Experimental Studies of Syntactic and Lexical Processes in Language Comprehension

by

Martin Sanford Chodorow

B.A. University of Texas at Austin

(1971)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF
PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 1976

Signature of Author

Department of Psychology, September 29, 1975

Certified by

Thesis Supervisor

Accepted by

Chairman, Departmental Committee
on Graduate Students
ABSTRACT

Experimental Studies of Syntactic and Lexical Processes in Language Comprehension

Martin Sanford Chodorow

Submitted to the Department of Psychology on September 29, 1975

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

A considerable body of evidence suggests that language comprehension consists of at least two partially dissociable processes, one involved with identifying individual lexical items and the other concerned with organizing words into structures. In Experiments I, II, and III, time-compressed speech was used in a paradigm adapted from Savin and Perchonock (1965) but redesigned to permit the measurement of the processing backlog that results from increasing the presentation rate of sentences. In Experiments I and II, a structural variable, the syntactic complexity of verbs, was used to test hypotheses about the listener's use of surface cues and verb subcategorization information for the prediction of subsequent syntactic structure. Results indicate that verbs which may appear in many different syntactic constructions produce larger processing backlogs than those that are more highly restricted in their number of possible environments. Surface cues, such as complementizers, reduce the processing backlog. In time-compressed materials, those substrings over which the listener is uncertain about the structural description tend to be recalled less accurately than other portions of the sentences. In Experiment III, lexically ambiguous sentences and their unambiguous controls were presented in the compression paradigm. Results were consistent with previous indications that ambiguities of word meaning must be resolved rapidly in the course of comprehension. Additional time provided at the end of a compressed sentence is not adequate compensation for the increased local processing load produced by the ambiguity. However, additional post-sentence time does allow the listener to complete much of the backlog resulting from verb complexity.

Thesis Supervisor: Merrill F. Garrett

Title: Associate Professor of Psychology
Table of Contents

Abstract 2
Biographical Note 4
Acknowledgements 5
Introduction 6
Experiment I 12
  Method 14
  Results 19
  Discussion 20
Experiment II 25
  Method 25
  Results 30
  Discussion 33
Experiment III 39
  Method 41
  Results 43
  Discussion 45
Footnote 48
Tables 49
Figures 54
Appendices 60
References 67
Martin Sanford Chodorow was born on September 25, 1949, in Waco, Texas. He attended the Waco Public Schools and in the fall of 1967, entered the University of Texas at Austin. In the spring of 1971, the author received a Bachelor of Arts degree in Plan II (a liberal arts honors program). In the fall of the same year, he entered the doctoral program in psychology at MIT.
I wish to express my appreciation to a number of people who have aided me in the development of this thesis. The members of my committee, Jerry Fodor, Susan Carey-Block, and Molly Potter, provided invaluable assistance through their comments and suggestions during both the formative stages of this work and in the preparation of the final version of the paper. Ken Forster kindly served as a reader on very short notice and offered many helpful suggestions.

I would also like to thank Dusty Billings for scoring the sentences for analysis in Experiments II and III, and Susan Cohen Leehey, who worked under considerable time constraints to produce the final typed copy. For the preparation of my experimental materials, I am indebted to A.W.F. Huggins, who patiently instructed me in the use of his speech compression program, and to Professor Francis Lee, who kindly allowed me to use his electronic compressor-expander.

Merrill Garrett, my thesis supervisor, deserves a great deal of thanks for his encouragement, insight, and guidance at every stage of this work. His optimism and cheerfulness proved to be quite contagious, especially when I encountered what seemed to be insurmountable obstacles.

Through it all, my family continually provided the love, understanding, and encouragement that is so important to one who is engaged in an effort of this type. For them, I offer a very special note of appreciation.
INTRODUCTION

Evidence from a variety of sources suggests that language comprehension is not a unitary process but instead consists of component procedures concerned with identifying individual words and organizing words in structures. These two processes, identification and structuring, are at least partially separable on the basis of their interaction with certain experimental paradigms. The evidence comes from auditory and visual studies of verbal materials presented in sentences and in lists. In each case, those variables which affect identification have immediate, highly localized effects on processing. Structural variables tend to have more global, less immediate consequences. It is the goal of the present study to examine an identification-related variable, lexical ambiguity, and a structure related variable, verb complexity, under conditions of rapid auditory presentation. The following is a review of the evidence and arguments which support the distinction between the two kinds of processing.

Identification in sentences. Studies of the subjective lexicon (Rubenstein, Garfield, and Milliken, 1970) reveal a strong effect of word frequency and ambiguity on lexical access time, i.e., the time to identify a string of letters as a word. Foss (1969) and Foss and Jenkins (1973) have used the phoneme-monitor technique to measure the listener's processing load immediately following a low frequency or ambiguous word appearing in a sentence. Reaction times are significantly longer than those obtained when high frequency or unambiguous items are substituted in the same constructions. This difference in response latency is significant only when measurement is made immediately following the low frequency or ambiguous word, i.e., at
the beginning of the next word (Cairns, in prep.). Garrett and Holmes (Garrett, personal communication) used the rapid serial visual presentation (RSVP) technique (Forster, 1970) to present sentences containing ambiguous words or frequency-matched unambiguous control items. Overall, the total number of words recalled did not differ for the two versions of the sentences; however, a comparison between ambiguous and control items indicated that the former were recalled much less often. The phoneme-monitor and RSVP results suggest that lexical ambiguity and frequency produce a transient, highly localized processing load, one which is specific to the individual word. When these two paradigms are use to study structure-related processes, a different pattern emerges.

Structural processing in sentences. Foss (1969) used phoneme-monitoring to examine the time-course of processing in sentences containing a wide variety of common syntactic constructions. His results show faster response times for late position targets than for early targets. Holmes and Forster (1970) report the same pattern of reaction times in a click-monitor study. If monitoring data reflect processing load, then for many kinds of syntactic constructions, the load decreases as the listener moves through the string. When the structure is made extremely complex, as in the case of right-branching and self-embedded relative clauses, the load remains high throughout the sentence and may even increase toward the end (Foss and Lynch, 1969). Using the RSVP paradigm, Forster (1970) found differences in the total number of words recalled from sentences of high and low syntactic complexity. When performance was analyzed by syntactic class (noun, verb, adjective, etc.), there were no differential effects of complexity on particular word types.
(Holmes and Forster, 1972). In contrast to the results for frequency and ambiguity, the measurements on syntactic complexity reveal a much more generalized effect of differences in processing load.

**Identification and structuring in lists.** Additional support for a distinction between identification and structuring comes from the study of word list processing. In a random list composed of a small closed set of familiar words (digits, for example), the relationship between elements is defined extrinsically, by precedence, i.e., by the linear order of presentation.

Aaronson (1974) has reviewed much of the literature dealing with word list processing. Two kinds of errors can be distinguished in subjects' ordered recall of lists. One type (item error) is the failure to recall an item; the other (order error) is a failure to place the item in its correct serial position. The first has generally been interpreted as reflecting those processes underlying the identification of words in the list. The second has been viewed as a measure of the processes responsible for structuring the input. The two seem to be at least partially dissociable: (1) Aaronson (1968) reduced the interitem interval in an auditorily presented digit list and found an increase in order errors that was not accompanied by an increase in item errors. (2) In a 1971 study, Aaronson, Markowitz, and Shapiro increased the pause time between words of a list by shortening the duration of each of the digits. Under these conditions, the number of order errors was reduced, but item errors remained the same. (3) When lists are presented in noise, reduction in the signal-to-noise ratio can produce an increase in item errors without an accompanying increase in order errors.
These results suggest that the interval following a stimulus item may be crucial for its incorporation into a structural representation. On the other hand, identification of an item seems to be relatively unaffected by certain limited manipulations of word duration or interword pause length. It is the acoustic quality of the list that can affect the identification process. The evidence cited above would seem to indicate that identification is normally a rapid process, highly dependent upon the sensory quality of the signal; structuring is temporally more removed from the stimulus.

Experiments designed to assess the time-course of processing in lists (Aaronson, 1966; 1968) have revealed a rather striking non-uniformity across serial positions. For presentation rates comparable to those of conversational speech (about three words per second), the reaction time to monitor for a predesignated target item was shown to increase with serial position beyond the third item. Since random digit lists were used in the experiments, serial position could not have affected the identification process as a consequence of differences in the predictability of an item's occurrence. Aaronson interprets the results as indicating a progressive accumulation of incomplete processing through the course of the string. At slower presentation rates, i.e., with longer intervals, this effect of serial position on monitor latency disappears. Taken together with the results of Aaronson, et al. (1971), this suggests that the backlog accumulated at fast rates is a backlog of structural processing, because reaction time and order errors decrease at the slower presentation rates, but item errors do not. The relationship between list processing and sentence processing. It is
interesting to note the similarities in the distribution of processing load for lists and syntactically complex sentences. In each, the load increases through the string. For both, there is evidence to suggest that processing is completed some time after the termination of the input. Aaronson and Sternberg (1963 [cited in Aaronson, 1974]) measured the time subjects took to indicate which of two lights had been turned on following the auditory presentation of a word list. Responses were faster when the interval between list and light was longer. The experimenters interpret greater response latencies immediately following the list as evidence for interference between displaced processing and the light-monitor task. Foss and Cairns (1970) found that requiring subjects to read two words aloud before repeating the right-branching or self-embedded sentence they had just heard reduced the accuracy of their repetitions. The experimenters suggest that the effect may be the result of interrupting structural processing of the sentences, processing which has continued beyond the termination of the signal.

Sentences differ from lists primarily in terms of their intrinsic structure. The elements of a sentence can be recoded into complex units; structure may be built from the bottom up, but it can also be imposed from the top down. The information necessary for a predictive, top-down organization is not available in lists; in multiply self-embedded relatives it is present, but apparently unusable. More complex, less predictive sentence processing ought to resemble list processing in the time-course of its accumulation.

Aaronson (1968) was able to produce a backlog in structural processing by reducing the intervals between list items. If a sentence were speeded up in its rate of presentation so that the intervals following the words were re-
duced in duration, we might expect a similar buildup in processing load.
This prediction seems to be consistent with the subjective phenomenon associated with listening to time-compressed speech (Foulke, 1969). Even though the listener recognizes the words of the message as he hears them, he still experiences a feeling of "falling behind." It is tempting to suggest that a dissociation of identification and structuring is responsible for this subjective experience. Experiments I and II were designed to measure displaced processing produced by time-compressing sentences of various structural types. In Experiment III, lexical ambiguity was used to increase the load at a particular locus in a compressed string.
EXPERIMENT I

In Experiment I, test sentences which varied in terms of their syntactic form were presented at a normal rate and at a much faster than normal speed. If increasing the rate of input produces a structural processing backlog, then the observed effects should differ for the various syntactic constructions.

Previous studies of structural complexity have revealed significant effects of verb subcategorization. Fodor, Garrett, and Bever (1968) found that anagram and paraphrase tasks were easier for sentences with simple transitive verbs than for sentences containing complement verbs followed by simple direct objects. The experimenters accounted for the difference by noting that complement verbs are more complex because of the greater number of syntactic environments in which they may occur, for example, before a direct object or before an embedded complement clause. When a listener encounters a simple verb, subcategorization information is sufficient to allow for the prediction of subsequent structure; for complement verbs, it is not sufficient. Accordingly, we might expect to find greater structural complexity for all sentences containing complement verbs. Forster (1970) compared complement sentences (sentences containing complement verbs followed by complement clauses) to single-clause sentences. The RSVP technique indicated no greater complexity for the complement sentences, although it did reveal higher complexity for two-clause sentences in general. The result led Holmes and Forster (1972) to propose that processing difficulty increases for complement verbs only when they are not followed by complement clauses. To test this hypothesis, Holmes and Forster used RSVP with single-clause sen-
sentences containing simple or complement verbs and with complement sentences
of various types. The results replicated the earlier finding of Fodor, et al. (1968), i.e., performance on single-clause sentences with simple verbs was superior to that for single-clause sentences with complement verbs. The results also replicated Forster's (1970) finding of no overall difference between single-clause and complement sentences. However, an examination of particular complement types revealed very large differences among them. Verb-phrase complement sentences (e.g., The angry officials urged the man to complete the form; hereafter designated CVP) were no more difficult than single clause sentences with simple verbs (e.g., Sally injured her left elbow during a game of squash; SS). But noun-phrase complement sentences containing that complementizers (e.g., The children believed that their mother was in serious trouble; CNP-that) were more difficult than SS and CVP. In fact, the mean number of words recalled from CNP-that was about the same as the number recalled from single-clause sentences with complement verbs (e.g., The doctors announced their strong support for the health scheme; SC). Noun-phrase complement sentences containing (for)...to complementizers (e.g., The lawyer couldn't bear his old father to tell jokes; CNP-(for)...to) proved to be most difficult of all. In summary, Holmes and Forster found evidence for the following order of complexity:

\[ \text{CVP} = \text{SS} < \text{SC} = \text{CNP-that} < \text{CNP-(for)...to} \]

On the basis of the three experiments cited above (Fodor, et al., 1968; Forster, 1970; Holmes and Forster, 1972), we can make the following predictions about the outcome of Experiment I: (1) The amount of structural processing backlog should be greater for single-clause sentences with comple-
ment verbs than for those with simple verbs. (2) Overall, the backlog should be the same for single-clause sentences and complement sentences. (3) Among the complements, CNP should be more demanding than CVP in terms of processing load.

Method

Materials. The sentences used in Experiment I are a subset of those in the Holmes and Forster RSVP experiment. Holmes and Forster asked nine Australian subjects to rate the sentences of their experiment for naturalness. The ratings were extremely high for all the sentence types. Twenty American subjects, however, classified the sentences of the CNP-(for)...to group as "highly unnatural." As a result, that syntactic category was not used in Experiment I. Four sentences were chosen from each of the remaining categories (see Appendix A). Note that sentences of the SC group have been divided into those containing noun-phrase complement verbs (SNP) and those containing verb-phrase complement verbs (SVP).

A technique similar to that employed by Savin and Perchonock(1965) was adapted for use in measuring displaced processing. Subjects heard a test sentence presented at either the normal rate or at twice the normal speed. After a variable interval following the sentence, they heard a list of eight words, each of which was drawn from a different noun category, such as "animal," "furniture," "weather," etc. For each category, there were five exemplars, e.g., "cow," "dog," "horse," "sheep," "cat." The eight categories were always presented in the same serial order, and subjects were allowed
to view a listing of this order throughout the experiment. During practice trials, the subjects were familiarized with the five exemplars of each category. After the sentence and its accompanying word list had been presented, the subject's task was to paraphrase the sentence and then recall the words of the list. It was hypothesized that at the fast sentence rate and with a short pre-list interval, structural processing would continue well beyond the termination of the sentence and would, consequently, interfere with encoding the list. Since both the order of the categories and the possible items at each serial position were known, the subject's task with regard to the word list was to encode for recognition, i.e., to encode sufficiently to be able to recall which one of the five nouns from a particular category actually occurred on a given trial. There was no demand on the subject for an extrinsic ordering of the list. Interference from sentence processing ought simply to reduce the number of items encoded. If the hypothesis concerning sentence processing is correct, the number of list errors should provide a measure of the backlog that the listener is required to make up for each structural type.

Compression. Materials for Experiment I were originally recorded at a mean rate of 3.4 words (4.8 syllables) per second. Presentation rate was increased by compressing the sentences to one-half their original duration. A computer program written by A.W.F. Huggins of the MIT Research Laboratory of Electronics performed the pitch-synchronous compression by first removing every other pitch period from the voiced portion of the signal and then abutting the remaining segments. For the non-voiced portions of the signal, every other 8-millisecond segment was removed. A sentence requiring the
original duration for presentation was produced by reduplicating each pitch period in the voiced portion and each 8-millisecond segment in the non-voiced portion of the compressed tape. The time-compression of speech involved discarding some acoustic information from the waveform. By expanding the compressed version, producing an expanded-compressed sentence, it is possible to produce a signal which occupies the same amount of time as the original utterance but contains only as much information as the compressed copy. Differences in performance with the two forms cannot, therefore, be attributed to intelligibility differences arising from disparity in the quality of the stimuli.

Presentation Conditions. Four conditions of presentation were used in the experiment. In one, an expanded-compressed sentence was followed by a delay interval of 750 milliseconds. This condition (Control A) was designed to provide an interference-free measure of list performance. In pilot studies, list recall was shown to be identical for (Control A) and other control presentations using delay intervals as long as 2.25 seconds. The result suggests that sentence processing can be completed within 750 milliseconds. The other three presentation conditions all employed the compressed rate. One of these (Control B) paired a compressed sentence with a delay interval equal to 750 milliseconds plus one-half the duration of the expanded-compressed version, so that the amount of time from the beginning of the sentence to beginning of the list was the same as in (Control A). Differences in performance between the two could not be attributed to the rate factor. It was predicted that results would be the same for (Control A) and (Control B) because the interval provided in the latter was assumed to be of suffi-
cient duration to enable subjects to complete any accumulated processing. In the two remaining conditions, a compressed sentence was followed by either a 750 millisecond or a 200 millisecond interval. If sentence processing can be completed within 750 milliseconds, then list performance for (Compressed + 750) will be comparable to that of the control presentations. The effects of interference ought to show up in performance on (Compressed + 200) if processing is not completed within 200 milliseconds. Those structural forms assumed to have the largest processing backlogs were expected to display the greatest decrements in list recall. According to the RSVP results, SS and CVP should be least affected.

In summary, the presentation conditions were as follows:

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>DURATION OF SENTENCE (as % of original)</th>
<th>POST-SENTENCE INTERVAL (in milliseconds)</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control A</td>
<td>100%</td>
<td>750</td>
<td>100% + 750</td>
</tr>
<tr>
<td>Control B</td>
<td>50%</td>
<td>750 + 50% original</td>
<td>100% + 750</td>
</tr>
<tr>
<td>(Compressed + 750)</td>
<td>50%</td>
<td>750 + 50% sentence duration</td>
<td>50% + 750</td>
</tr>
<tr>
<td>(Compressed + 200)</td>
<td>50%</td>
<td>200</td>
<td>50% + 200</td>
</tr>
</tbody>
</table>

The experimental sentences of Appendix A were embedded in test sequences of the following form: Each sentence was preceded by the warning "Get ready to listen to the sentence that follows." The warning was presented at the same rate as the test sentence, which followed after a one-second delay. The eight words of the list were presented, one every three-quarter second. Average duration for each list item was 0.480 seconds. The entire sequence consisted of (1) warning, (2) one-second interval, (3) test sentence,
(4) variable delay interval, (5) eight-word list. Following the presentation of the list, subjects first wrote down a paraphrase of the sentence and then the items of the list. During the course of instruction prior to the experiment, numerous examples of paraphrases were provided for several example sentences. Subjects were told that it was particularly important for them to provide adequate paraphrases of the test sentences. Six practice trials were run before beginning the experiment.

**Scoring.** The number of list items recalled correctly was computed for each response in which the paraphrase was adequate to allow recovery of the major grammatical relations in the test sentence. When more than one major relation (subject, verb, object) was non-recoverable, the response was treated as an error, and the missing data for the cell was filled in according to an estimation procedure described by Winer (1971, p. 489). In the subject analysis (see Results), for each filled cell, a degree of freedom was subtracted from the SUBJECT X STRUCTURAL TYPE X PRESENTATION CONDITION interaction.

**Design.** Four test sequences corresponding to the four presentation conditions were prepared for each of the twenty experimental sentences. Four tapes were constructed by selecting one test sequence for each sentence. On every tape, for each structural type, there was one sentence representing each of the four conditions. The sentences were randomly arranged in blocks of five, with every block containing one of each structural type and at least one of each presentation condition. Thirty-two subjects (eight per tape) were paid for their participation in the forty-minute experiment. Subjects were run individually or in groups ranging up to five in size.
Results

Two analyses of variance were computed, (1) by collapsing across sentences and using subjects to generate the error terms, and (2) by collapsing across subjects, using sentences to generate the error terms. For the subject (error term) analysis, the design contained three factors: one random between-subjects factor, tape(4 levels); and two fixed within-subjects factors, presentation condition(4) and structural type(5). The main effect for tape was not significant, and the interactions of tape with the other experimental variables also failed to reach significant levels. Therefore, tape was eliminated as a factor for the remainder of the subject analysis. The sentence analysis was carried out on a two-factor design: one fixed between-sentences variable, structural type(5); and one fixed within-sentences variable, presentation condition(4).

Insert TABLE I about here

The main effect for presentation condition was significant in the subject analysis \(F(3,93) = 3.384, p < .05\) and in the sentence analysis \(F(4,45)^{\dagger} = 4.402, p < .05\). Two orthogonal contrasts were constructed, one (c\(_1\) for CONDITION in TABLE I) to determine if a processing backlog generally remained after the 750 millisecond interval following a compressed sentence, and the other (c\(_2\) for CONDITION in TABLE I) to compare a condition believed to produce processing interference (Compressed + 200) with the other three
presentations. The first contrast was not significant for the subject analysis or the sentence analysis but the second contrast was highly significant on both.

The main effect for structural type was not significant on either analysis. Four orthogonal contrasts were constructed to correspond to the predictions concerning the interaction between presentation condition and structural type. The F-ratios based on error terms from each analysis are given in TABLE I. For each structural type, performance was compared for (Compressed + 750) and (Compressed + 200) in order to gauge the backlog of accumulated sentence processing. The difference between the two presentation conditions was significant for single-clause sentences with complement verbs (c₁ for STRUCTURE X CONDITION in TABLE I); there was no difference for SS and CNP-that; however, the difference for CVP was significant (c₄).

Twenty-three paraphrases (3.6% of the total) were judged unacceptable. The distribution of paraphrase errors follows: by presentation condition, (Non-compressed + 750) -5, (Compressed + 750 + tₜ/2) -5, (Compressed + 750) -8, (Compressed + 200) -5; by structural type, SS -6, SNP -5, SVP -3, CNP-that -4, CVP -5.

Discussion

The main effect and orthogonal contrasts for condition of presentation confirm the general predictions of the model. There was a significant decrement in performance on the word list for (Compressed + 200). An addition of 550 milliseconds to the delay interval proved sufficient to restore
list recall to its non-interference level. The difference between (Compressed + 750) and (Compressed + 200) was not uniform across structural types. Some, though not all, of the predictions concerning the nature of this interaction were confirmed. On the basis of the Fodor, et al. (1968) experiment, single-clause sentences with complement verbs were predicted to be more complex than those with simple verbs. In the current experiment, while the short interval had no effect on SS, it reduced recall significantly for SC. This pattern suggests that a processing backlog exists 200 milliseconds after a compressed SC sentence, but if such an accumulation did occur for SS, it is no longer present. Evidence from Forster (1970) and Holmes and Forster (1972) indicates that the overall complexity of complement sentences and single-clause sentences is about the same. In the current experiment, for (Compressed + 200) the difference proved to be negligible: 4.37 and 4.40 list items were recalled with single-clause and complement sentences, respectively.

It is with regard to complement type that the results of Experiment I differ from those of the RSVP study. List recall for CNP-that and CVP was opposite to prediction, i.e., CNP-that was unaffected by the shorter delay interval, while CVP showed a significant reduction in the number of items recalled. Following the line of argument discussed earlier, the results can be interpreted as evidence for a large processing backlog associated with CVP and a much smaller one (at most, less than 200 milliseconds) for CNP-that. Holmes and Forster found CVP to be less complex than CNP-that, which was, in turn, less complex than CNP-(for)...to. To explain this pattern, the experimenters suggested that a CVP sentence can be more easily
treated as a single unit for analysis, because the object of the matrix verb also serves as the subject of the complement clause. Consequently, there is no position in the surface string which corresponds to a break between clauses. Furthermore, verb-phrase complement verbs are semantically and syntactically more constrained than their noun-phrase counterparts. For these reasons, Holmes and Forster argued, CVP is treated more like SS and less like a two-clause sentence. Among the noun-phrase complements, CNP-that is less complex than CNP-(for)...to because the former has a more explicit surface cue to its underlying structure. This explanation of the RSVP results is not entirely satisfactory. If treating a complement sentence as a single unit facilitates processing then CNP-(for)...to should benefit from the fact that the two-clause structure is less explicit than in CNP-that. If verb-phrase complement verbs are more constrained and, therefore, less difficult to process, then single-clause sentences with verb-phrase complement verbs (SVP) should be less complex than ones containing noun-phrase complement verbs (SNP). However, this was not the case in either the RSVP experiment or in Experiment I.

The outcome of Experiment I can be handled in a much less complicated fashion. In fact, with a slight modification, the proposal made by Fodor, et al. (1968) is capable of accounting for the pattern of results. As noted earlier, when a listener encounters a complement verb, he is unable to determine immediately whether it will be followed by a direct object or a complement construction. This kind of structural ambiguity remains unresolved in some instances until the end of the sentence is reached. For example, the string The doctors announced their strong support for the
The doctors announced their strong support for the health scheme could no longer be counted on by the legislators. This problem does not arise for SS because subcategorization information for the verb specifies an unambiguous environment. Complement sentences provide the listener with structural ambiguity, but it is usually resolved sooner than in the single-clause sentences containing complement verbs. In Experiment I, all noun-phrase complement sentences contained a complementizer immediately following the matrix verb. In the CVP sentences, disambiguation was not provided as rapidly because the complementizer to appeared after the object noun phrase and therefore was separated from the main verb by several words. If the ambiguity produced by verb complexity prevents the assignment of structure to subsequent elements in the sentence, then the longer this ambiguity persists, the greater the processing backlog should be. Accordingly, SS and CNP-that should have small backlogs, and SC and CVP should have much larger accumulations. In terms of the model, when the listener reaches the matrix verb in SS, and its complementizer in CNP-that, he has sufficient information to structure the input which follows. For the other sentence types, structuring must wait for disambiguation.

Clark (1973) has pointed out the dangers involved in generalizing from results of analyses in which one of two random variables, subjects or sentences has been treated as a fixed effect. He recommends the use of the minimum $F'$ statistic to assess the degree to which results may be generalized simultaneously across both subjects and materials. In Experiment I, the main effect for condition of presentation is not significant by this more conservative measure ($\min F'(3,131) = 1.91$), although the contrast between
(Compressed + 200) and the other three conditions is significant (min $F'_1(1,128) = 5.07$, $p < .05$). Of the orthogonal contrasts used to test the predictions concerning the interaction of structural type and presentation condition, only the comparison of (Compressed + 200) with (Compressed + 750) for SC remains significant (min $F'_1(1,77) = 5.08$, $p < .05$). Since minimum $F'$ is designed to measure how well results can be generalized for new subjects and new materials, actually performing another experiment using different subjects and different sentences ought to answer the same question, How replicable are the results of Experiment I?
EXPERIMENT II

The claims made concerning the time-course of sentence processing, the effects of compression, and the role of structural ambiguity can all be tested more precisely by making certain modifications in the previous design. The alterations affect three aspects of the paradigm: (1) materials, (2) presentation conditions, and (3) response task.

Method

Materials. Both the Holmes and Forster RSVP study and Experiment I suffer from a materials-related weakness in design: A random variable, sentences, was nested in a fixed treatment variable, structural type. Two sentences belonging to different structural types differed not only by the fixed variable, but also in any number of presumably random ways. The RSVP technique is known to be sensitive to the effects of sentence plausibility (Forster and Ryder, 1971) and naturalness (Holmes, 1973). Materials used in the Holmes and Forster study were equated for naturalness, although no separate measure was taken of plausibility. Word frequency, a factor which could be expected to affect the lexical identification process (Rubenstein, et al., 1970), also varied among the sentences. In particular, the frequency of verbs was not equated across the structural variable. This design problem can be overcome by crossing sentences with structural type. Sentences remain a random factor, but now each sentence frame is represented in each type, with the only difference between structural types being the kind of verb inserted into the frame. Unfortunately, complete crossing of this sort is not possible because the same frame cannot serve as a single-clause sentence and as a complement sentence. However, within the single-clause
vs complement distinction, balancing of materials can be achieved. The listing in Appendix B contains twelve single-clause sentence pairs and nine complement sentence triples. Members of each pair differ by a single word, the verb, which is either of the noun-phrase classification, in SNP, or the simple transitive type, in SS. The complement triples differ by verb and/or complementizer. Two members of the triple share the same noun-phrase complement verb, one with a that complementizer (CNP-that) and the other without (CNP-∅). The third member (CVP) contains a verb-phrase complement verb with to complementizer. Verbs within each frame have been closely matched for frequency of occurrence. The CNP-∅ sentences were included in Experiment II to replace the CNP-(for)...to materials excluded from Experiment I. They afford an additional opportunity to test the structural ambiguity hypothesis of verb complexity. CNP-∅ and CNP-that differ only in the presence of the complementizer. The hypothesis must predict that at a fast presentation rate CNP-∅ will have a greater processing backlog than CNP-that. Independent support for this prediction comes from a phoneme-monitor study conducted by Hakes(1972). He found longer latencies to monitor the initial phoneme of the subject noun in the complement clause when the complementizer had been deleted. In terms of the structural ambiguity model, we would expect the size of the CNP-∅ backlog to be comparable to that of CVP, because both constructions are disambiguated at the same point, immediately after the subject noun-phrase of the complement clause.

Compression. Materials for Experiment II were originally recorded at a mean rate of 3.7 words (5.7 syllables) per second. Presentation rate was doubled through electronic time-compression by a Varispeech I compressor-expander
(Lee, 1972). This method differs from that used for Experiment I inasmuch as the segments deleted from the voiced portions of the signal do not correspond to glottal cycles. Expanded-compressed sentences were produced by reduplicating the segments left in the compressed versions.

Presentation Conditions. The phoneme-monitor work of Foss and Lynch (1969) and the sentence recall experiment conducted by Foss and Cairns (1970) suggest that even at the normal presentation rate, if structure is sufficiently complex, additional processing time may be required beyond the termination of the sentence. In Experiment I, there were no effects of structural type on list recall with the normal presentation rate and a 750 millisecond delay interval. However, if the interval were very brief, then differences might be found between the various types. Generalizing from the results of Experiment I, SNP, CNP-0, and CVP should be most strongly affected by the displaced processing. To test this prediction, a condition which combined the expanded-compressed rate with a 200 millisecond interval was included in Experiment II. In all, three conditions of presentation were used: (Expanded-compressed + 200), (Compressed + 750), and (Compressed + 200). For SS, SNP, CNP-that, and CVP, the pattern of differences between the two compressed conditions should be the same as that found in Experiment I, that is, SNP and CVP should be adversely affected by the short delay interval, while the effect on SS and CNP-that ought to be negligible. Furthermore, if the structural ambiguity hypothesis is correct, then CNP-0 should behave in much the same fashion as CVP. The pattern seen for (Compressed + 200) should emerge with (Expanded-compressed + 200) if, in fact, the backlog is also produced at the normal rate. These predictions for word list performance are sum-
Procedure. As part of the task in Experiment I, subjects were asked to paraphrase the test sentences. The resulting set of acceptable paraphrases spanned a considerable range of elaboration and complexity of structure. Although the importance of providing a good paraphrase was stressed in the instructions prior to the experiment, subjects may have been able to provide an "adequate" paraphrase even on trials in which they truncated sentence processing in order to reduce interference with list encoding. It seems reasonable to assume that the quality of the paraphrase may reflect in some way the amount of sentence processing which the subject performed. A scoring procedure which looks only for the recoverability of major syntactic relations will be insensitive to more subtle differences. For example, each of the following paraphrases given for SS #2 in Appendix A was judged acceptable according to the scoring criteria: "It was during a game of squash that Sally happened to sustain an injury to her left elbow;" "Sally hurt an elbow." In Experiment II, subjects were asked to write the test sentences verbatim, rather than give paraphrases. An exact recall response provides the data needed to grade performance on the sentence as well as the list. Evidence presented by Forster (1970) suggests that the number of words recalled from the sentence will reflect the extent to which it has been structured. He used the rapid serial visual presentation (RSVP) technique to show subjects
strings of words taken from grammatical sentences but scrambled into random orders for presentation. The number of words recalled from such strings was directly related to the degree to which subjects reorganized the items into grammatical sequences.

Instructions in Experiment II were altered so that subjects were told to respond first by writing down an exact repetition of the sentence and then the words of the list. In all other respects, the procedure was the same as in Experiment I.

Scoring. The sentence repetitions were scored independently by two judges who computed the number of words recalled correctly, in the proper order. The following criteria were used in scoring: Omission of a word counted as an error, as did incorrect positioning. A positional interchange of two items was counted as a single error. Inclusion of material not found in the test sentence did not affect the scoring. Incorrect inflectional endings were not treated as errors, provided the stem word was correct. On less than 1% of the responses the judges differed in their scoring by one point (one word). These differences were resolved through subsequent discussion until agreement was achieved. The test sentences used in Experiment II ranged in length from nine to eleven words. For the purposes of analyzing the sentence repetitions, the proportion of words correct in each response was expressed as the number recalled from a ten-word sentence.

Design. Three test sequences corresponding to the three presentation conditions were prepared for each member of the single-clause sentence pairs and the complement sentence triples. Nine tapes were constructed by selecting one sequence from each frame. On every tape, for each single-clause
type (SS and SNP), there were two sentences representing each of the three conditions. For each complement type (CNP-that, CNP-∅, and CVP), there was one sentence representing each presentation condition. Sentences were randomly arranged in blocks of seven, with each block containing at least one example of each structural type and at least two occurrences of each presentation condition. Instructions were altered so that subjects were told to respond first by writing down an exact repetition of the sentence and then the words of the list. Thirty-six subjects (four per tape) were paid for their participation in the forty minute experiment. They were run individually or in groups up to four in number.

Results

Two sets of analyses of variance were performed, one for list data and the other for sentence data. Within each set, a subject analysis, a single-clause sentence analysis, and a complement sentence analysis were computed. The design for each subject analysis consisted of one random between-subjects factor, tape(9); and two fixed within-subjects factors, structural type(5) and presentation condition(3). The single-clause sentence analyses contained two fixed within-sentences variables, structural type(2) and presentation condition(3). Similarly, the complement sentence analyses also had two fixed within-sentences variables, structural type(3) and presentation condition(3). In the two subject analyses, the tape variable was not significant in its main effects or interactions with the other independent variables and was therefore deleted from subsequent analysis.
List Performance. The main effects for structural type and presentation condition failed to reach significance on subject and sentence analyses. Orthogonal contrasts were used to test the predictions concerning the interaction of structural type and presentation condition. For SNP, CNP-∅,

and CVP, list recall in (Expanded-compressed + 200) did not differ from (Compressed + 200) \( (c_1) \) for CONDITION X STRUCTURE in TABLE II. Both conditions were marginally inferior to (Compressed + 750) \( (c_2) \). With SS and CNP-that, list recall in (Expanded-compressed + 200) did not differ from (Compressed + 200) \( (c_3) \), nor did these two conditions differ from (Compressed + 750) \( (c_4) \). The data are presented graphically in Figure 1. Although the effect is much weaker than in Experiment I, the pattern is the same, i.e., fewer list items recalled from SNP and CVP when the rate is fast and the delay interval short, but no comparable decrement in number of items recalled with SS and CNP-that. (Expanded-compressed + 200) seems to have been treated very much like (Compressed + 200) for both the high complexity and low complexity sentences.

Sentence Performance. The main effect for structural type was significant in
the subject analysis \( F(4, 140) = 11.728, p < .01 \) but non-significant in the single-clause and complement sentence analyses \( F(1, 11) = 1.172 \); and \( F(2, 16) = 0.196 \), respectively). The effect for structural type was produced almost entirely by a large difference between single-clause and complement sentences \( F(1, 140) = 42.815, p < .001 \). Materials were prepared by crossing sentence frame with structural type, but only within the single-clause and complement classifications. Therefore, the effect for type seen in the subject analysis is completely confounded with a difference in materials. The main effect for condition of presentation was significant in the subject analysis \( F(2, 70) = 13.243, p < .01 \) as well as in the single clause \( F(2, 22) = 7.178, p < .01 \) and complement \( F(2, 16) = 9.202, p < .01 \) sentence analyses. Sentence recall for (Expanded-compressed + 200) was superior to the other two conditions \( (c_2 \text{ for CONDITION in TABLE III}) \). Of primary interest are the effects of presentation condition on the structural variables. In the analysis of Experiment I, care was taken to compare presentation effects within particular structural types, so that, for example, CVP in (Compressed + 200) was compared only to CVP in other presentations. In the design of Experiment II, direct comparisons can be made among types, so long as the single-clause and complement division is maintained. The model must predict that at some stage interference from the list will impair sentence recall for \( \text{CNP-Ø} \) and CVP, but not for \( \text{CNP-that} \). Similarly, \( \text{SNP} \) should be adversely
affected by a condition in which the recall of SS remains relatively undiminished. These two predictions were tested for each of the compressed conditions. In (Compressed + 200), neither was confirmed, i.e., CNP-that did not differ from CNP-Ø and CVF (c₁ for CONDITION X STRUCTURE in TABLE III), and SS did not differ significantly from SNP (c₂). For (Compressed + 750) both predictions were confirmed (c₃ and c₄).

Discussion

The list data indicate that CNP-Ø, CVF, and SNP are all adversely affected by a short post-sentence interval, even when presentation is at the normal rate. This suggests, as do the earlier results, that even in the usual course of comprehension these structural types are relatively difficult. Conversely, the insensitivity of SS and CNP-that to the short delay indicates little, if any, processing backlog.

---

Insert Figure 2 about here

---

The sentence recall data are presented graphically in Figure 2. The predicted structural effects were found in (Compressed + 750) but not in the other two conditions, apparently because (Compressed + 200) and (Expanded-compressed + 200) happened to represent extremes for sentence processing. In the former, all constructions showed a sizable decrement in words recalled; in the latter, performance was uniformly high. During the additional 550
milliseconds of post-sentence interval contained in (Compressed + 750), the backlog was completed for CNP-that and SS but not for CNP-Ø, CVP, or SNP.

Results for sentence recall in Experiment II suggest that interference between sentence processing and list encoding was reciprocal. Although both the list and the sentence results display the predicted structural effect, they do so for different conditions. The two patterns of results can be accounted for if we assume that the listener must reach some criterial stage in sentence processing before he is willing to discontinue it in favor of list encoding. List recall for CNP-that is relatively unaffected by the (Compressed + 200) condition because the minimal acceptable level of sentence processing is reached within the short interval. For CNP-Ø and CVP more time is required to reach criterion. Sentence recall is equally low for all three complement types because each has had its processing interrupted before completion. Under the (Compressed + 750) condition, sentence recall displays the structural effect because processing has been completed for CNP-that and interrupted for the other two complement types. Word list recall is high for all three because processing has either been completed or has reached an interruptable stage within the 750 milliseconds. In the (Expanded-compressed + 200) condition, the sentence data indicate that processing was completed for all constructions. The list data suggest that for CNP-Ø, CVP, and SNP, it was completed at the expense of list encoding. Explanations based on partial sentence processing should, of course, be treated with appropriate caution. However, in view of the subject's goal of maximizing his response on both sentence and word list, it seems reasonable to consider partial processing as a possible strategy.
Exact repetition responses provide an opportunity to look at particular substrings of the test sentences. The structural ambiguity hypothesis of verb complexity should predict poorer recall of the ambiguous substrings in CNP-∅, CVP, and SNP when sentence processing encounters interference from list encoding. Ambiguous portions ought to be more vulnerable if they are held in a relatively unstructured form over a long duration. Figure 3 shows the proportion of words recalled at several positions in each of the constructions for the (Expanded-compressed + 200) and (Compressed + 750) conditions. Along the abscissa of each graph, a notational description is given for each of the positions examined. Beneath the description are the corresponding words taken from complement sentence #9 and single clause sentence #1 in Appendix B. Mean serial position is listed below each example item. For CNP-∅ and CVP, the main verb (V of S₀) and the following noun represent the ambiguous substring. The matrix subject (N of S₀) and the verb of the complement (V of S₁) lie in unambiguous regions. For the complement sentences in TABLE IV, the analysis indicates that the decrement in recall of the ambiguous portions is significant (c₁ and c₃), but in the un-
ambiguous regions, it is not \( (c_2 \text{ and } c_4) \). For SNP, three positions were examined: the main verb with its object, constituting the ambiguous sub-string, and the subject noun. Again the prediction is confirmed \( (c_1 \text{ and } c_2) \), i.e., the difference between ambiguous portions is significant, but for un-ambiguous portions it is not. In CNP-that and SS, performance over the corresponding substrings is unaffected by the (Compressed + 200) condition. The effects seen with CNP-Ø, CVP, and SNP cannot simply be explained on the basis of differences in serial position because the serial order of items is highly comparable among members of each sentence frame.

The substantial difference between single-clause and complement sentences (contrast for STRUCTURE in TABLE III) is contrary to the original prediction based on the results of the Holmes and Forster RSVP experiment. The sentences of Experiment II were not rated for naturalness prior to their use in the study; however, several readers had agreed informally that all versions of the test frames sounded very natural. Subsequent to the experiment, twelve additional subjects were asked to rate the sentences on a five-point scale from +2 "Perfectly natural and acceptable" to -2 "Very unnatural (clumsy or bizarre)." Each subject judged only one version of each test frame. The sentences of Experiment II were randomly embedded in a list consisting of materials used in Experiment I, other sentences used in the RSVP study, and constructions containing ambiguities of various kinds. For the materials of Experiment II, there was no difference between the mean rating for single-clause sentences (1.45) and complement types CNP-that and CVP (1.48). Therefore, the observed difference in recall cannot be attributed to a disparity in naturalness. In fact, naturalness seemed to bear no clear or con-
sistent relation to performance. The highest ratings were given to CVP (1.61) and the lowest were given to CNP-\(\emptyset\) (0.85), but in Experiment II, the number of words recalled was the same for the two complement forms.

The possibility that a subject will correctly report a given word from a sentence under rapid serial visual presentation is related, in part, to the syntactic class from which the word is drawn. Recall of nouns and verbs is generally superior to that of adjectives and function words (Holmes, 1973). For this reason, Holmes and Forster balanced the materials in their experiment for frequency of syntactic class. The materials of Experiment II were not balanced in this manner. Therefore, we might expect inequalities in the distribution of classes to account for the single-clause vs. complement difference. This would be the case if single-clause sentences happened to contain more nouns and verbs, and complement sentences contained more adjectives and function words. For the materials in Appendix B, the composition by class is as follows: In the single-clause frames, 43% of the words are nouns or verbs, with 57% adjectives or function words. In the complements, 50% are nouns or verbs. If the effect of word class were the determining factor in recall, then more words should have been remembered for the complement sentences and fewer for the single-clause constructions. The great superiority of single clause sentences in Experiment II cannot be explained by differences in naturalness or in word class composition.

The difference between RSVP and compressed speech performance with CNP-that and CVP also remains unexplained. Forster (1970) has argued that in RSVP the presentation of each successive word eliminates the viewer's sensory copy of the preceding item. Under these conditions, the reader must
process each word as it is presented if he is to be able to store it for future recall. Backward masking does not force the listener to process each word as it is heard in a time-compressed sentence. In fact, the results of Experiments I and II suggest that short-term storage is used to enable the listener to "catch-up," i.e., to complete displaced processing. Clearly, too little is currently known about the two paradigms to enable us to answer questions concerning differences in their results. As a first step, however, we may try to eliminate materials effects as an explanation for the disparity. One might wish to argue that the findings of Experiment I and the Holmes and Forster RSVP study do not reflect differences in the way the two techniques interact with various structural types, but rather reflect different interactions with materials-specific variables that are incidental to the experiments. A solution to the problem requires that both paradigms be applied to the same balanced materials, such as those of Experiment II.
EXPERIMENT III

In Experiment III, lexically ambiguous sentences and their control versions were compressed and presented to subjects under the same task conditions that were used to examine structural variables in Experiments I and II. Our predictions concerning the interaction of lexical ambiguity and presentation rate will depend upon the nature of the model we wish to construct for the treatment of ambiguity. The evidence presented below suggests that (1) the listener accesses the alternative interpretations of ambiguous lexical items, but (2) this access is costly in terms of processing capacity, and therefore (3) the ambiguity must be resolved very rapidly. The picture is consistent with the Garrett and Holmes finding of reduced recall for ambiguous items in RSVP, and the Cairns observation of a highly localized effect of lexical ambiguity on phoneme-monitor responses.

(1) Accessing alternative interpretations. Lackner and Garrett (1972) used a selective auditory attention paradigm to measure the influence of unattended biased contexts on the interpretation of attended ambiguous sentences. In the attended ear, the subject heard sentences containing various types of ambiguity. Simultaneously, an attenuated context sentence was presented to the unattended ear. The context was neutral or biased toward one or the other reading of the ambiguity. On the basis of subjects' paraphrases of the attended sentences, Lackner and Garrett concluded that the biased contexts were effective in influencing the subjects' interpretations. The experimenters suggest that the results can best be accounted for by a model in which the listener actively pursues all the options for an ambiguity until
he has some basis for choosing one.

(2) The cost of accessing alternative interpretations. Foss and Jenkins (1973) used the phoneme-monitor paradigm to measure the load immediately following an ambiguous word in a sentence which provided either a neutral or a biased context. Reaction times were compared to those obtained from sentences in which the ambiguous word was replaced by an unambiguous control word. The experimenters found that response latency was greater after the lexically ambiguous word, and that this was true for both neutral and biased contexts. Furthermore, the difference was present whether or not the subjects perceived the ambiguity in the sentences. Foss and Jenkins cite these results to support a model of processing in which all interpretations of an ambiguous word are always activated when the word is encountered in the sentence. The alternative readings are transferred into a limited-capacity working memory, resulting in a greater processing load than would normally be incurred in an unambiguous sentence.

(3) Resolving the ambiguity. Conrad (1974) played recorded test sentences to her subjects and followed each by the immediate presentation of a word printed in colored ink. The subjects' task was to name the color of ink in which the word was presented. Latency to respond was assumed to be an indication of the amount of interference between the word and the color-naming response. Test sentences contained a lexically ambiguous word either preceded by a biased context, in which case the ambiguous word was the last word in the sentence, or followed by a biased context, in which case the ambiguous word was the first noun in the sentence. When the ambiguous item was used as the post-sentence word, color-naming latency was
longer for test sentences than for controls. In the prior-context sentences, response times were also longer when the post-sentence word was the name of the superordinate category appropriate to either the biased reading or the alternative reading of the ambiguous item. The effect could not be found for the subsequent-context sentences. Conrad interprets these results as supporting a model in which both readings of a lexically ambiguous word are activated upon encountering that word, even if a prior biasing context is present. However, this activation is measurable for only a short time after the item's occurrence in the sentence.

On the basis of preceding work, the following predictions seem to be well motivated: At a normal presentation rate, there should be no difference in the post-sentence processing load for lexically ambiguous sentences and their matched unambiguous controls, because lexical ambiguity is resolved quickly. However, when an ambiguous sentence is compressed, the resolution of lexical ambiguity will occupy a larger proportion of total identification time. If the structuring processes are dependent upon input from an identification component, then structural processing should lag further behind when the compressed sentence contains an ambiguous word. This effect of compression ought to be observable with a short delay interval preceding the list, but not with a long delay which would allow displaced processing to be completed.

Method

Materials. The eight lexically ambiguous-unambiguous sentence pairs in
Appendix C were adapted from materials previously reported (Cairns, 1970) to be relatively unbiased in interpretation. Members of each pair differed by only one item, the ambiguous word or its frequency-matched control. The mean frequency of occurrence per 1.014 million words (Kucera and Francis, 1967) was 126 for the lexically ambiguous words and 111 for their control substitutes. In seven of the pairs, the ambiguous item was a noun, and in one it was an adjective. The sentences of Experiment III were included in the naturalness ratings list along with materials from Experiments I and II and RSVP studies. Each of the subjects asked to provide ratings saw only one version of each ambiguous-control pair. Overall, the materials of Experiment III were judged to be less natural than the other constructions in the list. However, the difference between ambiguous and control versions proved to be negligible (1.14 and 1.19, respectively). The syntactic form of the sentences in Experiment III was quite heterogeneous: #1-3 in Appendix C are SS; #4 is CNP-(for)...to; #5, CVP; #6, SNP; #7 and #8, CNP-that.

Compression. The materials were originally recorded at a mean rate of 3.8 words (5.3 syllables) per second. Presentation speed was doubled through pitch-synchronous time-compression, and normal duration copies were made by expansion.

Presentation conditions. (Expanded-compressed + 200), (Compressed + 750), and (Compressed + 200).

Procedure. The paradigm of Experiments I and II was used to present the test materials. Half the subjects were instructed to paraphrase the sentence before recalling the list. The other half were asked to give a verbatim repetition of the sentence before writing the list. Subjects were not told
that they would be hearing ambiguous sentences. However, two of the six sentences in the practice trials contained lexical ambiguities.

**Scoring.** The sets of criteria used for scoring in Experiment I and Experiment II were applied to the paraphrases and repetitions, respectively.

**Design.** Three test sequences corresponding to the three presentation conditions were prepared for each ambiguous and each control sentence. Six tapes were constructed by selecting one sequence from each materials pair. On every tape, for each value of ambiguity (ambiguous vs. control), there was at least one test sentence presented under each of the three conditions. A set of eight filler sentences was presented along with the experimental materials. The ordering of filler and test sequences was randomly established. Sixty subjects (five per tape for each of the two response tasks) were paid for their participation in the twenty-five minute experiment. They were run individually or in groups up to four in number.

**Results**

Two sets of analyses of variance were computed, one for performance on the word list, using data from both response-task groups, and the other for recall of the sentences, with data from the repetition group.

**List performance.** The design for the subject analysis of list data consisted of one fixed between-subjects factor, response task(2); one random between-subjects factor, tape(6); and two fixed within-subjects factors, presentation condition(3) and ambiguity(2). Once again, the random between-subjects tape factor was not significant in its main effect or its interactions and was
dismissed from the remainder of the analysis. In the sentence analysis of list data, there were three fixed within-sentences variables: response task(2), presentation condition(3), and ambiguity(2). The main effect for ambiguity was significant on both analyses ($F(1,58) = 6.641, p < .05$; and $F(1,7) = 13.186, p < .01$; by subjects and sentences, respectively). The main effects for response task and presentation condition were not significant. In TABLE V, the means are listed for the interaction of ambiguity with presentation condition, and the data are presented graphically in Figure 4.

Three orthogonal contrasts were used to test the predictions. As expected, there was no difference between ambiguous and control versions of the sentences for (Expanded-compressed + 200) ($c_1$ for AMBIGUITY X CONDITION in TABLE V). Contrary to prediction, the unambiguous-ambiguous difference was as large for (Compressed + 750) as it was for (Compressed + 200) ($c_2$). This difference proved to be significant in the latter two conditions ($c_3$).

Sentence performance. Two analyses were computed for the words recalled from the test sentences in the repetition task. The designs for the analyses were the same as those given for the list data, except, of course, for
the absence of the response task factor. None of the main effects or interactions in either analysis was significant. For the paraphrase task, 3% of the responses were judged unacceptable. The distribution of errors follows: by presentation condition, (Expanded-compressed + 200) -1, (Compressed + 750) -2, (Compressed + 200) -4; by ambiguity, ambiguous -3, control -4.

Discussion

The pattern of list recall for the unambiguous sentences in the three presentation conditions (see Figure 4) can be accounted for on the basis of their syntactic composition. As noted earlier, the sentences of Experiment III represented a variety of structural types. Overall, list performance was somewhat better in the (Compressed + 750) condition than in (Expanded-compressed + 200) or (Compressed + 200), although this difference was not significant. In Experiment II, when list recall was averaged across the structural variable (cf. means for CONDITION in TABLE II), a similar pattern was obtained. It is with the ambiguous sentences that recall departs from a pattern which can be accounted for by structural form.

The main effect for ambiguity was significant on both analyses. However, the difference between ambiguous and control versions for (Expanded-compressed + 200) was not significant. If lexical ambiguity is normally resolved very quickly, then it is reasonable to assume that it will have little or no effect when presentation is at the normal rate. The effect of ambiguity in the two compressed conditions is not so easily explained.
Why does the list recall decrement persist with the 750 millisecond delay interval? In Experiments I and II, the longer delay proved sufficient to restore performance on the word list; for the ambiguous sentences in Experiment III, it did not. Additional time provided after the sentence was not adequate compensation for the immediate increased load produced by the ambiguous item. The persistence of a significant ambiguous-unambiguous difference for (Compressed + 750) suggests an effect whose source differs from that found in the first and second experiments.

The analyses for total sentence recall provide no indication of a difference between ambiguous and control sentences. However, a post hoc examination of recall for the ambiguous or control item does suggest a difference. Seventeen of the thirty subjects in the repetition group remembered the same number of ambiguous and control words under the compressed conditions. Of the remaining thirteen subjects, twelve recalled more control items ($p < .005$, two-tailed sign test). For six of the eight sentence pairs, more control words were recalled. Under the (Expanded-compressed + 200) condition, six subjects remembered more control words, and four remembered more ambiguous ones. For four of the sentence pairs it was the control item that was reported more frequently, and for three of the pairs the ambiguous item was recalled more often. The results suggest an increased load for the ambiguous word when the sentence is compressed but not when it is presented at the normal rate. The effect of this load on list performance is not eliminated when time is provided at the end of the sentence. According to the hypothesis, when additional time is available immediately after the lexical ambiguity, the decrement in list
recall should disappear. The results for the expanded-compressed sentences confirmed this prediction. However, a stronger test of the hypothesis would involve a comparison between performance with uniform compression of the sentence and performance with compression of all but a small portion of the sentence immediately following the ambiguous or control word. The ambiguous-unambiguous difference for these conditions should be the same as it is in Experiment III for Compressed and Expanded-compressed, respectively. Notice that we would predict no significant improvement to result from adding processing time immediately following the complement verb of a compressed sentence, because at that locus the processing load is not primarily the result of insufficient time but rather of insufficient information.

The distinction drawn earlier between identification and structuring may now be re-stated in terms suggested by Experiments I, II, and III. Decisions about the identification of an item are made very shortly after the item's presentation. Increases in load at this stage seem to be relatively unaffected by subsequent manipulations of processing time. On the other hand, decisions about syntactic structure can be postponed and are therefore sensitive to the duration of the post-sentence interval.
1. The design of Experiment I precludes an examination of treatment effects (structural type and presentation condition) without collapsing data across either subjects or sentences. Performing the separate analyses allows us to look at the significance of effects using subjects as the factor of repeated measures and sentences as the factor of repeated measures. This is highly desirable since they represent the two populations across which we wish to generalize.
**STRUCTURE**

<table>
<thead>
<tr>
<th>SS</th>
<th>SNP</th>
<th>SVP</th>
<th>that</th>
<th>CVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.49</td>
<td>4.34</td>
<td>4.38</td>
<td>4.55</td>
<td>4.19</td>
</tr>
</tbody>
</table>

**CONDITION**

<table>
<thead>
<tr>
<th>Control A</th>
<th>Control B</th>
<th>Compressed + 750</th>
<th>Compressed + 200</th>
<th>Subject Analysis (df = 93)</th>
<th>Sentence Analysis (df = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.53</td>
<td>4.50</td>
<td>4.41</td>
<td>4.12</td>
<td>F: 0.749</td>
<td>F: 0.895</td>
</tr>
</tbody>
</table>

| c_1 | 0 | 1 | 1 | 0 | 0 | -1 | -1 | 0 | 0 |
| c_2 | 1 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 |

**CONDITION X STRUCTURE**

<table>
<thead>
<tr>
<th>Control A</th>
<th>Control B</th>
<th>Compressed + 750</th>
<th>Compressed + 200</th>
<th>Subject Analysis (df = 349)</th>
<th>Sentence Analysis (df = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c_1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c_2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c_3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c_4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*significant at the .05 level

**EXPERIMENT I**

mean number of words recalled from list

**TABLE I**
<table>
<thead>
<tr>
<th>STRUCTURE</th>
<th>CONDITION</th>
<th>Expanded-compressed</th>
<th>Compressed</th>
<th>Compressed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+200</td>
<td>+750</td>
<td>+200</td>
</tr>
<tr>
<td>SS</td>
<td>4.71</td>
<td>4.79</td>
<td>4.37</td>
<td>4.54</td>
</tr>
<tr>
<td>SNP</td>
<td>4.76</td>
<td>4.71</td>
<td>4.33</td>
<td>4.56</td>
</tr>
<tr>
<td>CNP-∅</td>
<td>4.81</td>
<td>4.96</td>
<td>4.56</td>
<td>4.56</td>
</tr>
<tr>
<td>CNP-that</td>
<td>4.63</td>
<td>4.69</td>
<td>4.22</td>
<td>4.50</td>
</tr>
<tr>
<td>CVP</td>
<td>4.52</td>
<td>4.72</td>
<td>4.50</td>
<td></td>
</tr>
</tbody>
</table>

**CONDITION X STRUCTURE**

**EXPANDED-COMPRESSION + 200**

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>SNP</th>
<th>CNP-∅</th>
<th>CNP-that</th>
<th>CVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>c₁</td>
<td>4.71</td>
<td>4.71</td>
<td>4.33</td>
<td>4.56</td>
<td>4.31</td>
</tr>
<tr>
<td>c₂</td>
<td>4.81</td>
<td>4.96</td>
<td>4.56</td>
<td>4.56</td>
<td>4.72</td>
</tr>
<tr>
<td>c₃</td>
<td>4.63</td>
<td>4.69</td>
<td>4.22</td>
<td>4.50</td>
<td>4.47</td>
</tr>
<tr>
<td>c₄</td>
<td>4.52</td>
<td>4.72</td>
<td>4.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sentence Analyses**

**Subject Analysis** (df=280) F

- c₁: 0.007
- c₂: 4.472*
- c₃: 0.239
- c₄: 0.181

**Single-Clause Complement** (df=22) F

- c₁: 0.003
- c₂: 2.981#
- c₃: 0.264
- c₄: 0.261

**Experiment II**

Mean number of words recalled from list

**Table II**
### EXPERIMENT II

Mean number of words recalled from sentence

#### TABLE III
### CONDITION X STRUCTURE X POSITION

<table>
<thead>
<tr>
<th>Noun</th>
<th>Verb</th>
<th>Noun</th>
<th>Verb</th>
<th>Noun</th>
<th>Verb</th>
<th>Subject Analysis (df=140) F</th>
<th>Sentence Analysis (df=44) F</th>
</tr>
</thead>
<tbody>
<tr>
<td>.90</td>
<td>.90</td>
<td>.83</td>
<td></td>
<td>.81</td>
<td>.78</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>c₁</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>15.766*</td>
</tr>
<tr>
<td>c₂</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>4.004#</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Noun</th>
<th>Verb</th>
<th>Noun</th>
<th>Verb</th>
<th>Noun</th>
<th>Verb</th>
<th>CVP Noun</th>
<th>Verb</th>
<th>CVP Noun</th>
<th>Verb</th>
<th>CVP Noun</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>.78</td>
<td>.56</td>
<td>.78</td>
<td>.81</td>
<td>.75</td>
<td>.31</td>
<td>.64</td>
<td>.83</td>
<td>.69</td>
<td>.58</td>
<td>.75</td>
<td>.75</td>
</tr>
<tr>
<td>c₁</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c₂</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c₃</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c₄</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Subject Analysis**

(1) F

| c₁   | 8.713* |
| c₂   | 0.000  |
| c₃   | 8.669* |
| c₄   | 1.097  |

**Sentence Analysis**

(1) F

| c₁   | 7.739* |
| c₂   | 0.123  |
| c₃   | 7.698* |
| c₄   | 0.975  |

#significant at .08 level
*significant at .05 level

**EXPERIMENT II**

mean proportion words recalled from selected positions in sentence

**TABLE IV**
<table>
<thead>
<tr>
<th>AMBIGUOUS</th>
<th>UNAMBIGUOUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.46</td>
<td>4.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EXPANDED-COMPRESSED</th>
<th>COMPRESSED + 200</th>
<th>COMPRESSED + 750</th>
<th>COMPRESSED + 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBIGUOUS</td>
<td>4.65</td>
<td>4.73</td>
<td>4.52</td>
</tr>
<tr>
<td>UNAMBIGUOUS</td>
<td>4.73</td>
<td>4.52</td>
<td></td>
</tr>
</tbody>
</table>

**AMBIGUITY X CONDITION**

<table>
<thead>
<tr>
<th>AMBIGUOUS</th>
<th>UNAMBIGUOUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.59</td>
<td>4.71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EXPANDED-COMPRESSED</th>
<th>COMPRESSED + 750</th>
<th>COMPRESSED + 200</th>
<th>Subject Analysis (df=116)</th>
<th>Sentence Analysis (df=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBIGUOUS</td>
<td>4.48</td>
<td>4.31</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>UNAMBIGUOUS</td>
<td>4.98</td>
<td>4.73</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| c₁ | -1 | 1  | 0  | 0  | 0  | 0  | 0.255 | 0.073 |
| c₂ | 0  | 0  | -1 | 1  | 1  | -1 | 0.061 | 0.035 |
| c₃ | 0  | 0  | -1 | 1  | -1 | 1  | 8.020* | 4.479# |

#significant at the .08 level

* " " " " .05 "

**EXPERIMENT III**

mean number of words recalled from list

**TABLE V**
EXPERIMENT II

mean number of words recalled from list

Figure 1
EXPERIMENT II
sentence recall

Figure 2
EXPERIMENT II
recall of selected positions in CNPØ

Figure 3a
Mean serial position

<table>
<thead>
<tr>
<th>Position</th>
<th>Patrolman</th>
<th>Allow</th>
<th>Sergeant</th>
<th>Push</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial</td>
<td>2.33</td>
<td>3.33</td>
<td>5.22</td>
<td>7.22</td>
</tr>
</tbody>
</table>

**Experiment II**
recall of selected positions in CVP

Figure 3b
EXPERIMENT II
recall of selected positions in SNP

Figure 3c
EXPERIMENT III

mean number of words recalled from list

Figure 4
APPENDIX A
Sentences of Experiment I

SS: Single-Clause Sentences with Simple Verbs
(1) Alan left a large pile of books in the library.
(2) Sally injured her left elbow during a game of squash.
(3) The cleaners couldn't empty all the ashtrays in the building.
(4) Many people attended the seminar on the government's foreign policy.

SNP: Single-Clause Sentences with Noun-Phrase Complement Verbs
(1) The actors didn't like the plot of the new play.
(2) The doctors announced their strong support for the health scheme.
(3) Your cousin doubted the truth of the rumors about Jim.
(4) The judge decided the case in favor of your daughter.

SVP: Single-Clause Sentences with Verb-Phrase Complement Verbs
(1) The prime minister inspired the members with a brilliant speech.
(2) Betty relied on her friends for advice on financial problems.
(3) Bob hired a group of young men for the job.
(4) John's brother trained all the animals for the local circus.

CNP-that: Noun-Phrase Complement Sentences with that Complementizers
(1) Susan realized that her brother had bought several new records.
(2) The children believed that their mother was in serious trouble.
(3) The lawyer didn't think that his client was a thief.
(4) Two strange men said that the elevator wasn't working properly.

CVP: Verb-Phrase Complement Sentences
(1) The angry officials urged the man to complete the form.
(2) Jim ordered his young brother to wash the dirty dishes.
APPENDIX A

Sentences of Experiment I

CVP: Verb-Phrase Complement Sentences

(3) Several people forced the foreign students to leave the hotel.

(4) The author didn't encourage anyone to read his early books.
APPENDIX B

Single-Clause Sentences of Experiment II

simple verb

noun-phrase complement verb

located

(1) The helicopter crew located the wreckage in the mountains.

discovered

(2) The structural engineer discovered the strength of the new design.

improved

(3) The bird watcher improved a very rare species in the woods.

observed

(4) The surveyor observed the length of the plot of land.

determined

(5) The contestants determined the number of jelly beans in the jar.

counted

(6) The fireman counted removed a potential fire hazard in the building.

noticed

(7) The students noticed covered most of the material in their textbook.

covered

(8) The foreign diplomat covered discussed the provisions of the treaty proposal.

revealed

(9) The astronomer revealed checked the accuracy of his original calculations.

doubted

(10) The waiter doubted dropped suggested a plate of corned beef and cabbage.

studied

(11) The physician studied reported some recent cases of the once-rare disease.
Single-Clause Sentences of Experiment II

simple verb

noun-phrase complement verb

defended

(12) The manufacturer \[\text{defended}\] the superior quality of his product.\[\text{guaranteed}\]
APPENDIX B

Complement Sentences of Experiment II

CNP0: Noun-Phrase Complement Sentence without Complementizer
CNP-that: Noun-Phrase Complement Sentence with that Complementizer
CVP: Verb-Phrase Complement Sentence

(1) The new law specifies taxpayers can take an extra deduction.
   The new law specifies that taxpayers can take an extra deduction.
   The new law enables taxpayers to take an extra deduction.

(2) The general assumed his men would fight bravely in battle.
   The general assumed that his men would fight bravely in battle.
   The general commanded his men to fight bravely in battle.

(3) The broker requested the company issue more stock.
   The broker requested that the company issue more stock.
   The broker encouraged the company to issue more stock.

(4) The banker recalled Joe had repaid the loan on time.
   The banker recalled that Joe had repaid the loan on time.
   The banker trusted Joe to repay the loan on time.

(5) The young minister hinted his congregation should be more charitable.
   The young minister hinted that his congregation should be more charitable.
   The young minister inspired his congregation to be more charitable.

(6) The hijackers demanded the airline follow their instructions.
   The hijackers demanded that the airline follow their instructions.
   The hijackers directed the airline to follow their instructions.

(7) The scientist predicted his assistants would verify the controversial experiment.
   The scientist predicted that his assistants would verify the controversial experiment.
   The scientist invited his assistants to verify the controversial experiment.
APPENDIX B

Complement Sentences of Experiment II

CNP: Noun-Phrase Complement Sentence without Complementizer
CNP-that: Noun-Phrase Complement Sentence with that Complementizer
CVP: Verb-Phrase Complement Sentence

(8) The judge insisted an experienced lawyer handle the case.
    The judge insisted that an experienced lawyer handle the case.
    The judge chose an experienced lawyer to handle the case.

(9) The rookie patrolman feared the sergeant would push him around.
    The rookie patrolman feared that the sergeant would push him around.
    The rookie patrolman allowed the sergeant to push him around.
APPENDIX C

Sentences of Experiment III

ambiguous
unambiguous

right
left
glasses
dishes

(1) The driver took the turn at the intersection.
(2) The artist put the glasses on the table carefully.
(3) The athletes met near the bridge late yesterday evening.
(4) They didn't want the committee to be too small.
(5) All the boys watched the bus leave the depot.
(6) The cleaner found the pipe under the bench today.
(7) They knew that the plant would be sold.
(8) The elderly farmer saw that the cane was damaged.
References


