DOCUMENT RESUME

ED 127 785 FL 007 796

AUTHOR Perfetti, Charles A.; And Others
TITLE Memory During Oral and Silent Reading.
PUB DATE 76
NOTE 17p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Cognitive Processes; Language Research; *Memory; Psycholinguistics; Reading; *Reading Comprehension; *Recall (Psychological); Retention; Semantics; Syntax; *Vocabulary

ABSTRACT

Following reading and listening tasks, adult long-term memory is high in semantic information and low in syntactic and lexical information. Comprehension during reading and listening must depend to some extent, however, on short term retention of linguistic information that is less abstract and shares more features of the input than the semantic products of comprehension observed in long-term memory. This paper describes a study focussing on short-term memory for discourse, with three basic issues: (1) the role of linguistically marked units in recallability of words just read, (2) information organization schemes as opposed to short-term memory handling of information, and (3) oral vs. silent reading. Thirty-two third and fourth graders from an urban parochial school in a white working class neighborhood were divided into two different skill levels determined by scores from the Metropolitan Achievement Test. Both simple and difficult reading tasks were given. The overall results support a model of reading comprehension skill emphasizing short-term memory quantitative differences in memory function during reading. The fact that differences were found between the groups in listening, silent reading and oral reading but not in probe digit performance suggests that differences in language memory functions are not wholly dependent on decoding or simple short-term memory capacity. (CLK)

Memory During Oral and Silent Reading

by

Charles A. Perfetti, Laura Bell and Susan Goldman

University of Pittsburgh

During normal reading, it is apparent that memory for particular words and phrases is a fleeting thing. The contents of adult long-term memory after reading are high in semantic information and low in syntactic and lexical information (Perfetti & Garson, 1973). The situation is similar in the case of listening (Sachs, 1967). However, comprehension during reading and listening must depend to some extent on the short-term retention of linguistic information that is less abstract and shares more features of the input than the semantic products of comprehension observed in long-term memory. The characteristics of this short-term memory for discourse is the focus of this study, centered around three issues.

One issue is whether linguistically marked units have any special status concerning the recallability of words that have just been read. Jarvella (1971) showed that in running discourse the probability that a recently heard word will be recalled by adult listeners is related to whether the word is from a previous sentence or from the current sentence. Perfetti and Goldman (1976) have shown a similar result in a probe discourse study of children grouped by level of reading achievement. Thus, one process in understanding spoken discourse is the use of sentence or clause units in the short-term retention of discourse. One interpretation of this is that a sentence or clause boundary appears to serve as a signal for semantic recoding of a linguistic unit which permits the next

segment to take its turn in a limited capacity working memory. In the case of reading, we would expect a similar state of affairs.

In addition to observing a sentence boundary effect in oral discourse, Perfetti and Goldman (1976) found that skilled readers recalled more words than less-skilled readers for both "near" probes, about five words back and "far" probes, about 14 words back, regardless of sentence boundaries. Differences between skilled and less-skilled readers were minimal for a single clause sentence and quite large for two clause sentences. Of special interest was the fact that reader groups did not differ in probe digit recall. Thus, the relationship between reading skill and memory for spoken discourse does not seem to be a matter of simple short-term memory capacity but of capacity to handle linguistic input of some complexity, e.g., sentences with two clauses.

Consider now the reading situation. It comes as no surprise to learn that reading skill, as measured by standard reading achievement tests, is related to the recall of information after it is read. Whether a test is one of "recall" or of "comprehension" it is measured on the retrievability of stored information. Thus, a general source of reader differences may be in the use of effective information organization schemes to facilitate memory performance. However, another model of reading achievement differences emphasizes the short-term handling of information. Given that a word is correctly decoded can it be recalled just after it is read? If not, then it is reasonable to conclude that short-term linguistic functions are involved rather than the application of organization schemes. Thus a model that emphasizes the difficulty in handling all the encoding and interpreting functions of comprehension receives some support.

The third issue is related to the above characterization of memory function during reading, and that is oral vs. silent reading. In oral reading, verbatim memory for just read material should be high. Decoding of every word has been assured and acoustic and phonetic features of the discourse segment are more available because of the oral encoding. By contrast, one might suppose that silent reading brings about a different situation. The reader is somehow selective about what gets fully decoded and encoded. Furthermore, one might expect reading skill differences to be confined to, or at least greater for, the silent reading situation. This argument is based on the assumption that it is in the silent reading condition that the whole range of problems that potentially trouble the unskilled readers can operate, thus magnifying the difference between skill levels. That is, in the oral situation both skilled and unskilled readers are at least reading every word. The contrasting view is based on the assumption that differences between skill levels in the range we are considering are of one of degree rather than kind. In this model, the limitations on unskilled readers lie largely in decoding (Perfetti & Hogaboam, 1975) and in short term linguistic encoding and memory (Perfetti & Goldman, 1976) and we do not expect to find radically different strategies that would produce certain interactions including one between memory performance and oral versus silent reading.

These aspects of reading process and reader skill are the object of the present experiment. There are two methodological considerations we held important. We wanted a normal reading task in so far as possible. And we wanted some control over where a child was in his reading at the moment of a memory test. The result was a task in which readers of two different skill levels are given a small book containing a story. Each

page contains only three or four lines of text and thus requires fairly frequent page turning. Turning a page occasionally reveals a single word probe rather than text, and the subject responds with the word that had followed it in the text. The distance of the probe in words and the location of a sentence boundary are independently varied as is whether the reading is oral or silent.

Method

Subjects

The subjects were 32 children from an urban parochial school in a white working class neighborhood. Half of the children were in the third grade and half were in the fourth grade. Within each grade two skill levels were identified according to scores on the Metropolitan Achievement Test administered by research staff. Skilled readers were those children whose Stanine score on the MAT was 6 or above, while less-skilled readers and Stanines 4 or below. In one case, an S was chosen whose Stanine was 5 because of too few subjects to fit this criterion. Table 1 presents means and ranges on MAT scores for third and fourth, skilled and less-skilled readers. Subjects were chosen so as to be matched

- - - - - - - - - - - - - - -
Insert Table 1 about here.
- - - - - - - - - - - - - - -

approximately on Otis-Lennon IQ, also administered by research staff. Mean IQs were 109.13 (range = 97-121) and 108.13 (range = 96-117) for the skilled and less-skilled third graders, respectively; and 104.25 (range = 96-110) and 103.0 (range = 93-106) for skilled and less-skilled fourth graders, respectively. There were 9 boys and 7 girls in each of the third and fourth grade groups. Mean age was 8 years, 7 months for the third grade and 9 years, 6 months for the fourth grade.

Materials and Design

Four stories were constructed based on third-seventh grade readers (not used in this particular school). The stories were edited to control for difficulty level and to implement the sentence boundary and probe distance variables. Each story was set in elite type of 4 lines per page. The stories ranged in length from 45-50 pages, with each page containing approximately 20 words. The difficulty level for the stories was determined by the New Flesch Reading Ease Index (Farr, Jenkins & Paterson, 1951). The easy stories had ratings of 82.0 and 83.5 while the difficult stories had ratings of 64.0 and 67.0. However, in the data to be reported, only the two easy stories are included, because only these were read by all subjects.

Each story was designed to have 18 probes at unpredictable pages in the text. A page with a probe was the same as a page with text except that it contained a single word. Nine of the 18 probes were for a near target and nine were for a far target. Thus, a near probe was followed by two content stories, excluding function words before the end of a page and a far probe was followed by five content words before the end of a page. The sentence structure variable was independently manipulated through the construction of 3 structural types as follows, where X indicates a content word (noun, verb, adjective, adverb) and X indicates the word that will appear as a probe on the next page.

Type A: X X X X X X . X End of Page

Type B: X X X X . X X X End of Page

Type C: X . X X X X X X End of Page

Thus in a Type A structure both near and far probes were across a sentence boundary and in Type C both probes were within the sentence. In Type B,

the near probe is within and the far probe is across. Whether a particular structure received a near or a far probe was randomly determined to define one experimental version and then the reverse assignment was used to produce a second version. Thus there were 4 stories of two versions each. The assignment of stories to the modality condition (oral vs. silent) was counterbalanced.

Procedure

For the fourth grade children, there were two successive test days. On the first day, subjects read orally and on the second day they read silently. Two stories were read on each day, the first one always easy and the second one always hard. Third grade subjects had the same procedure except they did not have the hard stories. They read an easy story aloud on the first day and the second easy story silently on the second day.

All subjects were tested individually in sessions of 30-40 minutes. Each day began with one short warm up following general instructions. Children were told they would be tested for their understanding of the story after they read it (they were), and that while they were reading they would sometimes encounter a single word. They were to say the word that had followed this word plus as much that came after that word as possible. Instructions were easily understood with the help of the warm up passage.

Results

There were two related dependent measures. One was the probability of correct verbatim target word recall and the other was the recall proportion of correct verbatim words following the probe, including the

target. A measure was also made of nonverbatim meaning-preserving paraphrase recall of the target word.

The analysis was made on the combined subject population, thus excluding for the fourth grade the hard story, which the third grade subjects did not receive. Thus, the data are from stories all Ss read. These data were analyzed in a 2 x 2 x 2 x 2 x 2 x 3 analysis of variance. The three between factors were reading skill, grade, and materials set, while the within factors were 2 Reading Modes, 3 Structure Types and 2 Probe Distances.

Verbatim Target Recall

Table 2 shows the probability of target recall. It can be seen that skilled readers recalled more targets than less-skilled readers.

- - - - - - - - - - - - - - -
Insert Table 2 about here.
- - - - - - - - - - - - - - -

$F(1,24) = 9.54$, $p<.001$. The data are combined for the third and fourth grades, because there was no significant grade effect, $F(1,24) < 1.0$. Recall was higher with oral reading than with silent reading, $F(1,24) = 5.30$, $p<.04$, and the effects of Probe Distance, $F(1,24) = 36.59$, $p<.001$, and Structure Type, $F(2,48) = 6.30$, $p<.004$, were both significant.

Of particular interest is the significant interaction of Structure Type with Probe Distance, $F(2,48) = 7.23$, $p<.002$, which reflects the effect of sentence boundaries. Tukey's honest significant difference (HSD) test on the cell means, $p<.05$ level, revealed that if the probe was in the same sentence, distance had no effect: Type C_{near} = Type C_{far} = Type B_{near}. Only if the probe crossed the sentence boundary did distance have an effect: Type B_{near} = Type A_{near} Type B_{far} = Type A_{far}. This

pattern of results did not interact with skill level, with reading mode, or with grade. Thus, the effect of Probe Distance is neglible in a C type structure where both probes are within the sentence, while its effect is large in the B type structure where a sentence boundary separates the probes. However, the effect of probe distance can be seen when both probes were from across a sentence boundary as in the A type structure. Thus, what seems to be the case is this: Within the boundaries of short sentences, memory for words is unaffected by the number of intervening words. Across a sentence boundary, memory is affected by number of intervening words.

There was however, a significant grade x skill x reading mode interaction on overall performance, $F(1,24) = 4.454$, $p = .045$. A Tukey's HSD test on the cell means, $p<.05$, revealed that the third grade less-skilled group performed significantly better on oral than silent reading while the other three groups performed equally well in each reading mode.

When the criterion for target recall was relaxed to include word substitutions that did not significantly alter the meaning of the target, the patterns of significant results are essentially the same for the effects of skill, grade, and the probe distance x surface structure interaction. However, the main effect of reading mode is not longer significant, $F(1,24) = 1.5$, $p = .23$, and the grade x skill x reading mode interaction is only marginally significant, $F(1,24) = 3.7812$, $p = .06$. A Tukey HSD test on the cell means for this interaction showed that the 3rd grade high skill group did significantly better on oral than on silent but the 3rd grade high skill group did significantly better on silent than oral reading. For the 4th grade groups, performance was the same in oral and silent reading. A comparison of target word paraphrase frequencies shows that the 4th grade groups paraphrased equally often in

the 2 reading modes while both groups of 3rd graders showed twice as many paraphrases in the silent as in the oral condition. This may suggest that oral reading enhances verbatim retention but that short-term memory for meaning is unaffected. However, it is important to note that neither the sentence boundary effects nor reader group differences are dependent on a verbatim recall measure.

Proportion Recall

Compared with the target recall measure, the proportion of words recalled following the probe, including the target, produced both a decreased Reader Group effect, $F(1,24) = 1.99$, $p = .17$, and a reduced Reading Mode effect, $F(1,24) = 1.33$, $p = .24$. However, the Probe Distance effect, $F(1,24) = 20.15$, $p<.001$ and the Structure Effect, $F(2,48) = 3.50$, $p<.04$ were significant as was the interaction, $F(2,48) = 10.03$, $p<.001$. Tukey HSD on cell means revealed a slightly different pattern of significance on this measure. For proportion recalled performance on Type B_{far} is significantly poorer than performance on Type A_{far}, even though both are across a sentence boundary and at the same probe distance. What differs is the position in the to be recalled sentences. For Type A_{far}, the probe is at the initial part of the previous sentence while for Type B_{far}, the probe is at the final portion of the previous sentence. There may be a tendency for subjects to recall to the end of a sentence and stop. This is consistent with a nonsignificant difference in proportion recalled for Type A_{far} = Type A_{near} = Type C_{far}. However, the boundary effect does appear: performance on Type A_{near} is significantly poorer than on Type C_{near} = Type B_{near}.

Thus, there are two problems in interpreting this measure. The first is noted above relates to the number of words between the probe and the end of the sentence. The second relates to the emphasis on target words rather than complete recall in the instructions to the subject. Thus, the criterion of a subject for producing the complete sentence fragment when he was uncertain of its verbatim form contributes to this measure. There was also a significant interaction of Reading Mode with Materials Set, $F(1,24) = 11.69$, $p<.003$ on this measure. One story showed a strong oral-silent difference favoring oral, while the other story showed no difference.

Reading Errors

The oral reading condition was recorded and an analysis of oral reading errors was made. Less-skilled readers, of course, made more errors overall the skilled readers. Reading errors were further classified according to whether they preserved meaning and syntax, syntax only, or neither. Total reading errors correlated negatively with reading level measured by the MAT, $r=.61$, ($p<.001$). It also correlated negatively with decoding speed of single words measured on these subjects in another experiment (Perfetti & Hogaboam, 1975), $r=-.71$, ($p<.01$). Errors during oral reading correlated also with performance on the probe memory task, $r=-.48$, ($p<.01$) for silent reading and non-significantly with oral reading. This result may reflect the fact that subjects were prompted on words they did not attempt to say during the oral reading.

Comprehension Test

The comprehension test was designed mainly so that expecting it would encourage subjects to understand what they were reading. The tests

were 8 short answer questions requiring factual information from the story. In general comprehension performance after an oral and silent reading produced a similar pattern of correlations. However, the correlations involving comprehension following oral reading tended to be non-significant except for a positive correlation with syntax preserving substitution errors, $r=.81$, $p<.001$. Silent comprehension correlated significantly with recacall performance measured by proportion of words recalled, $r=.47$, $r<.01$, and with oral recall performance.

Discussion

The results of the experiment support the hypothesis that sentence boundaries are functional in verbatim short-term memory during reading. In fact, since the sentence boundary effect was observed in all three dependent measures, we may conclude that it is not merely a question of verbatim retention. Verbatim target recall, paraphrase target recall, and proportion of total words recalled are all higher within a sentence boundary then across. These results thus agree with Perfetti & Goldman (1976) and Jarvella (1971) and strengthen the view that sentences are units of discourse processing in working memory.

These results provide no support for skill related qualitative difference in discourse memory. Skilled readers were significantly better at discourse recall by all measures, except proportion recall. In this case, separate analysis of third and fourth grade data revealed a significant effect for third grade, $F(1,8) = 11.76$, $p<.01$, but not for fourth grade.

There was some evidence in all the measures that reader group differences were smaller for the fourth grade than for the third grade.

A possible explanation for this might be in the decoding skills of the fourth graders. Table 3 shows a breakdown of the 16 fourth grade subjects according to reading comprehension level, defined by the MAT, and decoding speed, defined by vocalization speeds to pseudowords taken in another experiment (Perfetti & Hogaboam, 1975). Although a separate analysis of

- - - - - - - - - - - - - - - -
Insert Table 3 about here.
- - - - - - - - - - - - - - - -

variance on these four groups, did not show statistically significant group effects, it can be seen that the overall performance of the group identified as slow decoders is lowest, while the other group show little difference. Because of the close relationship between comprehension and decoding speed (Perfetti & Hogaboam, 1975(b), it is possible that a group such as that labeled Low-Medium represents anomalous group of unskilled comprehenders. They show neither slow decoding nor ineffective discourse memories and their difficulties in reading comprehension lie elsewhere.

This much is quite speculative, however, and the more general conclusions for the present study are that differences connected with reading skills are quantitative rather than qualitative. This generalization is evidenced in the result of no reader group interaction with Reading Mode nor with Structure Type nor with Probe Distance. Thus, skilled readers do not seem to differ from less-skilled readers in their use of sentence units in working memory. Perfetti & Goldman (1976) found this to be true in listening to discourse as well.

As in the case of the sentence boundary effect, the fact that there was no interaction between reading skill and Reading Mode supports the model of reading comprehension skill that emphasizes nonqualitative factors.

However, the possibility of some individual differences in the relationship between oral and silent reading can be seen in a separate third grade analysis. For the target recall measure on the third grade subjects, there was a significant three-way interaction of Reader Skill x Reading Mode x Probe Distance, $F(1,8) = 8.40$, $p<.02$. Near probes produced better recall than far probes except for skilled readers reading silently. During silent reading skilled readers recalled as often to a far probe as to a near probe, while during oral reading near recall was higher, as it was for unskilled readers in either silent or oral reading. This does suggest the possibility that silent reading preserves more distant discourse in memory at the expense of more recent discourse for skilled readers. Thus, there is a sense in which silent reading may be significantly different from oral reading. Developing a high level of reading skill includes acquiring a reading strategy does not simply mimic the oral reading process. The importance of such strategies for skilled reading has been generally recognized in accounts of reading processes (e.g., Gibson, 1975). For example, in the present case, skilled silent readers may have slightly less available recent segments to the extent that such recent segments are redundant or less useful in some other way. Correspondingly, somewhat more distant segments are kept active in memory as sentence components and adjacent clauses are integrated. All this is somewhat speculative, of course, and the effect may be of limited generality. After all, this interaction was not observed among 4th grade readers. In the context of the more general findings of the experiment, the finding of this particular interaction merely suggests a possible interesting type of processing difference.

Overall, the data gives support to a model of reading comprehension skill that emphasizes short-term quantitative differences in memory function during reading. The "on-line" discourse memory of skilled readers is better than the on-line discourse memory of unskilled readers. A large part of this difference is probably a question of decoding skill. A reader whose processing is occupied with effort decoding will have less processing to deviate to keeping recent discourse segments in memory. However, this cannot be the whole answer, in light of Perfetti & Goldman's (1976) finding of reading skill differences in a discourse listening task. The fact that these differences are found in listening, silent reading, and oral reading, but not in probe digit performance suggests differences in language memory functions not wholly dependent either on decoding or on simple short-term memory capacity.

Table 1

Mean MAT Scores for Third and Fourth Grade

	Fourth					
	Skilled			Less-Skilled		
	Raw	Percentile	Stanine	Raw	Percentile	Stanine
Mean	32.75	70.75	6.13	19.0	26.25	3.63
Range	30-37	60-86	6-7	9-24	2-40	1-5
			Third			
Mean	27.0	79.38	6.75	14.75	26.0	3.5
Range	22-32	62-90	6-8	11-17	10-38	2-4

Table 2

Mean Probability of Verbatim Target Recall for Skilled and Less-Skilled Readers in 3rd and 4th Grade

Type A: X X X X X X . X	Far	Near	Type B: X X X X . X X X	Far	Near	Type C: X . X X X X X X	Far	Near
Skilled	.63	.82		.69	.80		.82	.83
Oral	.60	.85		.65	.85		.83	.85
Silent	.65	.79		.73	.75		.81	.81
Less Skilled	.48	.79		.47	.79		.73	.74
Oral	.52	.81		.44	.88		.79	.75
Silent	.44	.77		.50	.69		.67	.73

Summary

Comparison	Mean	Comparison	Mean	Comparison	Mean
Skilled	.76	Near	.80	Near Within	.80
Less-Skilled	.67	Far	.64	Near Across	.79
Far Within	.78	Oral	.74		
Far Across	.56	Silent	.70		