LANGUAGE AND SPEECH, 1989, 32(1), 45 – 67

COMPREHENSION OF TEMPORAL TERMS BY GOOD AND POOR READERS*

PAUL MACARUSO,
EVA BAR-SHALOM,
STEPHEN CRAIN
and
DONALD SHANKWEILER
University of Connecticut

Poor beginning readers often have difficulty comprehending spoken sentences with complex syntactic structures. This study attempts to identify the reasons for this difficulty. Second-grade good and poor readers were tested for comprehension of spoken sentences containing the temporal terms *before* and *after*. Processing load was varied systematically while holding syntax constant in an effort to determine whether processing factors contribute to poor readers' comprehension problems, or whether poor readers are simply lacking the structural knowledge required to understand sentences containing temporal terms. The poor readers' high level of performance under conditions of reduced processing demands suggests that their misinterpretations in spoken language understanding may be due, in large part, to limitations in verbal working memory.

Key words: comprehension, reading problems, syntax, temporal terms

INTRODUCTION

Several studies have shown that children in the early school grades with poor reading ability have comprehension problems in spoken language as well. Some researchers have interpreted this association to mean that poor readers are delayed in the acquisition of certain aspects of syntax (Byrne, 1981; Stein, Cairns, and Zurif, 1984). This viewpoint,

---

* Three of the authors (Paul Macaruso, Stephen Crain, and Donald Shankweiler) are also affiliated with Haskins Laboratories. The research and preparation of the manuscript were supported in part by a Program Project Grant to Haskins Laboratories from the National Institute of Child Health and Human Development (HD-01994). We wish to thank the students, teachers and administrators of the Coventry, CT, elementary schools. Special thanks go to Clarence Edmondson of the George Hersey Robertson School, and to Mary M. Piazza and Annette Murphy of the Coventry Grammar School. We would also like to thank Brian Byrne, Helen Cairns, Ram Frost, and Bruno Repp for their helpful comments and suggestions on an earlier draft of this paper.

Address correspondence to Paul Macaruso, Cognitive Science Center, The Johns Hopkins University, Baltimore, MD 21218. Send requests for reprints to Haskins Laboratories, 270 Crown St., New Haven, CT 06511-6695.

which has been called the *Structural Lag Hypothesis* (Shankweiler and Crain, 1986), stems from investigations of children's comprehension of sentences that contain syntactic structures which emerge late in the course of language development. It has been shown that poor readers[1] have difficulty comprehending sentences with relative clauses (Byrne, 1981; Goldsmith, 1980; Stein *et al.*, 1984) and sentences containing adjectives like *easy* (Byrne, 1981). An alternative explanation of the difficulties these sentences pose, called the *Processing Limitation Hypothesis,* maintains that poor readers' inferior performance is due to the memory demands associated with these sentences in typical comprehension tasks (Crain, 1987; Shankweiler and Crain, 1986; Smith, Macaruso, Shankweiler, and Crain, in press). To investigate the nature of poor readers' comprehension problems, this paper examines in detail one putatively late-emerging construction: Sentences containing the temporal terms *before* and *after.*

Although temporal terms have not previously been employed in research on poor readers' spoken language comprehension, they have been studied extensively in the literature on the language development of preschool children. The basic finding in these studies is that children younger than five or six have difficulty comprehending sentences containing temporal terms, especially when they present conflict between the conceptual order of events and the order in which the events are mentioned (Clark, 1970, 1971; Johnson, 1975). An example is given in (1). On the other hand, for sentences containing temporal terms in which there is correspondence between conceptual order and order of mention, as in (2), preschool children usually show correct comprehension.[2] These results suggest that some sentences containing temporal terms may not be completely understood by most children before the age at which reading instruction begins.

(1) Before you push the car, push the horse.

(2) After you pick up the truck, pick up the car.

Beginning with the earliest research by Clark (1970, 1971), a number of explanations of preschool children's difficulties with temporal terms have been offered. Most of the explanations presuppose the existence of distinct stages in language acquisition. That is, it is generally assumed that some syntactic structures emerge later than others, and that preschool children initially lack the structural knowledge to comprehend the late-emerging structures. Consistent with this viewpoint is the finding by Amidon and Carey

---

1 We use the term "poor readers" to refer to children in early school grades who exhibit specific reading disability; that is, those who show a marked disparity between their level of performance in reading and the level that would be expected on the basis of their age, school experience, and general intelligence. There is a growing consensus that the problems underlying the reading deficit in most of these children reside in the language domain and do not reflect deficits in visual perception or general analytic skills. The poor readers are compared with age-matched controls who are proceeding at the expected rate of reading acquisition.

2 However, it might be the case that preschool children act out these sentences correctly for the wrong reasons; e.g., they might be following a simple order of mention strategy.

(1972) that preschool children perform less accurately on sentences like (1) than on sentences like (3), which pose the same conflict between conceptual order and order of mention, but which are structurally less complex. This difference was interpreted as evidence that sentences containing temporal terms emerge late in the course of language acquisition because of their structural complexity.

(3) Push the car last; push the horse first.

However, a processing explanation could also be put forward to account for the problems preschool children have in comprehending sentences containing temporal terms. It may be that these children err more on sentences like (1) than on sentences like (3) because, in all of the comprehension studies discussed thus far, the researchers failed to control for a presupposition associated with (1) – but not with (3) – that the child intends to push the car. Accommodating unmet presuppositions requires the listener to revise his or her mental model of the discourse setting so that it corresponds to the mental model of the speaker (Lewis, 1979). Revisions such as these may overtax the working memory resources of young children (see Crain and Steedman, 1985, and references therein).

To test this explanation of preschool children's difficulties with temporal terms, Crain (1982) developed a test procedure to satisfy the presupposition inherent in the use of temporal terms. To satisfy the presupposition, the experimenter in Crain's task asks the child to select a toy from an array of toys before each trial. The experimenter then incorporates the child's choice into the subordinate clause of the sentence to be acted out. For example, if the child selected a car, the experimenter could present the child with the sentence given in (1), in which the word *car* is incorporated into the subordinate clause introduced by the temporal term *before.* When contextual support was provided in this way, preschool children showed unprecedented success in comprehending sentences with temporal terms.[3]

In the present study we evaluated the opposing predictions of the Processing Limitation and Structural Lag Hypotheses by having second-grade good and poor readers act out sentences containing temporal terms. For these sentences, the Processing Limitation Hypothesis makes two predictions. First, second-grade poor readers, like normal preschoolers, should show a reduction in performance when sentences present conflict between conceptual order and order of mention, as in (1). We will refer to

---

[3] The results from Crain's (1982) study have been extended by Gorrell, Crain, and Fodor (in press) and Stevenson and Pollitt (1987). Gorrell *et al.* showed that preschool children benefit from choosing a toy in advance only if the choice is incorporated into the subordinate clause of the test sentence; the children do not benefit from having their choice incorporated into the main clause of the test sentence. This was expected, since only additional information about the subordinate clause serves the purpose of satisfying the presupposition inherent in sentences containing temporal terms. Further, Stevenson and Pollitt found that children as young as 3;4–3;9 act out sentences containing temporal terms in the appropriate sequential order, under conditions in which the child is only required to act out the main clause while the experimenter performs the action in the subordinate clause.

this as a sequencing problem (Crain, 1987). Sentences of this type place heavy demands on working memory because one of the clauses must be held in memory while a plan is formulated for acting out the sentence in the correct conceptual order (Hamburger and Crain, 1984). A second prediction of the Processing Limitation Hypothesis is that poor readers should have less difficulty when contextual support is provided, in keeping with Crain's (1982) findings with preschoolers. In contrast, the Structural Lag Hypothesis predicts that poor readers should have difficulty comprehending sentences containing temporal terms even when the appropriate contextual support is provided, because these sentences are structurally complex. That is, if poor readers are lacking the structural knowledge to deal with complex constructions containing temporal terms, their inability to comprehend them should be apparent in any context. Poor readers may adopt an order of mention strategy and give correct responses to sentences containing temporal terms that do not present a sequencing problem, but they should do less well than good readers on sentences which present a sequencing problem.

According to the Processing Limitation Hypothesis, one type of sentence containing temporal terms should be the most difficult for poor readers: Sentences like (4) pose a sequencing problem and, in addition, the subordinate clause appears second.

(4) Move the truck after you move the horse.

Such sentences do not indicate to the actor that execution of the main clause should be forestalled until the second clause is received. Reanalysis is required in order to arrive at the correct execution sequence, a fact which makes it likely that this sentence type places the heaviest demands on working memory of any of the sentences we have considered. In line with this argument, it is hypothesized that sentences like (4) should be more difficult to process than sentences like (1), which also pose a sequencing problem but in which the use of *before* in the initial clause signals that execution should be forestalled until the second clause is received. This prediction, too, is tested in the experiments that follow.

EXPERIMENT 1

The purpose of Experiment 1 was to obtain a baseline level of performance for second-grade good and poor readers with sentences containing the temporal terms *before* and *after*. This experiment, like most studies in the language acquisition literature, asked subjects to act out sentences containing temporal terms without controlling for the presupposition associated with these sentences. We wanted to determine the extent to which good and poor readers are able to comprehend constructions incorporating temporal terms despite missing contextual support.

Method

Subjects

The subjects were second-graders in a suburban public school. Each child initially received the Decoding Skills Test (DST) (Richardson and DiBenedetto, 1986) as a measure of reading ability and the Peabody Picture Vocabulary Test – Revised (PPVT) (Dunn and Dunn, 1981). The PPVT was given to ensure that all subjects assigned to a reader group had normal vocabulary knowledge, which we used as an indicator of normal IQ.

The good reader group comprised 14 children (ten girls, four boys) who obtained a DST score greater than 73 (mean 92.9, *SD* 13.9) and were assigned to the top reading level in their class by their reading instructor. The poor reader group comprised 14 children (eight girls, six boys) with DST scores less than 47 (mean 23.7, *SD* 14.9) and were placed at lower reading levels in their class. The DST score for each subject is the total number of items read correctly on a list containing 60 words and 60 pseudowords, chosen to give full representation of the syllable types that occur in the language.

The groups did not differ significantly in mean age (good readers 7; 9, poor readers 7; 8) (t (26) = 0.35, $p > 0.05$) nor in mean PPVT scores (good readers 110.6, *SD* 7.4; poor readers 105.0, *SD* 11.5) (t (26) = 1.54, $p > 0.05$).

Materials

Twelve test sentences containing temporal terms were presented. In six of these sentences there was conflict between conceptual order and order of mention (the *conflict* condition). Three of these contained *after* and three contained *before.* For the remaining six sentences, there was correspondence between conceptual order and order of mention (the *correspondence* condition). Three of these contained *after* and three contained *before.* In addition, there were two practice sentences, one containing *after* and one containing *before.* The practice and test sentences are listed in Appendix A.[4] Examples of the test sentences are given in (5)–(8):

(5) Push the smallest truck after you push the red car.

(6) Before you move the blue horse, move the smallest truck.

(7) After you push the black car, push the largest truck.

(8) Pick up the black car before you pick up the smallest horse.

All sentences were recorded on audiotape at a natural conversational rate. Each sentence was preceded by an alerting bell.

---

[4] It is important to note that these test sentences are not the least complex cases of sentences containing temporal terms, since the noun phrases in both clauses contain additional adjectival modifiers. We chose to introduce this slight added complexity in order to avoid inducing ceiling performance by all subjects.

Three sets of toy objects (cars, trucks, horses) were used in the experiment. Each set consisted of three items of the same type; members of a set varied in color (red, blue, black) and in size (small, medium, large), such that any particular object could be designated either by color or by size.

Procedure

The subjects were tested individually using an object manipulation task. Each subject was initially introduced to the two objects (any truck, any horse) used in the first practice trial. The objects were placed on a table top in front of the subject. The subject was instructed to listen carefully to the entire sentence before proceeding to act it out. If any practice trial was acted out incorrectly, it was repeated, and, if still incorrect after the second try, the experimenter performed the correct action for the subject. After completion of the first practice trial, the two objects (any car, any horse) needed for the second practice trial were introduced to the subject and placed in front of him or her. Once the second practice trial was completed, all nine objects were introduced to the subject and placed in front of him or her. The objects were grouped into three sets (cars, trucks, horses) and were randomly arranged within each set.

The test sentences were then administered in a fixed, random order. After each response, the experimenter recorded which objects were manipulated and noted the order in which the actions were performed. A sentence was repeated once if the subject requested it. Prior to each test trial, the experimenter rearranged the objects, always grouping them by type but shuffling the sets and the objects within each set.

Results

A response was considered correct if both the correct objects were chosen and the actions were acted out in the appropriate conceptual order. Errors were of two types: object selection errors and ordering errors. An example of an object selection error in response to sentence (8) would be if the subject chose the middle-sized horse instead of the smallest horse. Three sets of analyses were performed: one for overall errors, one for object selection errors, and one for ordering errors.[5]

Table 1 provides error percentages for each reader group for both the *conflict* and the *correspondence* conditions, further divided into sentences containing *before* and sentences containing *after.* One-way ANOVAs using subjects as a random factor[6] indicate that the poor readers made significantly more overall errors (across all sentence types) than the good readers (12.5% vs. 4.2%) ($F(1, 26) = 4.92, p < 0.04$).[7] It may be seen from

---

[5] If a subject made both an object selection error *and* an ordering error on a trial, the subject was credited with only one overall error.

[6] Since there were so few sentences of each type, item analyses are omitted. Thus, we have not established whether the findings can be generalized to items external to these experiments.

[7] All error analyses reported in this study are based on untransformed data. In addition

TABLE 1

Experiment 1: Percent Errors in the *Conflict* and *Correspondence* Conditions for Sentences Containing *Before* and *After*

	Conflict					
	Before C1, C2			*C1 After C2*		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	7.1	0.0	7.1	21.4	7.1	14.3
Good Readers	7.1	2.4	4.8	2.4	0.0	2.4

	Correspondence					
	C1 Before C2			*After C1, C2*		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	4.8	2.4	2.4	16.7	4.8	11.9
Good Readers	0.0	0.0	0.0	7.1	4.8	2.4

the table that the good and poor readers differed chiefly in ordering errors, not in object selection errors. Nevertheless the difference between reader groups in ordering errors failed to reach significance ($F(1, 26) = 3.77, p < 0.07$). The high level of success demonstrated by the good readers meant that subsequent analyses of their error pattern would be uninteresting, so we focused only on the error pattern of the poor readers.

The factors of *conflict* vs. *correspondence* and *before* vs. *after* were evaluated in the poor readers by a set of two-way ANOVAs. There was no main effect of *conflict* vs. *correspondence.* However, the analyses showed that the poor readers made more ordering errors and more overall errors for sentences containing *after* than for sentences containing *before* ($F(1, 13) = 6.90$, $p < 0.03$; $F(1, 13) = 4.85$, $p < 0.05$). No interaction was present because of the unexpectedly high rate of errors in response to sentences with *after* in the *correspondence* condition.

---

to these analyses, parallel analyses were performed on arcsine transformed data. The results of the analyses with transformed data did not differ from those obtained with untransformed data except in one case (see footnote 9).

Discussion

The results indicate that although the poor readers erred more in acting out sentences containing temporal terms than the good readers, they nonetheless displayed a high overall success rate. This occurred even though the test sentences contained padding in the form of added adjectival modifiers and even though the task was performed with unmet presuppositions. Only one poor reader failed to act out at least 75% of the test sentences correctly. Thus the findings demonstrate that nearly all of the poor readers possessed the necessary competence to successfully interpret temporal terms.[8] This undercuts the claim that children with specific reading disability evince a general lag in syntactic development.

Nevertheless, the results do not unequivocally support the Processing Limitation Hypothesis. Unlike the good readers, who performed well on all sentences, the poor readers had more difficulty with sentences containing *after* than with sentences containing *before,* even when the test sentences exhibited correspondence between conceptual order and order of mention. This finding is not readily explained by the Processing Limitation Hypothesis, and could be interpreted as evidence that the poor readers may be experiencing a delay in mastering the term *after.* It should be noted, however, that only one poor reader was unable to correctly act out at least two thirds of the sentences containing *after.*

There is limited support for the Processing Limitation Hypothesis in the finding that the highest percentage of errors by poor readers occurred on those sentences that we considered to be the most demanding of memory resources (namely, sentences containing *after* in which there is a conflict between conceptual order and order of mention). This result encouraged us to explore the effects of introducing more extreme variations in processing load while holding sentence structure constant. This strategy was adopted in a second experiment.

EXPERIMENT 2

The purpose of Experiment 2 was to examine the performance of good and poor readers in acting out sentences containing temporal terms when processing demands were varied. The baseline levels of performance established in Experiment 1 aided us in determining the extent to which processing demands could be manipulated such that performance would be above chance for both groups. We wished to discover whether poor readers continue to display a high percentage of correct responses in contexts that tax memory resources more severely than those in Experiment 1. If their performance

---

[8] Of course, anything substantially lower than 100% accuracy requires an explanation if the results are to be interpreted as indicating intact knowledge. In the present case, the poor readers' average success rate of 87.5% deserves comment. We take this to be evidence for competence because the sentences used in this experiment were complex cases of the structure being investigated, and no attempts were made to lessen memory demands.

under these conditions declined significantly, whereas the performance of good readers did not, then the data would lend support to the hypothesis that poor readers suffer from a limitation in processing. A widening discrepancy between reader groups as processing load is increased would defy explanation on the hypothesis that poor readers suffer from a developmental lag in the acquisition of the syntax underlying sentences containing temporal terms.

In this experiment processing demands were manipulated in both directions. In one case, we attempted to increase the processing load beyond the level imposed in Experiment 1 by adding a prenominal modifier to half of the test sentences. As exemplified in (9), these sentences contained the ordinal *second.* This addition increases the complexity of the plan that a listener must formulate in order to respond accurately to the noun phrase in which the ordinal appears (see Hamburger and Crain, 1984, for details) and thus should increase the likelihood that a subject would make object selection (and, possibly, ordering) errors.

(9) Push the second smallest horse before you push the blue car.

In the second case (which we call the *context* condition), we reduced the burden imposed on memory by making the context felicitous, i.e., by satisfying the presupposition associated with sentences containing temporal terms. In order to satisfy the presupposition, we followed the procedure employed by Crain (1982), which requires subjects to establish in advance their intent to perform the action mentioned in the clause introduced by the temporal term. In the *no context* condition, the presupposition inherent in the use of temporal terms was not satisfied; sentences were presented without context as in Experiment 1. As indicated earlier, in the absence of context any unmet presuppositions must be incorporated into the listener's mental model of the discourse setting, a procedure which is likely to tax working memory. In light of these considerations, the Processing Limitation Hypothesis anticipates a high rate of successful comprehension by both reader groups when the presupposition associated with temporal terms is satisfied, and, further, it predicts that poor readers' (but not good readers') performance will suffer in contexts that tax working memory, i.e., when the presupposition is unmet.

It should be apparent that if poor readers display a sufficiently high level of correct performance in any condition, this outcome would be inconsistent with the hypothesis that these children have not yet acquired the structural knowledge necessary to comprehend sentences containing temporal terms. In addition, a reduction in errors – either of object selection or ordering – in felicitous contexts would lend credence to a processing explanation of their performance failures in less than optimal contexts.

Method

Subjects

The subjects were second-graders who met the same selection criteria as those who participated in Experiment 1. Each child initially received the DST and the PPVT. The good reader group comprised 14 children (five girls, nine boys) who obtained a DST score greater than 69 (mean 97.2, *SD* 16.1) and were assigned to the top reading level in their class. The poor reader group comprised 14 children (four girls, ten boys) with DST scores less than 56 (mean 37.9, *SD* 12.0) and were placed at lower reading levels in their class.

The groups did not differ significantly in mean age (good readers 7; 8, poor readers 7; 11) ($t\ (26) = 1.44$, $p > 0.05$) nor in mean PPVT scores (good readers 115.4, *SD* 11.1; poor readers 108.9, *SD* 9.4) ($t\ (26) = 1.67$, $p > 0.05$).

Materials

For each condition (*context, no context*) there were 16 test sentences. In twelve of these sentences, there was a conflict between the conceptual order and the order of mention: Six of these contained *after* and six contained *before.* In the remaining four sentences conceptual order and order of mention corresponded: Two of these contained *after* and two contained *before.* Examples of each sentence type for the *no context* condition are given in (10)–(13).

(10) Pick up the second largest truck after you pick up the blue horse.

(11) Before you move the red car, move the smallest horse.

(12) After you push the black horse, push the largest truck.

(13) Pick up the blue truck before you pick up the second smallest car.

Half of the sentences of each type contained the ordinal *second* in one of the clauses, as exemplified in (10) and (13), and the other half omitted this extra modifier. The ordinal modified "largest" and "smallest" an equal number of times. We refer to the sentences with *second* as containing *complex NP*s (noun phrases) and the ones without as containing *simple NP*s.

The 16 sentences used in the *context* condition were nearly identical to those used in the *no context* condition except that in the *context* condition, the color adjective was relegated to the subordinate clause in all instances. In addition, for the *context* condition the noun in the main clause and the color adjective and noun in the subordinate clause were left unspecified, since these were determined by the subject's choice on each trial.

The test sentences followed five practice sentences for each condition. All sentences were read by the experimenter during the testing session at a natural conversational rate. The practice and test sentences for both conditions are listed in Appendix B. The nine objects from Experiment 1 were also used in each condition in this experiment.

Procedure

Subjects were tested individually in two sessions. Half of the subjects in each group received the *no context* condition in the first session and the *context* condition in the second session, and the other half received the conditions in the reverse order. There was a minimum of seven days between testing sessions.

No context. The procedure was identical to that followed in Experiment 1 except that there were now five practice sentences and 16 test sentences. Before hearing the first practice sentence, the subject was shown the three cars; before the second, the three trucks; and before the third, the three horses. For the last two practice sentences, only the two objects required for each sentence were placed in front of the subject. When the five practice trials were completed, all nine objects were placed in front of the subject. The test sentences were then administered in the same manner as in Experiment 1.

Context. The first three practice sentences were tested in the same way as the *no context* condition except that the objects used for each sentence differed across conditions. For the fourth practice sentence, a truck and a horse were placed in front of the subject. The subject was asked to point to the toy that he or she would like to pick up. The experimenter incorporated the object's name into the subordinate clause of the sentence and the remaining object's name into the main clause (as described in the next paragraph). The fifth practice sentence was administered in the same fashion as the fourth, using two different objects (a car, a horse). Once the five practice sentences were completed, all nine objects were placed in front of the subject and the test sentences were administered.

Prior to each test sentence the subject was asked to point to one of the nine objects he or she would like to move (pick up, or push). The experimenter subsequently incorporated this information into the subordinate clause introduced by the temporal term and made the entire sentence conform to this choice. For example, sentence (10) would have been read as it stands if the subject chose to pick up the *blue horse.* However, the same sentence might have been read: "Pick up the second largest horse after you pick up the blue truck" if the subject had chosen the *blue truck* instead. Before reading each test sentence, the experimenter noted the subject's color choice, incorporated it into the subordinate clause, and chose a different type of object to be described in the main clause.

The test sentences for both the *no context* condition and the *context* condition were given in a fixed, random order. The test orders were different for the two conditions. The procedure described in Experiment 1 for recording responses, repeating sentences, and rearranging the toys after each test sentence was followed in both conditions of this experiment.

Results

One-way ANOVAs indicate that, when errors were summed over conditions and sentence types, the poor readers made significantly more overall errors, more object selection errors, and more ordering errors than the good readers (poor readers 24.8%, 12.1%, 13.9%; good readers 8.1%, 4.7%, 3.6%) ($F(1, 26) = 14.21$, $p < 0.01$; $F(1, 26) = 7.98$, $p < 0.01$; $F(1, 26) = 8.08$, $p < 0.01$). Further analyses were performed to determine the extent to which varying processing demands affected the performance of each reader group.

TABLE 2

Experiment 2: Percent Errors in Each Context Condition for Sentences Containing *Simple* and *Complex NP*s

	No context					
	*Simple NP*s			*Complex NP*s		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	26.8	18.8	9.8	36.6	18.8	18.8
Good Readers	8.0	6.3	2.8	10.7	7.1	4.5

	Context					
	*Simple NP*s			*Complex NP*s		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	16.1	3.6	13.4	19.6	8.9	11.6
Good Readers	5.4	2.7	2.7	8.0	2.7	5.4

Table 2 provides error percentages for each reader group in the *no context* and *context* conditions for sentences containing *simple* and *complex NP*s. Two-way ANOVAs with *complexity* and *context* as within subject factors show that the poor readers made more overall errors when the test sentences contained *complex NP*s than when they contained *simple NP*s ($F(1, 13) = 5.37$, $p < 0.04$). In addition, the poor readers made more object selection errors and more overall errors in the *no context* condition than in the *context* condition ($F(1, 13) = 30.33$, $p < 0.01$; $F(1, 13) = 7.06$, $p < 0.02$).[9] The good readers, on the other hand, were unaffected by variations in *complexity* and *context.* These findings clearly indicate that, unlike the good readers, the poor readers' level of success in acting out sentences containing temporal terms is influenced by changes in processing demands.

Tables 3 and 4 present the results of *conflict* and *correspondence* in the *no context*

---

9 The analysis of ordering errors using arcsine transformed data revealed a significant *complexity* × *context* interaction ($F(1, 13) = 5.96$, $p < 0.03$) which failed to reach significance with untransformed data ($F(1, 13) = 4.50$, $p < 0.06$). The poor readers made the most ordering errors when sentences containing *complex NP*s were presented in the *no context* condition.

TABLE 3

Experiment 2: Percent Errors in the *No Context* Condition for Sentences Containing *Before* and *After* with *Conflict* and *Correspondence*

	Conflict					
	Before C1, C2			*C1 After C2*		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	19.0	14.3	4.8	42.9	20.2	25.0
Good Readers	14.3	11.9	3.6	8.3	7.1	1.2

	Correspondence					
	C1 Before C2			*After C1, C2*		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	32.1	32.1	3.6	35.7	14.3	21.4
Good Readers	0.0	0.0	0.0	7.1	0.0	7.1

and *context* conditions, respectively. Given that only two sentences contributed to the mean for each sentence type in the *correspondence* condition, we restricted our analysis to examining the effect of *before* vs. *after* for the *conflict* condition alone, and for the *conflict* and *correspondence* conditions taken together. Each analysis is based on difference scores for each subject using paired t-tests to determine whether the mean difference score is significantly different from zero.

In the *no context* condition the poor readers made the most errors on sentences containing *after* when *conflict* was present. They made significantly more ordering errors and more overall errors for these sentences than for sentences containing *before* ($t\,(13) = 2.73, p < 0.02$; $t\,(13) = 3.15, p < 0.01$). However it is also true that the poor readers made significantly more ordering errors and more overall errors for the entire set of sentences containing *after* than for the set of sentences containing *before* ($t\,(13) = 3.15, p < 0.01$; $t\,(13) = 3.22, p < 0.01$). The good readers' error percentages were not significantly different across sentence type.

Unlike in the *no context* condition, in the *context* condition the poor readers failed to show any significant differences in performance across sentence type.[10] The good

[10] Although the poor readers' percentages of ordering errors and overall errors were

TABLE 4

Experiment 2: Percent Errors in the *Context* Condition for Sentences Containing *Before* and *After* with *Conflict* and *Correspondence*

	Conflict					
	Before C1, C2			*C1 After C2*		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	8.3	6.0	3.6	25.0	2.4	23.8
Good Readers	0.0	0.0	0.0	14.3	6.0	8.3
	Correspondence					
	C1 Before C2			*After C1, C2*		
	overall	selection	ordering	overall	selection	ordering
Poor Readers	14.3	7.1	7.1	28.6	10.7	17.9
Good Readers	0.0	0.0	0.0	10.7	3.6	7.1

readers, however, did show some variations in performance despite their high level of correct responses. Sentences containing *after* when *conflict* was present produced more ordering errors and more overall errors than sentences containing *before* (t (13) = 2.19, $p < 0.05$; t (13) = 2.75, $p < 0.02$). In addition, the good readers made more ordering errors and more overall errors for the entire set of sentences containing *after* than for the set of sentences containing *before* (t (13) = 2.39, $p < 0.04$; t (13) = 2.90, $p < 0.02$). This finding is reminiscent of the results with the poor readers in Experiment 1.

We now examine in detail the error percentages for the sentences in which conceptual order and order of mention correspond. As noted, there were just four of these sentences in each context condition: two containing *after* (one with a *complex NP*) and two containing *before* (one with a *complex NP*). In the *no context* condition, the poor readers made 46.4% overall errors (seven selection, seven ordering) for sentences

---

higher for sentences containing *after* than for sentences containing *before* in the *context* condition, the differences were not significant. They mainly reflect the performance of two subjects who contributed to the high degree of variability in the poor reader group.

containing *complex NPs*, while they made only 21.4% overall errors (six selection, no ordering) for sentences containing *simple NPs*. In the *context* condition, the poor readers made 28.6% overall errors (four selection, four ordering) for sentences containing *complex NPs*, and only 14.3% overall errors (one selection, four ordering) for sentences containing *simple NPs*. Thus, we see that the relatively high error percentages for the poor readers in the *correspondence* condition are mainly due to the fact that one sentence of each type contained a *complex NP*.

Discussion

The fact that the poor readers were able to act out sentences containing temporal terms with only 68.3% accuracy in the *no context* condition, but were 82.2% accurate in the *context* condition confirms earlier indications that satisfying the presupposition associated with these sentences reduces the processing load which is encountered in the absence of context (Hamburger and Crain, 1982). Unlike the poor readers, the good readers were unaffected by changes in contextual support; they acted out the test sentences with greater than 90% accuracy in both conditions. Thus, a processing explanation gives a tenable account of the reader groups' pattern of performance across the *no context* and *context* conditions. The Structural Lag Hypothesis, on the other hand, is at a loss to explain why the poor readers were able to act out sentences containing temporal terms (with *simple NPs*) with a mean accuracy of 83.9% in the *context* condition. (Only two poor readers acted out these sentences with less than 75% accuracy.)

It is apparent that the poor readers' 83.9% success rate in the *context* condition is below the 87.5% success rate found in Experiment 1. However, the sentence materials were by design more difficult in Experiment 2: 75% of the sentences presented a sequencing problem while only 50% did so in Experiment 1.[11] This difference is reflected in the performance of the good readers also. For sentences containing *simple NPs*, the good readers were 92.0% accurate in the *no context* condition and 94.6% accurate in the *context* condition as opposed to an accuracy level of 95.8% in Experiment 1.

The significant effect of *NP complexity* can be interpreted as a result of processing difficulties. For example, the difficulties that the poor readers displayed with sentences containing complex noun phrases like "second largest horse" may be attributable to the circumstance that the plan required to choose the correct object cannot be compiled in a linear fashion. Phrases such as these present an example of what Hamburger and Crain (1984) call a "compiling discontinuity". It is reasonable to suppose that compiling discontinuities of this kind tax memory resources (see Hamburger and Crain, 1984, for discussion of compiling discontinuities in planning).

The Processing Limitation Hypothesis predicted that sentences that present a conflict

---

[11] Further, there were four more sentences in each condition of Experiment 2 than in the one condition of Experiment 1. These additional sentences, together with the intermixing of sentences containing and not containing *complex NPs*, made the task on the whole more burdensome than the task in Experiment 1.

between conceptual order and order of mention and also contain the temporal term *after* would be the most difficult for the poor readers. However, in both Experiments 1 and 2 the poor readers had more difficulty acting out all sentences containing *after* than sentences containing *before,* even those that do not require mental reversal. It is conceivable that the poor readers' difficulties with sentences containing *after* reflect a general failure to interpret this temporal term appropriately in the context of the experimental task. However, we would prefer to draw a different conclusion, based on the fact that most of the poor readers in this experiment (11 out of 14) were able to act out sentences containing *after* with 75% accuracy or better in the *context* condition, even when half of the sentences contained a *complex NP.* (Two of the remaining three poor readers were able to act out only one of the eight sentences containing *after* correctly, while the other poor reader performed at 50% accuracy.) Thus, with some exceptions, the poor readers demonstrated an ability to comprehend sentences containing *after.* In this connection, it is notable that the good readers also had more difficulty with sentences containing *after* than with sentences containing *before* in the *context* condition. But, clearly, their high overall level of correct performance across conditions with sentences containing *after* (89.3%) indicates competence.

Analyses which considered object selection errors and ordering errors separately yielded three facts of note. First, the results from both experiments showed that in the absence of context the poor readers were more prone to make ordering errors for sentences containing *after* than for sentences containing *before.* Apparently, when the presupposition associated with sentences containing *after* was left unsatisfied, the poor readers were sometimes unsure about the correct conceptual ordering of events. Although this suggests that the poor readers suffer from an incomplete understanding of *after,* the prevalence of these errors in the *no context* condition leads us to suspect, as suggested earlier, that their understanding of *after* was often masked by presenting these sentences in infelicitous contexts.[12] Second, the poor readers' higher percentage of object selection errors in the *no context* condition in comparison with the *context* condition indicates that when the presupposition was left unsatisfied, the poor readers had more difficulty identifying the objects to be placed in the two clauses than when the subject was allowed to choose the object for the subordinate clause. Third, the fact that the poor readers' percentage of overall errors, but not their percentage of object selection errors, was significantly greater for sentences containing *complex NPs* than for sentences containing *simple NPs* suggests that the addition of the extra adjectival modifier resulted in a general decrement in performance as anticipated by the Processing Limitation Hypothesis.

---

[12] The poor readers also made more ordering errors for sentences containing *after* than for sentences containing *before* in the *context* condition. However, this difference was not significant (see footnote 10).

General Discussion

Taken together, the findings of Experiments 1 and 2 show that under conditions of reduced processing demands, most of the poor readers in this study, like the good readers, were able to act out sentences containing temporal terms with a great deal of success. Further, the results of Experiment 2 indicate that as processing demands were increased, the poor readers' performance was eroded, whereas the good readers' performance remained unaffected. These results clearly conform to the predictions of the Processing Limitation Hypothesis. They are difficult to square with any general claim that young poor readers are delayed in the acquisition of complex structural knowledge. In the absence of structural knowledge, there would be no reason to expect the poor readers' performance to vary substantially under conditions of different processing demands. The fact that such variation was observed is support for the view that the majority of the poor readers in this study were not lacking the structural knowledge required to comprehend sentences containing temporal terms. Extrapolating from the findings in the present study, we expect that if an experiment were conducted including only *simple NP*s (as in Experiment 1) but with the presupposition associated with sentences containing temporal terms satisfied (as in the *context* condition of Experiment 2), we would find, as Crain (1982) found with preschool children, near-perfect performance by both reader groups.

The findings invite us to re-examine previous results that have been invoked in support of the Structural Lag Hypothesis. Let us consider, first, the results of several studies showing that poor readers tend to make more errors in comprehending relative clauses than good readers (Byrne, 1981; Goldsmith, 1980; Mann, Shankweiler, and Smith, 1984; Stein *et al.*, 1984). Although this finding is consistent with the Structural Lag Hypothesis, further questions need to be answered before a definite conclusion can be reached. For example, Mann *et al.* (1984) varied the type of relative clause presented to good and poor readers and found that although the poor readers made more mistakes in acting out relative clause sentences than the good readers, both groups were similarly affected by variation of relative clause type. That is, the poor readers performed uniformly worse than the good readers on all types of relative clauses. Since the poor readers also made significantly more errors in sentence repetition than the good readers, Mann *et al.* suggested that the difference in comprehension between reader groups could be explained in terms of verbal short-term memory capacity instead of a lack of structural knowledge on the part of the poor readers.

The tentative conclusion proposed by Mann *et al.* was bolstered in a study by Smith *et al.* (in press). Here additional steps were taken to positively distinguish between the two hypotheses. Smith *et al.* tested good and poor readers' comprehension of relative clauses using both an object manipulation task and a sentence-picture matching task. The critical difference between the Mann *et al.* and Smith *et al.* studies is that the latter attempted to ease processing demands to assess poor readers' level of grammatical competence in dealing with relative clause sentences. The methodological changes consisted of satisfying the presuppositions inherent in the use of relative clauses. Smith *et al.* found that the good and poor readers in their study made few errors on either task, and that both groups outperformed their (older) counterparts in the Mann *et al.* study. This led Smith *et al.* to conclude that the higher error rates found in earlier studies were mainly due to greater processing demands, rather than to a structural lag on the part of the poor readers. The findings from the object manipulation task of Smith *et al.* have been replicated as part of a larger study by Crain, Shankweiler, Macaruso, and Bar-Shalom (in press).

Differences between good and poor readers were also found in comprehension of sentences containing adjectives like *easy* (Byrne, 1981). Byrne discovered that the poor readers in his study, more than the good readers, tended to act out sentences such as "The bear is easy to reach" by having the bear reach for the rabbit rather than having the rabbit reach for the bear. The poor readers appeared to be treating the adjective *easy* as though it had the same structural properties as adjectives like *eager.* For sentences containing *eager,* such as "The bear is eager to reach", it is appropriate for the bear to reach for the rabbit. Byrne interpreted the poor readers' difficulties with adjectives like *easy* as evidence in favor of the Structural Lag Hypothesis. Using a sentence-picture matching task, a study was conducted to discover whether young poor readers' incorrect performance on sentences containing adjectives like *easy* would improve when processing demands were reduced (reported in Crain, 1987). Pictures were constructed to satisfy the presuppositions associated with the adjectives in question. For example, the correct picture for "The bear is easy to reach" was drawn such that the ease of reaching the bear was clearly indicated. In this task both the good and poor readers made virtually no errors. This suggests that the failure of the poor readers to correctly act out sentences containing adjectives like *easy* in the Byrne study was due not to a structural lag, but instead to the processing difficulties presented by unmet presuppositions. Byrne's task would not have established in advance that one of the animals was easy to reach.

The Processing Limitation Hypothesis is based on a modular view of language processing which proposes that low-level processors must rapidly feed their outputs to higher-level processors in a bottom-up fashion (Forster, 1979). Given this cognitive architecture, any obstruction in the system at a lower level would constrict information flow to higher levels. Just such a constriction appears to exist in poor readers at the level of phonology. Much research indicates that poor readers have a variety of problems in phonological processing that extend to spoken language. It is known, for example, that poor readers are less proficient than good readers in consciously manipulating the internal structure of spoken words, as required in syllable and phoneme segmentation tasks (Bradley and Bryant, 1983; Liberman, Shankweiler, Fischer, and Carter, 1974). As a consequence, poor readers fail to develop efficient word decoding skills which depend on identifying the internal components of words. Further evidence for the phonologic basis of the problems of poor readers comes from studies which show poor short-term recall of verbal material (Liberman, Mann, Shankweiler, and Werfelman, 1982; Mann, Liberman, and Shankweiler, 1980; Smith, Mann, and Shankweiler, 1986) and, in some cases, a deviant pattern of recall errors. There is evidence that these memory limitations are associated with wider phonological deficits in perception and production (Brady, 1986).

According to the Processing Limitation Hypothesis, poor readers' difficulties in phonological processing may impede the transfer of linguistic material upwards through the language processing system to higher-level processors. In other words, poor readers suffer from what Perfetti (1985) terms a "bottleneck" in the flow of information from lower-level processors to higher-level processors such as the syntactic parser. It is expected, then, that poor readers' deficits in phonological processing should give rise to problems in spoken sentence comprehension. In particular, their limitations in phonological processing should manifest themselves with sentences that tax working memory resources.

However, both good and poor readers should have little difficulty comprehending spoken sentences that do not place heavy demands on working memory regardless of their structural complexity (Crain and Shankweiler, 1988). More importantly, both groups should show successful comprehension of sentences which ordinarily pose problems when these sentences are presented in contexts that minimize processing load. Elsewhere we have elaborated a theory of working memory which generates predictions about specific sentence structures that poor readers and others with a phonological deficit should find difficult to process (Crain *et al.*, in press; Shankweiler and Crain, 1986).

The main findings of the present study are consistent with the predictions of the Processing Limitation Hypothesis. The poor readers had the greatest difficulty comprehending those sentences containing temporal terms that were the most taxing of working memory. But with memory demands lightened by satisfying the presupposition associated with sentences containing temporal terms, the poor readers performed with a high degree of accuracy. In sum, evidence from a variety of sources, including the two experiments described in this paper, calls into question the claim that young poor readers' comprehension problems with spoken sentences reflect a delay in the acquisition of structural knowledge. Instead, the present findings suggest that by the time most young children begin to learn to read, their syntactic knowledge is already in place; the special problems young poor readers face in comprehending certain spoken sentences most often stem from processing limitations arising from their phonological deficits.

(Received June 28, 1988; accepted April 4, 1989)

REFERENCES

AMIDON, A., and CAREY, P. (1972). Why five-year-olds cannot understand *before* and *after*. *Journal of Verbal Learning and Verbal Behavior*, **11**, 417–423.

BRADLEY, L., and BRYANT, P. (1983). Difficulties in auditory organization as a possible cause of reading backwards. *Nature*, **271**, 746–747.

BRADY, S. (1986). Short-term memory, phonological processing and reading ability. *Annals of Dyslexia*, **36**, 138–153.

BYRNE, B. (1981). Deficient syntactic control in poor readers: Is a weak phonetic memory code responsible? *Applied Psycholinguistics*, **2**, 201–212.

CLARK, E.V. (1970). How young children describe events in time. In G.B. Flores d'Arcais and W.J.M. Levelt (eds.), *Advances in Psycholinguistics* (pp. 275–284). Amsterdam: North Holland.

CLARK, E.V. (1971). On the acquisition of the meaning of *before* and *after*. *Journal of Verbal Learning and Verbal Behavior*, **10**, 266–275.

CRAIN, S. (1982). Temporal terms: Mastery by age five. *Papers and Reports on Child Development* (Stanford University), **21**, 33–38.

CRAIN, S. (1987). On performability: Structure and process in language understanding. *Clinical Linguistics and Phonetics*, **1**, 1–18.

CRAIN, S., and SHANKWEILER, D. (1988). Syntactic complexity and reading acquisition. In A. Davidson and G.M. Green (eds.), *Linguistic Complexity and Text Comprehension: Readability Issues Reconsidered* (pp. 167–192). Hillsdale, NJ: Erlbaum.

CRAIN, S., SHANKWEILER, D., MACARUSO, P., and BAR-SHALOM, E. (in press). Working memory and sentence comprehension: Investigations of children with reading disorder. In G. Vallar and T. Shallice (eds.), *Impairments of Short-Term Memory*. Cambridge, U.K.: Cambridge University Press.

CRAIN, S., and STEEDMAN, M. (1985). On not being led up the garden path: The use of context by the syntactic processor. In D.R. Dowty, L. Karttunen, and A. Zwicky (eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives* (pp. 320–358). New York: Cambridge University Press.

DUNN, L.M., and DUNN, L.M. (1981). *The Peabody Picture Vocabulary Test – Revised*. Circle Pines, MN: American Guidance Service.

FORSTER, K. (1979). Levels of processing and the structure of the language processor. In W.E. Cooper and E. Walker (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett* (pp. 27–86). Hillsdale, NJ: Erlbaum.

GOLDSMITH, S. (1980). *Psycholinguistic Bases of Reading Disability: A Study in Sentence Comprehension*. Doctoral dissertation, The City University of New York.

GORRELL, P., CRAIN, S., and FODOR, J.D. (in press). Contextual information and temporal terms. *Journal of Child Language*.

HAMBURGER, H., and CRAIN, S. (1982). Relative acquisition. In S. Kuczaj, II (ed.), *Language Development, Volume 1: Syntax and Semantics* (pp. 245–274). Hillsdale, NJ: Erlbaum.

HAMBURGER, H., and CRAIN, S. (1984). Acquisition of cognitive compiling. *Cognition*, **17**, 85–136.

JOHNSON, M.L. (1975). The meaning of *before* and *after* for pre-school children. *Journal of Experimental Child Psychology*, **19**, 88–99.

LEWIS, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, **8**, 339–359.

LIBERMAN, I.Y., MANN, V.A., SHANKWEILER, D., and WERFELMAN, M. (1982). Children's memory for recurring linguistic and nonlinguistic material in relation to reading ability. *Cortex*, **18**, 367–375.

LIBERMAN, I.Y., SHANKWEILER, D., FISCHER, F.W., and CARTER, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, **18**, 201–212.

MANN, V.A., LIBERMAN, I.Y., and SHANKWEILER, D. (1980). Children's memory for sentences and word strings in relation to reading ability. *Memory and Cognition*, **8**, 329–335.

MANN, V.A,, SHANKWEILER, D., and SMITH, S.T. (1984). The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language*, **11**, 627–643.

PERFETTI, C.A. (1985). *Reading Ability*. New York, NY: Oxford University Press.

RICHARDSON, E., and DIBENEDETTO, B. (1986). *Decoding Skills Test*. Parkton, MD: York Press.

SHANKWEILER, D., and CRAIN, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, **24**, 139–168.

SMITH, S.T., MACARUSO, P., SHANKWEILER, D., and CRAIN, S. (in press). Syntactic comprehension in young poor readers. *Applied Psycholinguistics*.

SMITH, S.T., MANN, V.A., and SHANKWEILER, D. (1986). Spoken sentence comprehension by good and poor readers: A study with the Token Test. *Cortex*, **22**, 627–632.

STEIN, C.L., CAIRNS, H.S., and ZURIF, E.B. (1984). Sentence comprehension limitations related

to syntactic deficits in reading-disabled children. *Applied Psycholinguistics*, **5**, 305–322.
STEVENSON, R.J., and POLLITT, C. (1987). The acquisition of temporal terms. *Journal of Child Language*, **14**, 533–545.

APPENDIX A

Test Sentences for Experiment 1

Conflict Condition

Before you move the blue horse, move the smallest truck.

Before you push the smallest car, push the red truck.

Before you pick up the largest horse, pick up the red car.

Push the smallest truck after you push the red car.

Pick up the largest car after you pick up the red horse.

Move the blue truck after you move the largest horse.

Correspondence Condition

After you push the black car, push the largest truck.

After you move the largest truck, move the blue horse.

After you pick up the blue horse, pick up the smallest car.

Push the black truck before you push the largest car.

Pick up the black car before you pick up the smallest horse.

Move the smallest horse before you move the black truck.

Practice Sentences

After you pick up the truck, pick up the horse.

Before you move the car, move the horse.

[continued over

to syntactic deficits in reading-disabled children. *Applied Psycholinguistics*, **5**, 305–322.
STEVENSON, R.J., and POLLITT, C. (1987). The acquisition of temporal terms. *Journal of Child Language*, **14**, 533–545.

APPENDIX A

Test Sentences for Experiment 1

Conflict Condition

Before you move the blue horse, move the smallest truck.

Before you push the smallest car, push the red truck.

Before you pick up the largest horse, pick up the red car.

Push the smallest truck after you push the red car.

Pick up the largest car after you pick up the red horse.

Move the blue truck after you move the largest horse.

Correspondence Condition

After you push the black car, push the largest truck.

After you move the largest truck, move the blue horse.

After you pick up the blue horse, pick up the smallest car.

Push the black truck before you push the largest car.

Pick up the black car before you pick up the smallest horse.

Move the smallest horse before you move the black truck.

Practice Sentences

After you pick up the truck, pick up the horse.

Before you move the car, move the horse.

[continued over

APPENDIX B

Test Sentences for Experiment 2

NO CONTEXT CONDITION

Conflict Condition

Before you push the red car, push the smallest horse.

Before you push the red horse, push the second largest truck.

Before you move the largest truck, move the black car.

Before you pick up the second smallest car, pick up the red horse.

Before you pick up the blue horse, pick up the smallest truck.

Before you move the second largest car, move the blue truck.

Move the black truck after you move the largest car.

Push the second largest horse after you push the blue car.

Pick up the smallest truck after you pick up the blue horse.

Move the black car after you move the second smallest truck.

Push the red horse after you push the largest car.

Pick up the blue truck after you pick up the second smallest horse.

Correspondence Condition

After you move the black car, move the largest truck.

After you push the second smallest horse, push the blue car.

Pick up the blue truck before you pick up the largest horse.

Move the black horse before you move the second largest car.

Practice Sentences

Push the largest car.

Move the second largest truck.

Pick up the second smallest horse.

After you pick up the truck, pick up the horse.

Before you move the car, move the horse.

CONTEXT CONDITION

Conflict Condition

Before you push the (color choice), push the smallest (other).
Before you push the (color choice), push the second largest (other).
Before you move the (color choice), move the largest (other).
Before you pick up the (color choice), pick up the second smallest (other).
Before you pick up the (color choice), pick up the smallest (other).
Before you move the (color choice), move the second largest (other).

Move the largest (other) after you move the (color choice).
Push the second largest (other) after you push the (color choice).
Pick up the smallest (other) after you pick up the (color choice).
Move the second smallest (other) after you move the (color choice).
Push the largest (other) after you push the (color choice).
Pick up the second smallest (other) after you pick up the (color choice).

Correspondence Condition

After you move the (color choice), move the largest (other).
After you push the (color choice), push the second smallest (other).

Pick up the largest (other) before you pick up the (color choice).
Move the second largest (other) before you move the (color choice).

Practice Sentences

Push the largest horse.
Move the second largest car.
Pick up the second smallest truck.
After you pick up the (choice), pick up the (other).
Before you move the (choice), move the (other).