proportion of recognition scores were the number of original sentences recognized when paired with a paraphrase, to the total number of sentences in the original passage. As noted, Chesapeake Bay contained 13 sentences and Computer Crime consisted of 19.

Placement agreement within paragraphs was a calculated statistic indicating the degree of agreement in placement of sentences with placement of those sentences in the original. In order to make these data comparable to results in our prior experiments, we made the statistic similar to Kendall's tau. To distinguish our statistic from Kendall's tau, we labelled it cluster tau, (C-tau). Final placements for each subject were recorded, while the subject performed the written recall task.

All sentence pairs were scored +1 if in agreement and -1 if in disagreement with the original with respect to placement in the same or different paragraphs. C-tau is therefore the total number of agreements less disagreements as a proportion of total pairs, as in Kendall's tau. An example might help. Let us suppose that the Chesapeake Bay original text contained in the first paragraph, Sentences 1, 2, and 3. However, a subject reconstructing the text actually placed Sentences 1, 2, and 4 in this first paragraph. Sentence 1 is correctly placed with 2 (+), but incorrectly placed with 4 (-). In addition, since Sentence 1 was not paired with Sentence 3, this also yields a (-) placement. However, it was not included in the same paragraph with any of the remaining 10 sentences (which was correct), hence all the other paired comparisons yield a (+) score. For Sentence 1 this yields 10 (+) and 2 (-) scores. This process was repeated for all sentences. Agreements and disagreements were calculated for paragraph clusters rather than sequences of sentence pairs.

As noted earlier, we argued that sequential requirements for comprehension at the paragraph level would probably be less demanding. We might assume, therefore, that clustering sentences by feedback would yield some differential results as compared to sequential ordering at the sentence level. Indeed, the data both parallels previous findings, but also shows some differences.

ANOVA for recall by passage and processing yielded a statistically significant main effect for passage ($F(1,42)=7.45$, $p<.01$) and process ($F(1,42)=3.2$, $p<.05$).

Insert Table 1 about here

The Chesapeake recall mean was 31.00 compared to 21.29 for Computer Crime. As in previous studies recall was better for the shorter passage. For the processing conditions, the reconstruct no-feedback mean (32.50) was superior to the read-only (22.19) and reconstruct-feedback (23.75).

ANOVA for recognition yielded a main effect for process ($F(2.42)=7.25$, $p<.01$). Again, the reconstruct-no feedback condition mean (87.25) was superior to read-only (72.19) and reconstruct-feedback (82.19). The feedback provided did not appear to assist the reconstruction process, and might have hindered the task.

Insert Table 2 about here

ANOVA for C-tau yielded statistically significant effects for passage ($F(1,32)=6.40$, $p<.01$) and process ($F(1,32)=7.77$, $p<.03$). This statistic was calculated for the two reconstruction groups only. C-tau for the Chesapeake

----

Insert Table 3 about here

----

passage was .69 compared to .58 for Computer Crime. The reconstruct-feedback mean was .68 compared to .55 for reconstruct-no feedback. As in previous studies the increased concordance associated with feedback did not lead to increased comprehension. The means and standard deviations for recall, recognition, and C-tau are given in Table 4. Again, for C-tau only the reconstruction conditions are included

----

Insert Table 4 about here

Conclusions

An almost axiomatic belief among institutional developers is that feedback enhances performance (Berliner, & Rosenshine, 1977; Langer, 1983). However, in a recent paper Mannes and Kintsch (1987) suggest that learner assistance does not lead to increased comprehension at all memorial levels. While their study did not deal with feedback, assistance was provided to the learner in the form of outlines either consistent or inconsistent with an experimental text subsequently provided. The consistent-outline students did better on comprehension tasks closely relevant to the materials provided, but the inconsistent-outline students did better on tasks involving deeper processing. The results not only parallel earlier findings (Gallagher, 1981; Battig, 1966), but are reflected in our data.

In this study as in our previous work feedback assisted concordance, but this was not necessarily reflected in improved comprehension. For shorter passages it is likely that assistance is simply not needed. That is, the subject is likely to have adequate working memory resources to retain the passage.

While feedback assists concordance, the fact that concordance does not necessarily lead to increased comprehension is possibly a function of our measures. The reconstruction process which involves continual sorting probably does enhance recognition, which is based on familiarity with the sentences. Recall requires the generation of content, and matching to sequence does not appear to enhance or assist storage or subsequent retrieval. Indeed, semantic as well as episodic content implies preexisting schemata available to the learner. The problem lies in the kinds of learning outcomes sought.

At this point it is becoming clearer that there is no universal set of relationships between discourse structure, knowledge representations in text and learner, and feedback which can be logically assumed to exist. The commonly and widely held belief that feedback is better must be hedged on a number of critical issues.

References

Battig, W. F. (1966). Facilitation and interference. In E. A. Bilodeau (Ed.), Acquisition of skill. New York: Academic Press. Pp. 215-254.

Berliner, D. C., & Rosenshine, B. (1977). The acquisition of knowledge in the classroom. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Ed.), Schooling and the acquisition of knowledge. Hillsdale, NJ: Erlbaum. Pp. 375-404.

Bovair, S., & Kieras, D. E. (1981). A guide to propositional analysis for research on technical prose. Tucson, AZ: University of Arizona. (Tech. Rep. No. 8)

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. Cognitive Psychology, 2, 331-350.

Britton, B. K., & Black, J. B. (1985). Understanding expository text: from structure to processes and world knowledge. In B. K. Britton & J. B. Black (Eds.), Understanding expository text. Hillsdale, NJ: Erlbaum. Pp. 1-8.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 12, 335-359.

Frase, L. T. (1969). Paragraph organization of written materials: the influence of conceptual clustering upon level of organization. Journal of Educational Psychology, 60, 394-401.

Frederiksen, C. H. (1977). Semantic processing units in understanding text. In R. O. Freedle (Ed.), Discourse processes: advances in research and theory. Norwood, NJ: Ablex. Pp. 58-87.

Gallagher, J. P. (1981). The effectiveness of man-machine tutorial dialogue for teaching attribute blocks problem-solving skills with an artificial intelligence CAI system. Instructional Science, 10, 291-332.

Getsie, R., Langer, P., & Glass, G. V. (1985). Meta-analyses of the effects of type and combinations of feedback on children's discrimination learning. Review of Educational Research, 55, 9-22.

Hintzman, D. L. (1978). The psychology of learning and memory. San Francisco, CA: W. H. Freeman.

Irwin, J. W. (1982). Effects of coherence explicitness on college readers' prose comprehension. Journal of Reading Behavior, 14, 275-284.

Just, M. A., & Carpenter, P. A. (1987). The psychology of reading and language comprehension. Boston, MA: Allyn and Bacon.

Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: Erlbaum.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. Psychological Review, 85, 363-394.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. Nilsson (Ed.), Perspectives on memory research. Hillsdale, NJ: Erlbaum. Pp. 33-62.

Kulhavy, R. W., Schmid, R. F., & Walker, C. H. (1977). Temporal organization in prose. Review of Educational Research, 14, 115-123.

Langer, P. (1983). Modifications of current feedback strategies: a text synthesis approach. Annual Report of the Air Force Office of Scientific Research Summer Faculty Research Program. St. Cloud, FL: Southwestern Center for Electrical Engineering Education Press.

Langer, P., Keenan, V., & Culler, J. (1987). Effect of content of feedback in processing discourse. Psychological Reports, 60, 1191-1200.

Langer, P., Keenan, V., & Culler, J. (1988). Effects of signalling and feedback on the comprehension of reconstructed text. Perceptual and Motor Skills, 66, 27-39.

Langer, P., Keenan, V., & Medosch-Schonbeck, C. (1986). Some possible limitations of feedback in instruction. Psychological Reports, 59, 731-741.

Lorch, R. F., Jr., & Chen, A. H. (1986). Effects of number signals on reading and recall. Journal of Educational Psychology, 78, 263-270.

Mannes, S. M., & Kintsch, W. (1987). Knowledge organization and knowledge text. Cognition and Instruction, 4(2), 91-115.

Meyer, B.J.F. (1985). Prose analyses: purposes, procedures, and problems. In Britton, B. K., & Black, J. B. (Eds.), Understanding expository text. Hillsdale, NJ: Erlbaum (pp. 11-64).

Meyer, B. J. F., & Freedle, R. O. (1984). Effects of discourse type on recall. American Educational Research Journal, 21, 121-143.

Reder, L. M. (1980). The role of elaboration in the comprehension and retention of prose: A critical review. Review of Educational Research, 50, 5-53.

Schultz, C. B., & DiVesta, F. J. (1972). Effects of passage organization and note taking on the selection of clustering strategies and on the recall of textual material. Journal of Educational Psychology, 63, 244-252.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 9, 77-110.

Tulving, E. (1983). Elements of episodic memory. New York: Oxford Press.

van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

Woodward, B., & Bernstein, C. (1976). The final days. New York: Simon & Schuster.

Young, S. R. (1984). A theory and simulation of macrostructure (Tech. Rep. No. 134). Boulder, CO: University of Colorado, Institute of Cognitive Science.

Appendix A

Chesapeake Bay Paragraph Arrangement

1. Ten percent of the water in Chesapeake Bay contains virtually none of the dissolved oxygen normally found in clean water. This is because every year the spring runoff flushes large amounts of fertilizer residues and other pollutants into the bay.

2. A research team that studies oxygen depletion reports that the amount of oxygen rich water has steadily decreased since the 1930s, when oxygen depletion occurred only in the months of July and August. Because of oxygen depletion, enough damage has now been done to seriously threaten the Chesapeake Bay commercial fisheries. This lead President Reagan to call for a "long, necessary effort to clean up" the Chesapeake last January.

3. The affected area extends from the Susquehanna River to below the mouth of the Potomac. All of the water in between these areas has been highly polluted and suffers from oxygen depletion.

4. The pollution provides nutrients that support large populations of microscopic plants. When the plants die, they sink to the bottom of the bay and begin to decay. From spring until fall, the process of decay consumes the bay's oxygen faster than it is replenished from the atmosphere.

5. The lack of oxygen is the leading suspect in the reduced catch of valuable food fish. All the deep water blue crabs are gone. Additionally, Virginia oystermen have reported hauling in "black bottoms" or foul smelling sediment containing only dead shellfish.

Appendix B

Computer Crime Paragraph Arrangement

1. Seven young computer hackers in Milwaukee got some unwelcome publicity by managing to tap into the computers at Los Alamos National Laboratories in New Mexico, where many of the nation's nuclear bombs are designed. Although the group of kids did not uncover any top secret weapons plans, they caused dismay in government agencies, corporations, hospitals and other institutions across the country by spotlighting the vulnerability of all computerized information.

2. The members of the group, who ranged in age from 15 to 22, proved the contention of artificial intelligence expert Marvin Minsky that "computers are dumb machines, and people can still easily outwit them." The group of kids were not computer wiz kids or mad geniuses bent on cracking almost unbreakable electronic codes. Rather, they were just ordinary home computer hobbyists who used a few standard programming techniques to dial up private files. Among the other computerized data they tinkered with were the records of radiation therapy patients at Manhattan's Sloan-Kettering Cancer Institute.

3. The group began their electronic raids a while back, but stepped up their raids after seeing the movie "War Games," where a clever young kid taps into the nation's defense system computer and begins playing a game of global thermonuclear war. Like the kid in the movie, all the group needed was a personal computer and a modem, a device which links the computer to a telephone. With this equipment they could call up all the computers connected by a giant network known as telenet. Although each institution subscribing to telenet supposedly has its own security system, sneaking around that security proved to be only too simple.

4. At both Los Alamos and Sloan-Kettering, the tampering was noticed quickly, and the FBI notified. Decoy files were then created in the data banks to try to trap the unauthorized users. Anonymous tips finally led the FBI to the group of kids from Milwaukee. As of recently, no charges have been brought against the group, but in the world of computers, the panic was on.

5. Donn Parker, a computer crime expert at SRI International in Menlo Park, California, views the incident as part of a larger fad. There is an epidemic of malicious computer break-ins. Everybody's privacy is in jeopardy. Just about the only people not upset by the rumpus were the cryptographers who design high security computer codes. It seemed likely that their skills would be in great demand for years to come.

Table 1

ANOVA: Proportion of recall scores

Source	SS	Df	MS	F	Sig
Main Effects	2119.56	3	706.52	4.66	.007
Passage (P)	1131.02	1	1131.02	7.45	.009
Condition (C)	988.54	2	494.27	3.26	.05
2-way Interactions	595.29	2	2917.65	1.96	.15
P x C	595.29	2	297.65	1.96	.15
Explained	2714.85	5	5642.97	3.58	.009
Residual	3673.13	42	151.74		
Total	9087.98	47	193.35		

Table 2

ANOVA: Proportion of recognition scores

Source	SS	Df	MS	F	Sig
Main Effects	1896.38	3	632.13	4.88	.005
Passage (P)	16.33	1	16.33	.13	.72
Condition (C)	1880.04	2	940.02	7.25	.002
2-way Interactions	167.54	2	83.77	.65	.53
P x C	167.54	2	83.77	.65	.52
Explained	1063.92	5	412.78	3.19	.02
Residual	5444.00	42	129.62		
Total	7507.92	47	159.74		

Table 3

ANOVA: C-tau

Source	SS	Df	MS	F	Sig
Main Effects	.17	2	.084	5.58	.009
Passage (P)	.097	1	.097	6.40	.02
Condition (C)	.072	1	.072	4.77	.04
2-way Interactions	.007	1	.007	0.48	.50
P x C	.007	1	.007	0.48	.50
Explained	.18	3	.06	3.88	.02
Residual	3.43	28	.02		
Total	.50	31	.02		

Table 4

Means and standard deviations: Proportion recall, proportion recognition, and C-tau

Source	Recall		Recognition		C-Tau	
	Mean	SD	Mean	SD	Mean	SD
Chesapeake Bay	31.00	16.53	81.13	13.53	.68	.15
Computer Crime	21.29	8.32	79.96	11.94	.55	.13
Read-only	22.19	8.30	72.19	12.55	n/a	n/a
Reconstruct-no feedback	32.50	17.32	87.25	9.51	.58	.12
Reconstruct-feedback	23.75	13.39	82.19	11.17	.68	.15