Good and poor readers drawn from seventh-grade classes read one prose passage and listened to a second one. They were tested following each passage, for comprehension and recall of that passage. Under both reading and listening conditions, good readers recalled a greater proportion of the stories, and the likelihood of their recalling a particular unit was a clear function of the unit's structural importance; poor readers recalled less of the stories, and their recall protocols were not as clearly related to variations in structural importance. Performance following reading was significantly correlated \( r = .85 \) with performance following listening. The results indicate that poor readers suffer from a general comprehension deficit and that similar processes are involved in reading and listening comprehension. (Author)
Technical Report No. 23

RECALL OF THEMATICALLY RELEVANT MATERIAL BY ADOLESCENT GOOD AND POOR READERS AS A FUNCTION OF WRITTEN VERSUS ORAL PRESENTATION

Sandra S. Smiley, Drew Daulph Oakley, David Worthen, Joseph C. Campione, and Ann L. Brown

March 1977

University of Illinois
at Urbana-Champaign
1005 West Nevada Street
Urbana, Illinois 61801

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

This research was supported by Grants HD-06864 and HD-05951 from the National Institute of Child Health and Human Development.

*Designation as a Center pending approval.
Recall of Thematically Relevant Material by Adolescent Good and Poor Readers as a Function of Written Versus Oral Presentation

Sandra S. Smiley, Drew Dauph Oakley and David Worthen

Western Washington State College

Joseph C. Campione and Ann L. Brown

University of Illinois
Recall of Thematically Relevant Material by Adolescent Good and Poor Readers as a Function of Written Versus Oral Presentation

Adequate comprehension of prose passages is an essential academic skill. Knowledge in schools is largely acquired via the medium of written prose and relatively independent of practical action (Olson, 1976). In accommodating to the demands of a literate tradition, the child must learn to extract meaning from decontextualized messages, whether they are presented in print or in oral (lecture) form. The basic processes of comprehension that underlie this essential skill are largely unexplored. In particular, there is considerable controversy concerning whether reading and listening comprehension tap the same unitary process or depend on totally different mechanisms (Banks, 1974).

In this paper we are concerned with the comprehension processes of good and poor readers. Although it is commonly assumed that the problems of the disabled reader reside mainly in inadequate decoding skills, Guthrie (1973) has shown that poor readers also suffer from inadequate comprehension during reading. Of interest is why this is so. The poor reader could experience difficulty understanding material he is reading because of his struggle with the decoding mechanisms. Guthrie suggests, however, that there may be a group of poor readers whose decoding skills are relatively intact but who are primarily deficient in comprehension ability. One would expect that these poor readers would also be poor listeners; therefore, we decided to look at prose comprehension in children of different reading ability when the text was spoken or presented in print.

Several nontrivial methodological problems confront those who would compare listening and reading comprehension (Banks, 1974; Sticht, Beck, Hauke Kleinman, & James, 1974). We were concerned with three principal problems in the design of this study. The first was the selection of the measure of
comprehension. Most previous comparisons of listening and reading have compared children's performance on separate standardized reading and listening comprehension tests which contain a potpourri of items which could not be regarded as representative of basic underlying processes common to comprehension. For example, many auditory tests include such items as following a sequence of sounds or recognizing non-speech sounds (Witkin, 1971), competencies which should have little to do with reading comprehension. Therefore, we restricted our attention to one measure, the recall of information as a function of its thematic relevance, surely an essential prerequisite for comprehension via any presentation mode.

Children and adults favor the main theme in their retention of prose passages; ideas rated important to the theme are recalled most frequently, but information rated as less crucial is seldom if ever featured in recall (Brown & Smiley, 1977; Johnson, 1970). Even children as young as third grade are sensitive to several degrees of importance and, therefore, are able to extract relevant information from texts (Brown & Smiley, 1977). Below this age there is some evidence that children extract thematically relevant information (Christie & Schumacher, 1975; Korman, quoted by Yendovitskaya, 1971), although their sensitivity to fine levels of importance has not been assessed. The ability to concentrate on main events to the exclusion of nonessential material is a basic cognitive process essential for all comprehension activities, whether in the context of listening or of reading. It is for this reason that we focused on sensitivity to importance as an index of adequate comprehension.

All the children in the Brown and Smiley study listened to the stories they would later be required to remember. We do not know, therefore, whether children possess the same theme sensitivity when reading as they do when listening. Sticht and his associates (Sticht et al., 1974) suggest that listening and reading share the same underlying processes of cognitive...
competency, and that there should thus be a close correspondence between listening and reading skills: good readers should be good listeners. All the students in the Brown and Smiley study were of average reading ability, so that the listening competency of poor readers was not assessed. Therefore, in this study children of average or good reading ability were compared with a group of poor readers in terms of their ability to extract the major theme of a passage when listening or reading.

The use of a recall measure can also be defended in terms of prior methodological problems experienced in this research area. A common measure of comprehension is that the student answer correctly a series of questions concerning text context. The choice of such questions is important. For example, many reading comprehension tests contain standardized curriculum content, with the result that students have been known to score high on such tests without even reading the target passages (Carver, 1971; Tuinman, 1972; Sticht et al., 1974). Thus, although one can deduce that the children could read the test questions, little or nothing is known concerning whether they could comprehend the target passage. Recall as a measure of comprehension is applicable to both listening and reading, but it does have the problem that it measures both comprehension and memory efficiency. For this reason we selected recall as a function of rated importance as a relatively uncontaminated measure, for here the pattern of responding, rather than the absolute amount of recall, is the metric of main interest. A student who recalls little can still favor the important aspects of the story in his reconstruction of the passage.

The second methodological concern was with the selection of the target passage. Texts designed to be spoken differ in many essential features from those designed to be read (Olson, 1976). Spoken messages rely on several extralinguistic factors to determine the total significance of the message (Grice,
1972), such as voiced intonation and stress, gestures, pragmatic implicature, and shared contextual knowledge. Much of the message need not be explicitly conveyed by words, as both the conveyor and receiver can and do depend on "speaker coherence factors" (Wertsch, 1974) in that the listener relies upon (and the communicator presupposes that he will) general background knowledge to disambiguate utterances. Mature communicators obey rules concerning the relationship of what is said, and what is implicated in a particular context. The speaker assumes the listener's mutual understanding of conversational implicature (Gordon & Lakoff, 1971) and only specifies the implicit in cases when he has reason to doubt these assumptions.

In contrast to oral messages, written statements must be explicit and context-free. The writer must convey a message which the reader, removed from him in time and space, and unable to ask questions, can understand. These differences between oral versus written messages present a difficult problem for the comparison of listening and reading skills, for what type of text should one use as target passages, oral or written messages? Reading aloud a written text may unfairly penalize the listener, for such material may be more easily understood via the mode in which it was intended to communicate. However, the opposite solution, writing out a spoken communication, should be equally detrimental to reading, as the rich contextual support which accompanies the spoken message in face to face communicative situations is missing. Reading transcripts of conference discussions will confirm that reading oral communication is not an easy task. We do not believe there is an obvious answer to this dilemma, but decided on a compromise. Certain folk tales originated in an oral tradition and have been handed down by word of mouth from generation to generation. Even when presented as written texts, these tales retain many of the qualities of spoken messages, including dramatic emphasis, etc.: indeed, many of the books of children's folk tales are explicitly intended to be read.
aloud to children. For these reasons we selected two traditional Japanese folk tales, unfamiliar to the children of our culture. Both listening and reading comprehension were assessed on the same type of passages, but the bias toward either a written or spoken message was somewhat attenuated by the choice of folk tales from an oral tradition.

The final non-trivial methodological problem encountered in a comparison of reading versus oral comprehension is the timing of the presentation of material. Ideally, if one wants to compare memory for materials, it is necessary to present the to-be-remembered text for an equal amount of time. But are reading and listening rates comparable? In mature adults, Sticht suggests that they are, but he deliberately excludes speed reading and skipping in this comparison. One control that has been attempted is to present the reading material at the same rate as it would take to listen to the same material. Goldstein (1940) adopted this control; however, he presented the material line by line, thus vitiating the benefits that skilled readers might gain from looking back, skipping, etc. Again we settled on a compromise. The to-be-read material was presented for the same amount of time as the oral presentation; however, sections of text (approximately 6 lines) were presented as a unit, thus allowing the type of guided search said to be characteristic of skilled reading.

Seventh-grade children were selected for study for several reasons. First, Sticht's model is a developmental one, with listening competency preceding reading competency, and it is not until the seventh grade, when basic decoding skills are thoroughly mastered, that Sticht predicts that reading and listening comprehension become comparable. In their extensive review of the literature, Sticht et al. (1974) demonstrate that it is at the middle of the seventh grade that the number of studies reporting listening better than reading is equal to the number of studies finding reading better than listening.
An additional reason for the selection of seventh-grade students is that the existing Brown and Smiley data afford reasonable comparison groups. We already have recall data from third-, fifth-, and seventh-grade children. The original seventh-grade data can be compared with those of the average readers in this study, thereby providing a desirable replication factor. Furthermore, the poor readers in this study were reading two or more years below grade level; therefore, comparisons could be made between the poor readers in this study and normal readers in the Brown and Smiley study of approximately the same reading age.

In summary, the main feature of this study was to compare good and poor readers' comprehension of oral versus written prose under conditions where major methodological problems of prior studies had been eliminated or reduced.

Method

Subjects

The subjects were 36 junior high school students attending the Vista Middle School in Ferndale, Washington. Twenty-one seventh-grade children reading at or above grade level constituted the Non-Title I group. The Title I group consisted of 15 students of approximately equal chronological age (three in sixth grade, nine in seventh grade, and three in eighth grade) who were taking part in a remedial reading program. To enter the program, each student had to be reading two or more years below grade level. All students were in regular classes in the school with the exception of the special tutoring for reading difficulties. Further, each child in the Title I group was in an age-appropriate grade and was singled out for special attention due only to their reading performance.
Stimulus Materials

Two fairy stories were selected as stimulus material, both unfamiliar to the children in this study but both having been used in a prior study on story recall in children (Brown & Smiley, 1977). The two Japanese folk tales, "The Dragon's Tears" and "How to Fool a Cat,"2 were of comparable reading difficulty of grade five (Dale-Chall readability scores of 5.2287 and 5.3682), which should be easily read by normal seventh-grade children. The passages were also of approximately the same length (390 and 403 words, 34 and 28 lines) and of the same number of idea (pausal) units (59 and 54). The stories were divided into subunits following a procedure used by Johnson (1970) and Brown and Smiley (1977). Twenty-one Western Washington State College students were asked to read the stories thoroughly and then to divide the text into individual units by placing a vertical line at a division point. An individual unit was defined as one which contained an idea and/or represented a pausal unit, i.e., a place where a reader might pause. Agreement concerning the divisions into independent units was achieved by eleven or more raters for each unit. After division into independent units, each story was retyped with one unit per line, and a second group of college students was asked to rate the importance of each unit to the theme of the story using a four-point scale. First they were asked to eliminate one quarter of the units which they judged to be least important to the theme of the passage. This procedure was then repeated twice more until only one quarter of the units remained. These last remaining units were judged the most important to the theme, while the set eliminated first were the least important. (For fuller details of the rating procedure, see Brown and Smiley, 1977).

Twenty-seven Western Washington State College students rated the Dragon story and 34 rated the Cat story. On the basis of these importance ratings the structural (pausal) units of each story were rank-ordered from least to most important and divided into four levels of importance in such a way as to ensure that the number of units at each level was approximately equal. The number
of units and range of importance ratings for each level of structural
importance are shown in Table 1. The resultant four sets of units, corresponding
---------
Insert Table 1 about here
---------
to the four levels of importance, were used as the measure of rated importance
against which the recall performance was compared.

Procedure

All students attempted recall on two stories. The students were seen in
small groups. Each subject was randomly assigned to one of two conditions,
Listen-Read or Read-Listen, depending on whether they read or listened to a
story first. In addition, half the students read the Dragon story and listened
to the Cat story, while the reverse was true for the remaining children. The
students in the read condition were presented with slides, each containing
approximately five to six lines of text. Each slide was projected for the
same amount of time that it took to play that part of the text by tape.
Students in the listen condition heard a tape-recorded version of the story
recorded by a female native-speaking American.

Immediately after listening to or reading a story, the children attempted
to write their recall. They were instructed to try to remember as many
details as possible but that they could use their own words to retell the story
if they wanted to.

Results and Discussion

The children's recall protocols were scored by two independent raters
(interrater reliability = .91) who were instructed to score leniently for
gist rather than for exact reproduction of words or phrases. The judges rated
whether or not the main point of each idea unit was retained, irrespective of
wording.
Preliminary inspection of the data revealed no differences attributable to sex, story, or order of presentation, so these variables were not considered further. The mean proportion of units recalled as a function of Group and Importance Level are illustrated in Figure 1. Title I students appear to differ from normal readers both in the efficiency of their recall (amount recalled) and in terms of their sensitivity to level of importance of the constituent units.

Confirming the visual impression, a 2 (Groups) x 2 (Mode: Listen vs. Read) x 4 (Importance Units) mixed analysis of variance revealed significant main effects of Groups, $F(1,34) = 108.22$, $p < .001$, Mode, $F(1,34) = 18.12$, $p < .001$, and Importance Units, $F(3,102) = 75.60$, $p < .001$. Non-Title I students outperformed Title I students (recalling .49, as compared to .18, of the units), and listening (.40) produced better recall than reading (.32).

Of primary importance here is the significant Groups x Importance Units interaction, $F(3,102) = 5.56$, $p < .005$, which is depicted in Figure 1. To describe the interaction more fully, a number of follow-up analyses were conducted. First, simple effects analyses revealed reliable effects due to Importance Units within each group, Non-Title I $F(3,102) = 66.22$ and Title I $F(3,102) = 14.95$, both $p$s < .001. Scheffé comparisons within each group revealed different patterns of differences. For the Non-Title I students, the level 4 units (most important) were recalled significantly better than all other levels, $S^2(3,102) > 18.70$, $p < .001$. In addition, level 3 units were recalled more often than either level 2 or level 1 units, $S^2(3,102) > 13.31$, $p < .01$, which themselves did not differ. Thus, level 4 units were recalled more frequently than level 3 units, which were further differentiated from level 2 or level 1 units.
for the Title 1 group, the only reliable contrasts were between level 4 units and each of the other levels, $S_k^2(3, 102) = 10.34$, $p = .05$; none of the differences between levels 1, 2, and 3 approached significance, $S_k^2(3, 102) = 2.45$, $p = .10$. Thus, whereas the average-to-good readers differentiated the most important (level 4) units from level 3 units, and level 3 units from level 2 and 1 units, the poor readers distinguished only level 4 units from the remainder.

Considering the poor reader, the failure to distinguish between level 1, 2, and 3 units was obtained following both reading and listening. The same relations of units recalled as a function of importance level were also found, e.g., 6, 7, 8, 9; Level 1 after reading, and 7, 8, 9. 10, 11, and 12 after listening. Thus, the same pattern is obtained following both listening and reading, which suggests the absence of any interactions in these data. More importantly, is obtained at two different overall levels of reading.

The absence of the interaction implies the lack of variance indicates that the differences in importance of units on the two groups were obtained following either reading or listening, and is consistent with the notion that comprehension of text or real material involves the same processes. While poor readers cannot distinguish as many levels of structural importance as good readers when they read, neither do they do so when they listen. As an additional test of the relation between listening and reading comprehension, a correlation coefficient was computed, and the resulting $r$ of .86 was highly significant at .001. In addition, the correlation within each population was significant, Title 1 $r = .52$ and non-Title 1 $r = .39$, in both cases were $p < .01$. Good readers thus appear to be good listeners, as Sticht et al. (1974) suggest.

In summary, the main findings are that good readers show better comprehension than poor readers, and that the good readers are sensitive to more
gradients of importance. The fact that both of these effects are obtained when the to-be-comprehended material is either read or heard, together with the correlation between listening and reading scores, suggests that the same processes are involved in the two tasks.

For comparison purposes, we can consider the performance of the present students with that obtained from normal readers in an earlier study. The data from the seventh- and third-grade children in the Brown and Bialer (1977) study are included in Figure 1, seventh grade because they are of comparable age to the two reading groups in this study and third grade because they were the youngest children tested in the original study. As can be seen in Figure 1, the non-titel 1 seventh-grade children are performing essentially at the same level as the seventh graders in the original study, a nice replication. The only point where the scores appear to diverge is at the lowest level of the test. The non-titel 1 students outperform the original seventh-grade children at the lowest level, unlike the original sample, show no further decrement in reading ability with the lowest levels of importance.

Third-grade students perform at a much lower level, not only in comparison to normal readers and seventh graders but also in relation to third-grade children. The differences are both in terms of levels in that the Title I children read at a lower level but in terms of patterns, with the normal readers and third graders showing a greater sensitivity to importance levels.

The third-grade sample was one of the Title I children, while the fourth-grade sample was from the same first-grade sample. The procedure and results of this second study will be described before a general description of all the data.

Experiment 2B

Subjects, in a single session of twenty-three first-grade children (wet
girls and 17 boys) participated in the study. Their mean chronological age at the time of testing was 6 years 8 months.

**Apparatus and stimulus materials.** Each child listened to one of two Japanese children's stories, "The Dragon's Tears" or "How to Fool a Cat," through earphones connected to a cassette tape recorder.

**Procedure.** With the exception of one male subject, all children were tested in single-sex pairs in a room alone with the second author. Subjects were seated on either side of the experimenter and told that they were going to listen to a study, one at a time, with earphones on. Each child listened to either the Dragon or the Cat story. The first child listened to his story and told the first half when it was completed. He then attempted to verbally retell the first of the story to his classmate. Then the second child heard the remaining story and the first child became the listener. The listener-experimenter procedure was found necessary to ensure attention to the story and reasonable attempts to recall in narrative form (Brown, 1976).

**Results.**

Two independent raters listened to each subject's recorded recall and rated that recall for the presence or absence of each pause unit. The agreement between the two raters was 90, and only the ratings of the first rater (the second author) were included in the analysis.

The mean proportion of units recalled as a function of Order of Listening and Importance units were subjected to a 2 x 4 mixed analysis of variance. The effect due to Importance units was significant, F(3,60) = 32.01, p < .001. The mean proportion of units recalled as a function of degree of Importance are also presented in Figure 1. Scheffe post-hoc tests revealed that level 1 recalled differed significantly from all other levels (p < .05), but that levels 1, 2, and 4 did not differ from each other (p > .10).
Thus the first grade level and pattern of recalls are extremely similar to those of the seventh-grade Title I students. Both show low absolute levels of recall and differentiate only between the most important units (level 4) and the remaining three levels, which themselves do not differ.

Discussion

The results of these experiments lead to two major conclusions. First, poor readers suffer from a comprehension deficit when compared to average-to-good readers, a deficit which does not appear to involve decoding skills and which, at least in this research is strikingly large. Second, the data are consistent with the assumption that auditory and reading comprehension depend upon the same basic process(es).

Regarding the first point, the fact that good and poor readers show differential sensitivity to degrees of structural importance confirms Guthrie’s (1973) conclusion that poor readers are deficient in comprehension skills as well as in their decoding skills. That the comprehension difficulties exist independent of decoding problems is clearly indicated here by the finding that the groups’ differential sensitivity to importance is also obtained following auditory presentation of the target passage. Finally, as can be seen in Figure 1, the magnitude of the group difference is large in terms of absolute amount recalled. More important in our view, however, is the fact that it was necessary to test children as young as first grade (Exp. 1b) before finding another group which showed as little sensitivity to importance variations as the Title I seventh graders of the main experiment.

Considering the comparison of reading and listening comprehension, the fact that the effects of structural importance are the same on each at least suggests similarity of underlying processes. This conclusion is of course strengthened by the reliable correlation obtained here between reading and listening scores. As we remarked earlier, poor readers also appear to be poor listeners.
References


Sticht, T. B., Beck, L. J., Hauke, R. N., Kleiman, G. M., & James, J. H. 
Auding and reading: A developmental model. Human Resources Research 

Tuinman, J. J. Inspection of passages as a function of passage-dependency of 

Wertsch, J. V. Simply speaking. In Papers from the tenth regional meeting. 

Witkin, B. R. Reading development through auditory perception training: End 
of budget period report. ESEA Title III Almeda County School Department, 
Hayward, California, October, 1971.

Yendovitskaya, T. V. Development of memory. In A. V. Zaparozhets & D. B. 
Elkonin (Eds.), The psychology of the preschool child. Massachusetts: 
Footnotes

1. This research was supported by Grants HD-06864 and HD-05951 from the National Institute of Child Health and Human Development. The authors would like to thank Mr. Jack Groom, reading specialist, and Mr. Dave Boeringa, Assistant Principal, and the staff and students of the Vista Middle School in Ferndale, Washington, and Mr. Ed Brown, Principal, Ms. Deborah Dougherty, first-grade teacher, and the children of the Silver Beach Elementary School, Bellingham, Washington. Thanks are also due to John Montgomery for assisting in scoring the data. We deeply regret the tragic death of Dr. Sandra Smiley, a close personal friend and highly valued collaborator.

Requests for reprints should be sent to: Dr. Ann L. Brown, Center for the Study of Reading and Cognition, 1005 W. Nevada, Urbana, Illinois 61801.

2. Taken from Japanese Children's Stories, Florence Sakade (Ed.), Rutland, Vermont/Tokyo, Japan: Chas. E. Tuttle Co., 1957.
Table 1

Independent Ratings of Structural Importance for the Two Target Stories

<table>
<thead>
<tr>
<th>Importance Level</th>
<th>Number of Units</th>
<th>Mean Rating</th>
<th>Rating Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The Dragon's Tears</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1 (least)</td>
<td>13</td>
<td>1.43</td>
<td>1.15 - 1.70</td>
</tr>
<tr>
<td>Level 2</td>
<td>16</td>
<td>2.08</td>
<td>1.82 - 2.41</td>
</tr>
<tr>
<td>Level 3</td>
<td>15</td>
<td>2.77</td>
<td>2.44 - 3.11</td>
</tr>
<tr>
<td>Level 4 (most)</td>
<td>15</td>
<td>3.49</td>
<td>3.19 - 3.85</td>
</tr>
<tr>
<td><strong>How to Fool a Cat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1 (least)</td>
<td>13</td>
<td>1.48</td>
<td>1.06 - 1.91</td>
</tr>
<tr>
<td>Level 2</td>
<td>14</td>
<td>2.08</td>
<td>1.97 - 2.29</td>
</tr>
<tr>
<td>Level 3</td>
<td>14</td>
<td>2.93</td>
<td>2.44 - 3.21</td>
</tr>
<tr>
<td>Level 4 (most)</td>
<td>13</td>
<td>3.61</td>
<td>3.24 - 4.00</td>
</tr>
</tbody>
</table>
Figure Caption

Figure 1. Mean proportion recalled as a function of structural importance. The Title I and Non-Title I groups are from the main experiment; the first graders are from Experiment 1b; and the third and seventh graders are from Brown and Smiley (1977).
LEVEL OF STRUCTURAL IMPORTANCE

MEAN PROPORTION CORRECT RECALL

LEVEL OF STRUCTURAL IMPORTANCE