INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
THE LIMITS OF ABSTRACT LINGUISTIC IDEAS:
THE TIME COURSE AND IMPORTANCE OF REFERENCE IN INTEGRATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Jonpaul Rogers Martin, B.A., M.A.

* * * * *

The Ohio State University
1980

Reading Committee:
Neal F. Johnson, Ph.D.
Dean H. Owen, Ph.D.
Lester E. Krueger, Ph.D.

Approved By
Neal F. Johnson
Advisor
Department of Psychology
Dedication
To The Club Lawrence (Larry's), and All Who Made It a
Superb Wampeter and Wonderous Grandfalloon.
Acknowledgments

My thanks and appreciation are given to almost all the instructors and professors I ever had, they virtually all helped make learning and solving new problems fun and exciting. Of course, a special few deserve special mention for spending so much time and effort educating a smallish third baseman who could hit, field, throw, but not move well to his right: Neal F. Johnson, Dean H. Owen, Delos D. Wickens, James Erickson, Harvey Shulman, George Briggs, Lester Kreuger, Don Meyer, Mari Jones, Joe Spitzner, and more.

Along the way, from the first Fall days of 1964 when Columbus felt so lonely, through all the good times, rewards, and growth, til now when Ohio State is home— with warm, wonderful memories, all the effort was worth it. My grateful and life-long thanks to those friends and buddies who were so dear: Jennifer Gille, Lefty Ruff, Giff Rector, Rick and Patty Grant, Sam Ladrach, Ron Huiatt, Harvey Marmurek, Andy Logan, Jim Fritzen, Anne O'Keefe, and some others. Finally, I wish to acknowledge a debt probably unpayable to three unselfishly supportive folks, Helen Dean and Anthony Aloysius Martin, and Laura Daniele.
VITA

September 8, 1946. . . . . Born - Warren, Ohio

1969 . . . . . . . . . . . . B.A., The Ohio State University, Columbus, Ohio

1970-1974. . . . . . . . N.I.M.H. Pre-Doctoral Trainee Fellowship, The Ohio State University, Columbus, Ohio

1973 . . . . . . . . . . . . M.A., The Ohio State University, Columbus, Ohio

1974-1976. . . . . . . . Graduate Teaching Associate, Introductory Psychology Co-ordinator, The Ohio State University, Columbus, Ohio

1977-1978. . . . . . . . Assistant Professor of Psychology, State University of New York College at Cortland, New York

1979-1980. . . . . . . . Director of Research, Evaluation, and Grants, Western Reserve Psychiatric Center, Northfield, Ohio

PUBLICATIONS


iv

FIELDS OF STUDY

Major Field: Cognitive Psychology

Studies in Memory and Psycholinguistics: Professors Neal F. Johnson, Delos D. Wickens, J.R. Erickson

Studies in Perception and Information Processing: Professors Dean H. Owen, Harvey G. Shulman, Lester E. Kreuger, George Briggs, D. Gilson

Studies in Statistics and Design: Professors George Briggs, Mari Jones, Paul Isaac, Robert Wherry, J.R. Erickson
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>EXPERIMENT 1</td>
<td>45</td>
</tr>
<tr>
<td>Method</td>
<td>49</td>
</tr>
<tr>
<td>Results</td>
<td>56</td>
</tr>
<tr>
<td>EXPERIMENT 2</td>
<td>65</td>
</tr>
<tr>
<td>Method</td>
<td>68</td>
</tr>
<tr>
<td>Results</td>
<td>72</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>81</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>86</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>91</td>
</tr>
</tbody>
</table>

# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Complex or ( L^{-4} ) Sentences.</td>
<td>51</td>
</tr>
<tr>
<td>2.</td>
<td>A Sample Set of Related Sentences for One ( L^{-4} )</td>
<td>52</td>
</tr>
<tr>
<td>3.</td>
<td>Mean Recognition Ratings in a ( 2 \times 3 \times 5 \times 3 ) ANOVA</td>
<td>58</td>
</tr>
<tr>
<td>4.</td>
<td>Main Effect and Some Interaction Means in ( 2 \times 3 \times 5 ) ( \times 3 \times 3 ) ANOVA</td>
<td>62</td>
</tr>
<tr>
<td>5.</td>
<td>Interaction Means for ( 2 \times 3 \times 5 \times 3 \times 3 ) ANOVA</td>
<td>63</td>
</tr>
<tr>
<td>6.</td>
<td>Theme Sets for Experiment 2's Materials</td>
<td>70</td>
</tr>
<tr>
<td>7.</td>
<td>Mean Recognition Rating for List, Type, and Length of Test Sentence</td>
<td>75</td>
</tr>
<tr>
<td>8.</td>
<td>The Mean Number of Correct Recognitions for all Combinations of Tests.</td>
<td>79</td>
</tr>
<tr>
<td>9.</td>
<td>Source Table for ( 2 \times 3 \times 5 \times 3 ) ANOVA on Recognition Ratings for ( L^{-2} ) sentences</td>
<td>87</td>
</tr>
<tr>
<td>10.</td>
<td>Source Table for the ( 2 \times 3 \times 5 \times 3 \times 3 ) ANOVA on Recognition Ratings</td>
<td>88</td>
</tr>
<tr>
<td>11.</td>
<td>Experiment 2: ( 2 \times 3 \times 4 ) ANOVA on Recognition Rating.</td>
<td>89</td>
</tr>
<tr>
<td>12.</td>
<td>Experiment 2: ( 2 \times 3 \times 4 ) ANOVA on Total Number of Correct Recognitions</td>
<td>90</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mean Recognition Rating by Level of Experience</td>
<td>60</td>
</tr>
<tr>
<td>2.</td>
<td>Mean Number of Correct Recognitions by Sentence Type</td>
<td>74</td>
</tr>
<tr>
<td>3.</td>
<td>Recognition Rating of all Types and Length of Test Sentences</td>
<td>78</td>
</tr>
</tbody>
</table>
INTRODUCTION

The present experiments are concerned with memory for sentences conveying some theme or idea (Bransford & Franks, 1971; Cofer, 1973). The issue to be discussed is how the memorial representation of sentences or discourse comprehension should be conceptualized. Though the questions involved have a considerable history in the literature, the previous arguments may be reduced to two alternative views of language processing: an interpretative (linguistic object) hypothesis and a constructive (assimilative) hypothesis (Barclay, 1973; Fredericksen, 1975a).

In the past, the interpretative hypothesis has been more dominant. Each such theorist has emphasized different particulars, but all share a belief in the existence of some type of individual sentence record. The psychologists (e.g., Miller & McNeil, 1965; Anderson & Bower, 1972; Bower, 1967; Johnson-Laird, 1970, 1974) have focused on the proper encoding dimensions, while the linguists (Blumenthal, 1967; Chomsky, 1965; Katz, 1972; and others) have stressed how a surface structure interpretation is developed. Almost exclusively, interpretative theorists have worked with single sentences and have borrowed heavily from both computational linguistics (Simmons, 1968, 1970, 1971, 1972) and the general M.I.T.—Harvard transformational tradition (cf. Weimer, 1974, or Greene, 1972, for reviews).

The constructive hypothesis on the other hand argues that
individual sentences are not usually processed or retained in their entirety. Instead, it is assumed that inputs are selectively processed, using information derived from the current item together with contextual information and stored knowledge about the world to generate a semantic interpretation which fits the input (Fredericksen, 1975a, 1975b, 1975c; Jenkins, 1973). Thus a sentence is not thought to be treated in isolation, but processed into a larger, schematic representation in accordance with non-linguistic (a-linguistic) cognitive information.

Briefly then, the linguistic object theory maintains than an independent trace is formed for each sentence, while the constructive view suggests that memorial representations are much less specific, more general, more schematic. The primary distinction involves the hypothesized nature and detail of the sentence storage in memory (Melton, 1963).

Before presenting the history of these issues, it may prove helpful to discuss the close similarity between these views and the dichotomy of memory proposed by Tulving (1972; Tulving & Bower, 1974). A brief review of this currently popular distinction may prove useful in tracking the history of the language processing theories. Tulving (1972) speculated there are two types of encoding or memory systems, episodic and semantic or conceptual. Each of these systems closely resembles one of the general theories of sentence memory previously introduced. According to Tulving, every piece of information attended to in some fashion enters episodic memory. There it is represented in
terms of its temporal, spatial, frequency, and possibly orthographic attributes (Bower, 1967; Morton, 1969, 1970; Underwood, 1969). This encoding includes semantic information, but most critically, each sentence (or event) is tagged in terms of some set of temporal and spatial properties. This store compares quite favorably to the type of sentence memory called for in the abstractive or linguistic object hypothesis.

The second store in the Tulving dichotomy, semantic memory, also contains the semantic information of input events. However, here there is a recoding process in which the specific inputs are integrated (if possible) with the multi-dimensional structure of concepts stored in long-term memory (Atkinson, Hermann, & Westcourt, 1973; Lindsay & Norman, 1972; Rummelhart, Norman, & Lindsay, 1972). The same events as are registered in episodic memory lose their perceptual-temporal tags and are transformed in terms of their cognitive references so that their sum (theme) maps into the semantic memory of the individual. Thus, semantic memory (Tulving, 1972) for a set of sentences would appear to be very similar to a constructive hypothesis in suggesting an absence of individual sentence traces in a typical reading situation. The qualification, typical reading, is a recognition of the possibility that a reader could under some circumstances adopt an unusual strategy such as verbatim memorization (Tulving & Bower, 1974; Neisser, 1977).

It is important to realize that Tulving's concept of semantic memory is different, perhaps deeper (Craik & Lockhart, 1972) than
the more common usage of the term to mean lexical memory (Collins & Quillian, 1972; Meyer, 1975). Tulving does not refer to just a lexical-dictionary. He states that his semantic memory is a mental thesaurus containing all the organized knowledge a person has about words, other verbal symbols, their references, and all the rules and formulas for the use of this information. Not only are definitions involved, but also much more complex conceptual relations.

In summary, Tulving's episodic memory may be conceived of as a veridical record of the input events, with the time and place of the input being the most important aspects of that record. Semantic memory for sentences may be conceived of as a reconstructive schema (Barlett, 1932; Neisser, 1967, 1977) embedded in a multi-dimensional memory structure. The types of representation involved in episodic and semantic memory clearly parallel the linguistic object and constructive theories respectively. The experiments described within are an attempt to investigate the relationships between these types of encoding.

The overall model to be developed here is that both types of memory information (episodic and semantic) are available to the reader (listener) under particular conditions of time and experience, and perhaps strategy—therefore both general theories are correct in some sense. The contention is that an episodic encoding is a reasonable representation of initial (shallow) levels of memory, but that typical reading strategies and thematic material lead to the development and retention of a semantic
representation. Thus with multiple inputs, especially when a common theme is involved, there is a loss of episodic information for individual sentences. It is supposed that primarily a representative gist or schema results from processing thematic sentences. One of the major goals of the present experiment is to clarify this hypothesized constructive development.

The first experiment examines the time course of the constructive process using a continuous recognition paradigm with highly related sentences such as those used in Bransford and Franks (1971, 1973). The study attempts to discover how much thematic experience a reader must have before he or she evidences only a semantic as opposed to an episodic memory for a sentence. The experiment holds constant the amount of time between the presentation of a sentence and its recognition test, but systematically varies the amount of prior experience with the theme involved. The second experiment is concerned with the nature of the constructed schema. Previous studies which have observed integration (constructive) phenomena with these materials have consistently confounded words and their meaning (references).

In the second experiment, a typical Bransford and Franks (1971) methodology is used to study sentences which are still highly related in terms of references, but no longer have a complete confounding of words and those references. The test is essentially whether readers formulate schema when the input events are more discriminable in terms of word events, but no more different than typical Bransford and Franks (1971) sentences
in terms of their semantic relationships. The study thereby investigates an aspect of the necessary relationships among the input sentences in order to observe constructive effects.

Before detailing these experiments, a review of the appropriate literature is provided. The history of the constructive theory is reviewed, along with a descriptive sampling of related contemporary research, analogous developments in linguistics, a record of the specific paradigm and materials of these experiments, and finally, a description of the thematization processes believed to be involved in constructive or integrative phenomena.

Literature Review

General Early Sentence Research. Barlett (1932) is perhaps the most famous of the initial investigators of sentence memory. His work with sentences reflected a more general commitment on his part concerning the proper study of psychology. Barlett felt that the typical experimental materials of his day were too simple and quite possibly artificial. Consequently he chose to use items such as pictures, complex patterns, anecdotes, and prose stories as stimuli in his research. Barlett was interested in memory for structure—a characteristic he believed was shared by all these stimuli.

The typical sentence memory experiment reported by Barlett (1932) involved the method of repeated reproduction of a story. Subjects were read a piece of thematic material twice, then attempted an initial reproduction (recall) about fifteen minutes later, with several subsequent recalls at longer time intervals.
Before discussing the overall results and the interpretation he developed, it might be helpful to review a sample story. It is important to note again that Bartlett believed any such story had a definite form (Woodworth, 1938) or character, represented by its plot or what was sometimes called its atmosphere. Bartlett was interested in how this underlying character was retained in comparison to the details of the prose.

Below is one of Bartlett's most famous stories, an American folk tale entitled "War of the Ghosts:"

"One night two young men from Egulac went down to the river to hunt seals, and while they were there it became foggy and calm. Then they heard war cries and they thought: 'Maybe this is a war party.' They escaped to the shore, and hid behind a log. Now canoes came up to them. There were five men in the canoe, and they said: 'What do you think? We wish to take you along. We are going up the river to make war on the people.' One of the young men said: 'I have no arrows.' 'Arrows are in the canoe,' they said. 'I will not go along. I might get killed. My relatives do not know where I have gone. But you,' he said to the other, 'may go with them.' So one of the young men went up the river to a town on the other side of Kalama. The people came down to the water, and they began to fight, and many were killed. But presently the young man heard one of the warriors say: 'Quick, let us go home; that indian has been hit.' Now he thought; 'Oh, they are ghosts.' He did not feel sick, but they said he had been shot. So the canoes went back to Egulac, and the young man went ashore to his house, and made a fire. And he told everybody and said, 'Behold! I accompanied the ghosts and we went to fight. Many of our fellows were killed. They said I was hit, and I did not feel sick.' He told it all, and then became quiet. When the sun arose, he fell down. Something black came out of his mouth. His face contorted. The people jumped and cried. He was dead."

Rightly understood, the story centers around the ghosts. However, even a casual examination of the reproductions in the Bartlett
(1932) studies indicates that the true plot was usually not perceived (remembered). All mention of the ghosts tends to be omitted in the reports of the English students who served as subjects. Evidently (Woodworth, 1938), the allusion to ghosts was considered an incidental aspect of the story.

For the most part, it appears that each subject reacted in his own way to a story, that is, he or she seemed to form an idiosyncratic conceptualization. In his or her recall of this personalized (selective) storage the subject omits a great many details, modifies facts, adds some information, and generally improves the internal consistency of the story. This individualized memory exemplifies an important aspect of the general constructive point of view, that each subject's cognitive contributions are important determiners of the understanding which is developed (constructed) in comprehension. These effects appear to be even more pronounced across time and repeated reproductions. Bartlett reports that virtually none of the many recalls he examined were completely accurate, though nearly all retained at least a few striking details, even if those particulars were not well fitted into the true plot.

As a result of a long series of studies with several types of structured stimuli, Bartlett concluded that when subjects attempted to reproduce an item such as a story, they constructed their version from limited direct memorial information. Bartlett felt that subjects draw on their general
memory of the world to formulate a story matching a few retained
details plus the well-remembered atmosphere (attitude) of the
passage. Such results led Bartlett (1932) to conclude:

"Remembering is ... an imaginative reconstruc-
tion or construction, built of our attitude towards
a whole mass of organized past reactions or exper-
ience (a schemata or schema) ... and a little
retained detail which commonly appears in image or
language form."

Bartlett's view, which is essentially the basis of construc-
tive versions of sentence memory, ran counter to the conventional
memory theory of his day. The more accepted view, as exemplified
by Ebbinghaus (1885) or Wundt (1907), conceived of memory as
reduplicative, with a distinct trace for every input event.
Woodworth (1938) described this traditional view of remembering
as the excitement of so many definite impressions—a set of
entities which are fixed, lifeless, and have only the capacity
to be re-excited. Bartlett was instead hypothesizing the active,
constructive memory process which results in the development and
storage of what he termed schema. These schemata are not a
chronological record of previous perceptions, each preserved
separately, but rather an integration of those perceptions into
an organized setting relating to some particular sort of exper-
ience (Mattingly, 1975). The structure or the nature of the
constructive process itself is left ill-defined. According to
Bartlett, a comprehender in his or her initial reading somehow
forms a schema which lacks detail, but retains a fundamental
(organizational) attitude. It is clear that Bartlett believed
remembering was literally a constructed justification of this attitude (schema).

Bartlett was by no means the first nor the only researcher to be interested in sentence memory, or to posit a schema-like theory. Herbart (cf Boring, 1957) for example, had much earlier introduced the concept of apperceptive mass, a construct very similar to Bartlett's organized knowledge. Herbart's theory of memory suggested that any learning presupposed the availability of previous memory structures with which the new information could be integrated. Learning was not supposed to be only a passive recording of information, but necessarily involved apperception—the connection of new information with a person's background of experience. Several contemporaries of Bartlett, Crosland (1921) and Lewis (1933) for example, also posited highly similar general theories of memory representations.

One of the most interesting and perhaps the most thorough of these contemporaries of Bartlett was Henderson (1903). Henderson's dissertation at Columbia was a recall study of materials quite similar to Bartlett's prose studies. Henderson was interested primarily in education, and he emphasized the differential retention of this type of material across several age groups. Henderson discussed at length the various processes by which subjects simplify and generalize (condensation) in memory. These processes combine similar statements, modify details to make the stories more meaningful, omit incongruent aspects, and generally make the major theme of the story dominant throughout. It is not
difficult to substitute the constructs of contemporary psychology, and quite reasonably characterize Henderson's analysis as semantic memory (Tulving, 1972) for thematic material. Henderson's article is to be recommended to those interested in such theories, both because of its clarity and captivating prose.

To reiterate the general position of such theories, memory for a set of semantically related sentences is best viewed as a constructive schema, not a set of reduplicative traces. The typical description suggests some kind of actively molded, ever-changing, attitude towards the massed effects of a series of past events. Bartlett (1932) implied that all cognition, including both perception and memory was the result of such constructive processes. Neisser (1967) has interpreted this to mean that all meaning is the result of constructive processes. It is important to note that both initial understanding, and the subsequent use of this information was thought to be a function of the aforementioned schematic constructions. These schemata are also supposed to be constantly elaborated or altered, by either later related inputs or the recall process itself.

According to both Bartlett and Neisser, all recall and recognition involve these schemata— they constitute all the available records of the original experience(s). It is possible for schema to introduce bias and distortion into both the initial encoding and all later retrieval of a given set of information. Neisser (1967) explained that the reproductive changes evidenced by Bartlett and Henderson were clear evidence of the distortion which can be
affected by a schema storage. However, despite the attention paid
to the theoretical construct of schema, neither Neisser nor Bartlett
provide much detail about the nature of this mechanism. Experiment
1 of this report is an attempt to examine how and when the develop­
ment of such integrative effects occurs. This study examines the
effect of related experience on the retention of later input sen­
tences.

**Contemporary Thematic Sentence Research.** The Bartlett (1932)
and Neisser (1967, 1977) theorizing, or what has been here termed
the constructive memory hypothesis, has not gone unchallenged.
Most directly, Zangwill (1972) and Anderson and Bower (1972) have
opposed this model. Each reviews a considerable amount of liter­
ature which appears to contradict the constructive view. To capsu­
lize their views, it is argued that Bartlett's experiments show an
extraordinary high, probably misleading proportion of errors. They
report a number of sentence recall studies following Bartlett that
suggest the most common error in recall is omission, not the dis­
tortion and amalgamation reported by Bartlett.

Both Woodworth (1938) and Cofer (1973) have also commented
that Bartlett's reproduction studies do not seem to have stressed
accuracy—implying that his results are possibly not represen­
tative of recall tasks in general. The point seems to be that in
their opinion, recall is characteristically more accurate than a
schema would allow. Consequently, Zangwill and others have argued
that an individual trace (linguistic object) theory is a more
reasonable and justifiable hypothesis.
Two important qualifying points must be made about the objections raised by Zangwill and Anderson and Bower. The first qualification involves Cofer's contention (1973, 1976) that recall tasks in general may not be the best method for ascertaining how an experience is represented in memory. Using an intuitive rationale derived in part from the work of Kintsch (1967, 1972) and Mandler (1970, 1973), Cofer maintains that recognition tests may be more reasonable (sensitive) tests of memory. While not denying the role of retrieval in recognition, Cofer implies that recognition tests probably allow one to tap more the nature of the trace storage and less its retrieval (Melton, 1963). Cofer notes the amazing sensitivity of recognition memory demonstrated in Shepard and Teghtsoonian (1961) and Shepard (1967). He also discusses the advantage of how recognition tasks lend themselves to response bias analysis such as a forced-choice paradigm and signal detection analysis (Banks, 1970; Green & Swets, 1974). Use of these analytical procedures might aid in separating the effects of memory and response bias in recognition. In summary, Zangwill (1972) and others may be positing a legitimate criticism concerning the inaccuracy of Bartlett's studies, but there may in general, be better ways of asking the questions discussed about the nature of memory. Research which is reviewed later suggests the reasonableness of Cofer's suggestion.

The second qualifying issue involves the type of experimental materials used in either recognition or recall tasks. It appears critically important for the issues at hand whether related
(thematic) or unrelated sentences are studied. The majority of the experiments supporting the linguistic object hypothesis have used carefully chosen, semantically unrelated sentences. For example, Shepard's (1967) often quoted demonstration of remarkable memory involved sentences which were as semantically distinct as possible. The problem is that such effects or observations are not necessarily generalizable due to the unique nature of these sentences. It may well be that the loss of episodic encoding is most probably evidenced when related items are investigated. That is, related materials which more resemble everyday discourse or communication, lend themselves to an integration process. It may prove difficult for an experimental subject to integrate items which have been selected with the goal in mind of having each be as different as possible from all other to be remembered items.

Several theorists (Crothers, 1972; Carroll and Freedle, 1972; Kintsch, 1974; Rummetveit, 1970) have restated this point. Single words or single unrelated sentences may not display phenomena involved in the typical processing of meaningful verbal materials. Shank (1973) has possibly made the most forceful presentation of this view, he argues that paragraphs and connected prose are the best (perhaps only) means of illuminating memory structure.

There have been a few studies in the literature which have directly tested the effect of semantic relatedness in recall tasks. Howe (1970) for example, had subjects read and recall 160 word passages. The subjects recalled an average of 81.5 words, of which only 19.8 (19%) actually appeared in the original paragraph
while 51.7 words were intrusions. This is a large intrusion rate compared to more typical verbal learning studies using unrelated materials (Ausubel, 1963; Kintsch, 1970). Tulving (1972) for one, has interpreted these results to indicate episodic memory for related materials is not very long-lasting. In another study, Jarvella (1970) investigated related sentences, and found similar effects. Jarvella's subjects listened to lengthy prose passages. At random intervals, he interrupted and asked for backward recall. Jarvella found that subjects could at most report only a sentence or two, with only a semantic gist remaining for all earlier material. Such studies seem to indicate that with related sentences, subjects do not evidence a very stable episodic encoding even in recall tasks. Griggs (1974) also reports supportive recall data, but since he was studying Bransford and Franks (1971) type materials, discussion of his experiments is postponed till later in this report.

It is appropriate to emphasize again that the more vocal proponents of a linguistic object hypothesis, researchers such as Zangwill (1972) and Anderson and Bower (1972), have almost exclusively used only minimally related materials. It seems likely that the use of such materials biases experiments in favor of demonstrating episodic encoding. That is, because of the materials, it seems likely that people would stop or remain at this level of analysis.

Consider for example the many cross-over experiments of Anderson and Bower. Their HAM model, one of the most developed
linguistic object theories, calls for each sentence studied to be analyzed into a set of labeled associations between words in the sentence. These labels correspond roughly to the case relations introduced by Fillmore (1966) and now so popular with computer simulators. While HAM's large number and different type of associations represents a considerable elaboration of classical associationalism (Anderson & Bower, 1973), it retains the principle that each sentence is stored as a set of functionally independent associations between the words of each sentence.

The cross-over experiments they report involve presenting subjects with pairs of sentences which share the same object noun. Two such sentences might be:


Anderson and Bower hypothesized that [1] and [2] are each stored as a string of labeled associations. Since these sentences share a word, each sentence's representation in memory should involve an association connection to the word "postman." After studying each sentence embedded in a list of the same type of item, subjects are asked to recall the objects of all sentences when they are provided with various types of cues or prompts. For [1] above, subjects would alternately receive the following types of cues: (a) the sentence's subject (boy); (b) the sentence's verb (bit); (c) the subject and verb of the same sentence (boy bit); and (d) the subject and verb between matched pairs of sentences (i.e., crossed, as boy talked). Anderson and Bower provided some evidence that the
associative logic of HAM predicted the unintuitive results that cross-over cues were the most effective prompt.

This type of evidence for linguistic object theories like HAM involves the presentation of minimally related sentences. These materials and the task demands of accuracy amount to a considerable biasing in favor of episodic encoding. The accuracy demands and the limited semantic relatedness would appear to deter constructive or thematic processes which might exist. It is also true that Foss and Harwood (1975) have argued that all models of the HAM class are inadequate to account for even the results of this type of cross-over experiment. They suggest the particular sentence sets of Anderson and Bower contained unexamined confoundings.

The remaining studies to be discussed in this section are examples of the use of highly related sentences in different types of recognition tasks. The results are in general quite supportive of the constructive hypothesis. The evidence to be presented involved: (i) studies of memory for individual sentences embedded in prose; (ii) studies of context effects in both word and sentence interpretation; and (iii) a few extant experiments on memory for word occurrence in sentences.

Perhaps the most often quoted study on memory for individual sentences embedded in prose is Sachs (1967). She tested subjects' memory for sentences either immediately after presentation, or after 80 or 160 syllables of thematic (prose) material. Sachs found that subjects seem to well remember both the form (syntax)
and meaning of a sentence on the immediate test, but they demon-
strate little memory for form (surface structure) at the longer
retention intervals. It appears that subjects are able to retain
only a semantic gist for sentences (possibly a schema) at the
longer retention intervals. There are many experiments suggesting
this same general finding, that subjects consistently remember
meaning while forgetting the supposedly (by linguistic object
theories) equally important syntactic or form attributes. Such
evidence is available for example in Anderson and McGraw (1973),
Brewster (1975), Bock and Brewster (1974), Dooling and Mullet
(1973), Honeck (1971, 1973a, 1973b), Honeck and Honeck (1976),
Honeck, Riechman, and Hoffman (1975), and Sachs (1974).

One very interesting experiment was reported by Sulin and
Dooling (1974) in which they also tested sentences embedded in
paragraphs. The procedures are somewhat suspect, but the exper-
imental task is quite clever and the results deceptively powerful.
Sulin and Dooling presented identically worded seven-sentence
paragraphs to two different groups of subjects. The experimental
group was told that the passage was about a well known historical
character (e.g., Adolf Hitler or Helen Keller). The control
group was told that the passage was about an imaginary and unknown
person (Ron Huiatt or Laura Daniele for example). The results of
a recognition query indicated that the experimental subjects were
much more likely to falsely recognize new thematic material—sen-
tences which were never presented but were appropriate (though
vague) to the famous person. These subjects made significantly
more recognition errors in trying to remember seven sentences!
The study thus indicates that the pre-experimental knowledge of
a well known character is assimilated in memory when thematic
material is presented. Sulin and Dooling posited that a subject
constructs a schema to represent actually presented materials,
and that this construction also included pre-experimental know-
ledge of the presented topic.

The second major line of research to be reviewed here concerns
the effects of sentential context on the speed and nature of sen-
tence memory. The studies are quite diverse, and the following
should be viewed only as a sampling of various approaches. A
considerable amount of contextual research has been done by John
Bransford and co-workers. These studies are discussed first.

Bransford, Barclay, and Franks (1972) tested whether the
semantic descriptions developed by subjects correspond strictly
to the information specified by a linguistic object analysis.
They discovered that the semantic representation constructed as
a function of reading related sentences included more information
than was directly expressed in the input items. Bransford et al.
(1972) found it was likely that subjects make assumptions about
(infer and remember) the spatial relations among objects men-
tioned in study materials.

Johnson, Bransford, Nyberg, and Cleary (1972) and Johnson,
Bransford, and Solomon (1974) have extended this finding in
showing that subjects may assume the existence of objects not
actually mentioned in a passage—but objects necessary for actions
described in the sentences. The subjects read a small story that was designed to allow inferences about necessary instruments to perform specified acts. For example, subjects might read either [3] or [4], then at recognition, view [5]:

[3] John was trying to fix the bird house. He was pounding the nail when his father came out to help him do the work.

[4] John was trying to fix the bird house. He was looking for the nail when his father came out to watch him and help him do the work.

[5] John was using the hammer to fix the bird house when his father came out to watch him and help him do the work.

The subjects reading [3] were likely to falsely recognize [5], but those hearing [4] were almost always relatively positive that they had not heard [5]. Glanzman and Pisoni (1973) replicated and extended these findings, emphasizing the role of case relations (Fillmore, 1968) in the assumptions that subjects seem willing to make in comprehending a set of sentences. Several other theorists (Barclay, 1973; Fillenbaum, 1974; Frase, 1972; Kintsch and Monk, 1972; Tzeng, 1973) have focused on the inferences made by subjects in detailing spatial relations among an object set. The general result of these studies is a strong suggestion that subjects do not simply store the information explicit in sentences. Instead subjects appear to use the linguistic inputs in conjunction with alinguisitic or cognitive information. The contention frequently presented is that any adequate approach to comprehension must consider the cognitive contribution of a reader as well as the linguistic character of the input sentences.
It is possible to posit that subjects first comprehend (episodically encode) sentences completely, then make any or all cognitive elaborations such as realizing spatial implications, inferring consequences, or any other developments. This version of the processing supposes that whatever these cognitive contributions are, they occur after linguistic comprehension (or linguistic object analysis) takes place. The constructive theory described here and elsewhere (for example, Bransford and McCarrell, 1974; Bransford and Johnson, 1973; Clark, E.V., 1974) suggests that comprehension results only when the reader has sufficient alinguistic (cognitive) information to use the cues specified in linguistic input to create a semantic content which allows understanding. This view could be called a strong version of constructive theory. It does not deny that there are many levels of understanding and memory, but suggests that usual reading strategies involve critical contributions (combinations) of linguistic and alinguistic information in the initial process of comprehension.

This notion of the importance of pre-requisite information has been tested in several ways. For the most part the studies have varied the nature and availability of contextual information. Bransford and Johnson (1973) tried to manipulate subjects' ability to comprehend a set of sentences by varying the availability (mention) of what was termed pre-requisite information about the topic being discussed. Bransford and Johnson presented passages such as the following:
"The procedure is actually quite simple. First you arrange things into different groups. Of course one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run, this may not seem important but complications may easily arise. A mistake that can be expensive as well. At first the whole procedure will seem complicated. Soon however, it will become just another facet of life. It is difficult to foresee any end to the necessity of this task in the immediate future, but then one can never tell. After the procedure is completed one arranges the materials into different groups again. Then they can be put into their appropriate places. Eventually they will be used once more and the whole cycle will then have to be repeated. However, this is a part of life."

There were three context conditions for the presentation of such paragraphs: subjects were told the context (washing clothes) either before reading, after, or not at all. Bransford and Johnson found that the no-topic-group showed quite poor comprehension and recall, as did the subjects who received the information after reading but before recall. The subjects in the topic-before-reading group had significantly higher scores. The experimenters discuss how such results show that having appropriate pre-experimental (contextual) information in memory is not enough, this data or record must be activated in the comprehension process in order to facilitate memory.

Subsequent studies by Johnson, Doll, Bransford, and Lapinski (1974) and Barclay, Bransford, Franks, McCarrell, and Nitsch (1974) elaborate somewhat on both the methodology and theory of these results. These later works provide evidence that familiar,
unambiguous words are interpreted differently depending on their sentential contexts. Again the emphasis is on the subjects constructing and storing information. These experiments concentrate on the differential encoding or interpretation that individual words (and sentences) receive as a result of their context. Haviland and Clark (1974, also Clark & Haviland, 1974) present a very similar finding and theory using a different methodology. Clark and Haviland hypothesize a Given-New strategy to account for the results. Aside from differences in nomenclature and perhaps emphasis, the strategy represents another constructive hypothesis of prose comprehension.

The final sentence memory research to be discussed in this section concerns tests of a comprehender's memory for the word events of the sentences studied. Recall that Sachs (1967, 1974) and related data imply that the memory available for sentences changes over time. The two experiments to be reported here examine the nature of the semantic information said to be available to subjects. Pompi and Lachman (1967) demonstrated that subjects make thematic errors in the retention of prose. Their subjects read a highly thematic passage about an operation for cancer. The individual words used were low associates to that theme. On subsequent recognition tests, thematic words that had not been presented in the passage were falsely recognized. The results suggest that the memory representation with such highly related sentences does not include specification of prior word events.
Johnson-Laird, Robbins, and Belicogna (1974) have also tested how well the words used in a sentence are remembered. The results are less dramatic than Pompi and Lachman (1967), but they do support the same general conclusions. Johnson-Laird et al. examined how well subjects distinguish whether some particular piece of information was conveyed by a noun or a verb. The results suggest that even when listeners can successfully recall the meaning of a sentence, they can remember whether some part of that meaning was conveyed by a noun or a verb only if they know they will be tested for such information. Unless instructions are used to emphasize a verbatim memorization, it appears lexical categories for meaning are quickly forgotten. This non-remembering of individual words may represent serious problems for any simple linguistic object hypothesis.

A constructive theory on the other hand is in fact supported by these data. According to a view such as Bransford and McCarrell (1974), input words are not items of storage, but only abstract constraints which guide the construction of meaning. Thus the emphasis in memory should not be on the individual word and its individual meaning, but on the comprehender and his/her construction. Bransford and McCarrell feel that this conceptualization of words as constraints helps explain how words can have so many different senses of meaning. Consider the word "use." As Glanzman and Pisoni (1973) describe, one can use a rope or a knife or a box, etc., and in each case, what the word "use" means depends on the entities involved in the event "using." Also the same
object can be used for different functions. Thus to use a crate to carry oranges is different from using that crate to stand-on to paint the ceiling. In all cases, the word imposes a set of abstract constraints on the nature of the event, but the referents involved are crucial in creating the semantic interpretation of the relationships being discussed.

When all the studies in this section are considered, it becomes clear that there is at least a beginning of evidence for a constructive theory of how people remember related materials. Certainly a great deal of this type of processing remains considerably vague. The current experiments are an attempt to develop an understanding of some aspects of this processing. It is a hope of this work to come to an increased knowledge of integrative phenomena at least within the Bransford and Franks situation. The next topic of this review describes an analogous development of interest within linguistics, and finally the general paradigm and materials of these experiments are described.

**Developments Within Linguistics.** Until recently the interpretive semantics of Katz and Fodor (1963) and Katz and Postal (1964), later incorporated by Chomsky (1965), represented the only accepted theory of linguistic semantics. This general linguistic theory, sometimes referred to as the Harvard-M.I.T. transformational tradition (cf Weimer, 1974) or the Extended Standard Theory (McCawley, 1974), has a linguistic object emphasis. According to this theory, the grammar of a language (more specifically, its base structure component) generates deep structures
which are input to a semantic component. The formation of these deep syntactic structures is the initial step in language production. The semantic component's role is to interpret the syntactic structure, i.e., to describe their meanings. Before proceeding, two features of the theory should be noted: (a) it employs a deep syntactic structure, distinct, and in some sense prior to, the level of semantic representation; and (b) this syntactic component is the generative element of the system.

Syntactic interpretation of a sentence involves the following mechanisms and processes. First there is a semantic component which consists of two parts. One component or part of the semantic interpreter is a dictionary of lexical elements. A single entry in this dictionary consists of both a grammatical portion, which provides the parts of speech appropriate for each lexical element; and a semantic portion, which represents each of the distinct senses the lexical item has as a part of speech. This dictionary entry is represented by a small set of universal features such as (+) concrete, (+) animate, (+) male, etc. The use of these theoretical feature sets has been criticized by both Bierwisch (1969) and Bolinger (1965) because of its arbitrariness, but that argument will not be developed here.

The second component of the semantic system was a set of rules (i.e., projection rules) which select the appropriate sense of each lexical item in the sentence in order to derive the correct readings for each grammatical structure in turn. This process of internal (within a sentence) disambiguation is a complex one; it
is accomplished by a series of amalgamations in which all possible readings for the lexical item in a grammatical structure are systematically compared. The acceptable readings of each grammatical structure are then selected, combined with those of the next structure, and so on, until the entire sequence has been associated with a single set of readings. The dictionary entries provide all the necessary information to interpret a sentence, the projection rules systematically combine some subset of all possible entries to derive the sentence's actual meaning(s).

One of the most controversial aspects of interpretive semantics theory is the distinction it makes between the subject's knowledge of his/her language (language competence) and his/her knowledge of the world. Katz and Fodor (1963) and more recently Katz (1971, 1972) hypothesize that a theory of meaning should concern itself only with the former type of knowledge. They contend that an individual's knowledge of the world is too idiosyncratic and unsystematic to have a place in a theory of how the ideal speaker understands the meaning of a sentence presented in isolation. Katz and Fodor explicitly argue against the notion of including reference or other alinguistic (in their model) knowledge within a theory of semantics. This emphasis on linguistic information and individual sentence semantics exemplifies the linguistic object nature of this thesis. Katz (1972) not only excludes these ideas, but also relegates them to a yet to be conceived of linguistic communication theory (although some might consider Rommetveit, 1968, 1974; or Halliday, 1971 as a beginning).
However, even if people are able to understand sentences without a context being specified by the experimenter, it is not unreasonable to imagine that the comprehenders conjure an appropriate context of their own design and make use of it in trying to understand the supposedly isolated sentences. Katz does not recognize this possibility. Rommetveit (1974) is especially convincing in arguing against the Katz (1972) position.

Recently, several linguistics have proposed alternatives to the theory of interpretive semantics. While their individual theories differ somewhat in particulars, these researchers all reject the rigid distinction made in interpretive semantics between a language user's linguistic and cognitive knowledge. Each of these theories somehow allows for an integration of these two kinds of knowledge at various levels of language processing. It is this recent tradition, which will be termed generative semantics, that is considered most important for the present review. The development of this type of theory is an analogy to constructive theorizing in psychology. The two points of view cannot be equated, but both represent an emphasis on an active and semantically based language process. In fact, a few cognitive theorists with psycholinguistic interests such as Olson (1970, 1972) and Reid (1974) have presented theories almost identical to generative semantics from a more psychological point of view. The emphasis is also that subjects make use of reference and other situational contexts to be able to make semantic judgments about the meaning of almost any statement in
everyday usage. These theories also attempt to explain the role of semantic memory in language. The emphasis is clearly on reference, context, and knowledge of the world.

In linguistics, the inclusion of these cognitive mechanisms came through the work of Bach (1968), Lakoff (1971), McCawley (1968), Postal (1971), and others—it constitutes a major revision in the relation between syntax and semantics and in how semantic relations are to be defined in language. There has been a steady progression from the purely syntactic conception of deep structure in Chomsky (1957, 1965) tradition through a quasi-semantic position in abstract syntax (Ross, 1967) to a purely semantic one in generative semantics (Mattingly, 1975; Ross, 1975). Basically generative semantics proposes that the deep structure of a sentence is its semantic representation. The generative component is no longer considered to be the basic syntactic structure; in fact, most generative semantic theories eliminate this device entirely. Instead the semantic component is converted directly into surface syntactic representations by an as yet unspecified set of transformational rules.

Generative semantic grammars thus appear to suggest a somewhat radical departure from the more traditional transformational theory. Chomsky (1970) has stated that the generative semantics position is merely a notational variant of his theory. Greene (1972) for one, agrees with Chomsky, but others such as Weimer (1973) argue emphatically that while the two views may in some sense be equivalent, i.e., able to be stated in the same
uninterpreted calculus, the two positions differ considerable in explanatory adequacy. For present purposes, what is crucial is the formal renouncement of the sharp division between linguistic and cognitive psychology. The important point is not that the generative semantics position be different from Chomskyan theory, but that it conceives of semantic representations to be a function of both linguistic (within a sentence) factors and the language user's extra-linguistic knowledge. This merger provides the basis for an investigation of the nature of semantic processing in the communication and retention of more complex information than single word lists or simple categorization tasks.

One such generative semantics theory will be discussed in order to provide an example of this general movement in linguistics. Of course, this coverage will be by necessity quite short and incomplete, but hopefully it will provide a bit more perspective on these developments. Fillmore (1966, 1968, 1971a, 1971b) introduced one of the more influential versions of generative semantics. His theory, more formally called case grammar, rather heavily emphasizes the role of extra-linguistic considerations in understanding language. In his theory, the base component is again not a deep syntactic unit, but rather a set of case relations which may only be indirectly evident in the surface structure of a sentence.

The case relations referred to:

"... comprise a set of universally, presumably innate concepts which identify certain types of judgments that human beings are capable of making about
the events which are going on around them, judgments about such matters as who did it, who it happened to, and what got changed” (Fillmore, 1968).

Fillmore thus posits a theory of semantic structure which has the charming feature of being based on aspects of an external event that are perceptually available even to a pre-language child (Halliday, 1975).

The fundamentals of the system, the cases, are a description of the observable semantic relationships between objects in the environment. Linguistic mechanisms need only be applied to these concepts to communicate in language form.

Fillmore’s theory is by no means complete at this time, it is still being developed or altered as circumstances demand (McCawley, 1974; Ross, 1974). In particular, the transformational rules which would convert a set of case relationships into a surface structure have as yet not been clearly enumerated. Fillmore’s most recent (1971a) list of cases included:

Agent (A): the instigator of the event
Counter Agent (C): the force or resistance against which the action is carried out
Object (O): the entity that moves or changes or whose position or existence is in consideration
Result (R): the entity that comes into existence as a result of the action
Instrument (I): the stimulus or immediate physical cause of an event
Source (S): the place from which something moves or comes
Goal (G): the place to which something moves
Experience (E): the entity which receives or accepts or experiences or undergoes the effect of an action.

The base component of any sentence of Fillmore’s grammar comprises
two components: (i) the proposition which is a tenseless set of relationships involving verbs and nouns (i.e., contains a verb and at least one case category); and (ii) the modality constituent which includes negations, tense, mood, etc.

Fillmore (1971a) has also identified the lexicon or form of dictionary of individual word meaning which would be appropriate to a case grammar. Fillmore asserts that the combination of case relations a word can enter into and the presuppositions which it requires for usage in a message more or less define that word. He does suggest at one point, that there may exist some semantic features stored with each word, but the important theoretical constraints on word meanings are implied by its potential cases and presuppositions. This model appears to be yet another way to suggest that the words within a semantic memory may all be defined in terms of other words in the system. Fillmore provides some examples of simplified lexical entries in such a system, two of them are:

[6] pilot profession: A of (V 0 A) where V is to navigate, with the presupposition that O is an air vessel

[7] knife use: I of (V O I A) where V is cut, with the presupposition that O is a physical object.

Some additional details of these entities were provided, but since Fillmore does not suggest that even those completely encompass a lexical entry, these comments are not provided here. The emphasis in this type of definition is clearly that the meaning of a word is a matter of its function in a semantic analysis (Halliday, 1974,
1975). Described in another way, the emphasis is not on semantic features, but on the restrictions governing the use of a word on its assumed denotation (Perfetti, 1972; Perfetti & Goodman, 1975). This kind of definition and the entire focus of case grammar is very close conceptually to the constructive views in the psychological literature reviewed earlier.

Many current attempts at computer simulation of language memory have borrowed heavily from the case grammar of Fillmore. Kintsch (1970, 1974) for one has adopted a case grammar propositional base as a lexical deep structure. Also Rummelhart (1974), and Shank (1973, 1974) have programs which represent memory for sentences or sometimes paragraphs as a network of case-like relationships among concepts. Frijda (1972) provides an excellent review of this class of computer simulation models. The most notable aspect of the models is their lack of supportive data. These computer simulations are almost entirely based on intuition. The exceptions being the recent work of Rummelhart and Lindsay (1974) and the LNR group plus Shank's (1974) most recent reports. Most importantly for this review, these models do not make the strong assumption that one's knowledge of language can be separated from the role this knowledge plays in the comprehension and production of language. It is worthy of note that most of these types of models ignore the issues dichotomized in the competence-performance distinction of Chomsky (1965) and other linguistic object theorists.

In sum, the focus of generative semantics is very similar to
developments in psychology which emphasize sentential context and holistic (schematic) representation of information. There are other types of theories, such as message analysis by Rommetveit (1968, 1974), speech acts by Searle (1971), and others which could be presented as further examples of a dynamic, cognitive point of view. One of themes bringing these data together is their reaction against the major source of evidence from the interpretive model, judgments of grammaticality, ambiguity, and anomaly. The critics seem puzzled by this choice of data, and have challenged it in several ways. The generative semanticists and related theorists (e.g., Fillmore, 1968; McCawley, 1970; Halliday, 1975; Rommetveit, 1974) have emphasized that people do not speak in order to be grammatical per se, but to convey meaning. Their resultant emphasis on intention, context, meaning, and message should thereby be clear.

The Current Paradigm. The final section of this introduction deals with a particular two-stage recognition paradigm introduced by Bransford and Franks (1971, 1973). This methodology and the experimental materials serve as the basis for the research in this paper. The experiments described later are an attempt to clarify the integration phenomena demonstrated in the typical Bransford and Franks situation. These experimenters and their associates have repeatedly shown that if a subject is presented with a long series of sentences, some subset of which possess a common theme, a subject appears to actively construct an idea (or schema) to represent all the information suggested in the
related (though randomly presented) sentences.

The usual Bransford and Franks experiment (1971 for example) involved presenting subjects with 24 sentences, 6 of which related to 1 of 4 so-called linguistic ideas. That is, a subject studied four sets of related sentences (i.e., on the same theme), and then attempts to recognize these items embedded in a long list of sentences. Each theme is referred to as a linguistic idea, or a semantic episode (Rumelhart, Norman, and Lindsay, 1972), which can be described in a single, complex idea-sentence. One such linguistic idea might be, "The rock which rolled down the mountain crushed the tiny hut at the edge of the woods."

Initially the subjects were presented with 24 simple sentences of varying length which were derived from four such complex sentences. Each of these simpler sentences represents a distinct subset of the total idea. A subset sentence of size one (length-1 or 1-1) for the aforementioned idea sentence might be "The rock rolled down the mountain" or "The hut is tiny." Notice that an 1-1 sentence contains one piece of semantic information. A subset sentence of size two (1-2) and three (1-3) respectively might be "The rock crushed the hut at the edge of the woods" and "The rock which rolled down the mountain crushed the hut at the edge of the woods." Again, subjects were presented with six such subset sentences of varying lengths (two 1-1, two 1-2, and two 1-3) for each of four ideas. In most of the Bransford and Franks experiments, subjects were never presented with the original complex idea sentences (a so-called 1-4).
After the initial study phase, subjects were presented in a recognition test, a series of possible semantic subsets and the complex idea sentence. Bransford and Franks, using a Bartlett-like schema logic, predicted subjects would not only be able to integrate the to-be-remembered (TBR) items, but that such an integration was habitual and unavoidable. Consequently they hypothesized that subjects would not be able to distinguish which sentences were presented. Also, since they believed the integrated representation would include all studied information, they predicted that subjects would most likely select as previously seen (old), the very long, complex sentences which best represent the entire idea. These predicted results were observed. The complex L-4 sentences which were never presented, but were the best representation of the linguistic idea, were in fact the most likely to be selected as old. The results are all the more dramatic when they are compared to the considerable sensitivity previously demonstrated by Shepard (1967).

In summary, Bransford and Franks (1971, 1972) have shown that if a subject is asked to recognize highly related sentences, he or she will more likely choose the composite idea constructed (or the most similar available sentence), than any of the individual items actually presented. The theoretical explanation posited by Bransford and Franks proceeded from the notion that a person's comprehension of a sentence is a function of both his knowledge of the words involved in the events described, and the contexts in which the items are perceived (semantic relations).
It is important to re-emphasize that Bransford and Franks stress the importance of a generalized (across sentences), constructed memorial representations—a semantic encoding of these types of inputs.

The nature of the schema or constructed-whole remains rather vague. Certainly Bransford and Franks have provided several replications of the gross effects, but there have been few attempts to detail the processes involved. The limit of their analysis appears to involve the rather molar construct of a schema. There have been several related studies by other investigators which are described below, but even a brief review demonstrates that these studies add little to an understanding of linguistic integration. Often these studies fail to converge on the topic because they are based on very different interpretations of the meaning of the original Bransford and Franks experiments.

The one exception to the stipulation, and perhaps the most interesting study of the set, was done by Peterson and McIntyre (1973). These experimenters noted the considerable differences in recognition memory reported by Bransford and Franks (1971) and Shepard (1967), and they attempted to study more directly the influence of semantic relatedness in a recognition task. Peterson and McIntyre wishes to control procedural differences, equate two groups of sentences for length and study time, and vary the relatedness of the items to be studied. They presented two groups of subjects with comparable lists of sentences (length, number, etc.) except that one set was constructed of semantically
independent items (not related) while the other set was semantically related as in Bransford and Franks. Their results, both in terms of recognition scores and a signal detection analysis ($d'$) indicated that accurate recognition memory for sentences occurs only when semantic intersential relations (relatedness) are not present. Their experiment represents a demonstration of the importance of semantic relatedness for the integration effects of Bransford and Franks (1971). It is perhaps worth noting at this time, that the issue of how and in what way the sentences must be related, remains unanswered. The Peterson and McIntyre (1973, 1974) items, as the original Bransford and Franks sentences, were constructed from the same set of words for each theme. That is, all related sentences are combinations of one set of words for each theme. Thus the semantic relatedness evidenced so far involves an integration of sentences which not only share meaning (references), but also word events. This point is discussed again at the end of this section.

The remaining research to be discussed on integration is a series of papers by Singer and Rosenberg (1973) and Katz (1973; Katz, Atkeson, & Lee, 1974; Katz & Grenewald, 1974). The primary reason for reviewing these papers is to indicate the different conceptualization sometimes given to the linguistic idea phenomena. The reader should be aware that these experiments do not converge on the nature of constructive analysis.

Singer and Rosenberg (1973) re-analyze earlier Bransford and Franks (1971, 1972) data and describe that their often used
confidence-rating scoring procedure is ambiguous. Bransford and Franks had shown that recognition ratings (a combination score which represents confidence rankings and recognition probability) are highest for L-4 sentences, and then decrease in a linear fashion from L-3, L-2, and L-1 sentences. Bransford and Franks have stressed this linear effect in providing evidence for integration.

Singer and Rosenberg contended that it is possible to explain these data in terms of a grammatical analysis which they felt was more parsimonious than integrated schema. Singer and Rosenberg posited that subjects are simply more likely to say old to any sentence which contains the main semantic clause of each theme, and that the longer sentences are more likely to contain this key element. Singer and Rosenberg (1973) saw no need to hypothesize more complicated explanatory mechanisms. While this point was well taken, there had been a confounding of sentence length and probability of containing a main semantic clause, this argument has been well countered by another Franks and Bransford (1974) paper. Very briefly, Franks and Bransford replicated their previous results while controlling for the main semantic unit confounding. They also noted that there is a need for further research to provide more detailed accounts of the structure and processes involved in this integration. The present experiments are such an endeavor.

The final studies to be discussed were reported by Katz and several co-workers. Katz has contended that these experiments
refute the notion of integration or abstract linguistic ideas. Before detailing these works, it should be made explicit that Katz's interpretation of the Bransford and Katz research is quite different from the present author. Katz has generally argued that the integration phenomena have been mislabeled by Bransford— that the phenomena are explainable without reference to semantics. Katz appears to be trying to prove that the linguistic or language aspects of Bransford's experiments are misleading. Katz appears to be saying that language is not necessary to observe integration. His contention is poorly aimed. Bransford and others have continually argued that the integration phenomena being studied may be observed with language materials, or may other types of stimuli. It is overly simple and erroneous to criticize Bransford's research because integration effects can be demonstrated with other types of materials. As a point of fact, the first Franks and Bransford (1971) study involved visual form prototypes and pattern-schema. However, Katz completely ignores this research and the almost constant Bransford (and his co-workers) references to Bartlett's (1932) schema. The entire series of Katz experiments are not reasonable pertinent unless one accepts his confused definition of Bransford's work.

It is also true that each of Katz's experiments rests on a different, yet equally tenuous assumption about the materials being used in each case. In the first such experiment, Katz (1973) utilized active and passive forms of the same sentence base (deep structure) and instructed his subjects only to judge if new
sentences had the same meanings as prior studied items. His results indicated a breakdown in the typical integration effects found by Bransford and Franks. Essentially he found that subjects were not as willing to integrate active and passive sentence information. Katz's conclusions rests critically on the validity of his assumption that active and passive forms of a deep structure have identical meaning. While it is true that some transformational grammars do define them as such, there is a store of psychological investigations which suggest otherwise (see Anisfeld & Kornblut, 1973; James, Thompson, & Baldwin, 1973; and Watson & Johnson-Laird, 1972 for reviews).

In a second study, which supposedly disproves the semantic nature of the integration results, Katz, Atkeson, and Lee (1974) have studied nonsense sentences in a Bransford and Franks type paradigm. These sentences are items which are composed of nonsense words tagged with English inflections (bound morphemes). Katz's major assumption in using these materials is that pseudo-sentences have no meaning! Katz made this explicit claim. He presented this type of item, and then observed a mild integration effect. Katz then concluded that the Bransford and Franks' claim of semantic construction was disproved. That is, if integration necessarily involved semantics (Katz's assumption), and these sentences have no semantics (Katz's unsupported contention), yet they display an integration, then Bransford must be wrong (Katz's conclusion). The inferences of Katz appear incorrect for at least two reasons: (i) the experimental materials of Katz et al. are
in all probability not meaningless (Epstein, 1961, 1962; McCawley, 1970; O'Connell, 1970); and (ii) even if the sentences were somehow completely meaningless, they possess a form as indicated by the syntactic cues—which could be used as a basis for integration. The experiment of Katz and his associates seems again to miss the point that constructive theory as outlined by Bransford and Franks (and others) views integration as more than a semantic process.

In his final, and perhaps most interesting study, Katz and Grenewald (1974) demonstrated a type of limit on the integration of linguistic ideas. In this experiment, they presented approximately twenty sentences (L-1 only) to subjects. The experimental lists contained at least six different L-1 items pertaining to each of four different themes. Like Bransford and Franks (1971), they then provided a series of recognition trials in which subjects viewed related items varying in size from L-1 to L-4. All of the items in the recognition phase were new—a weak procedure borrowed directly from early Bransford experiments. Katz hypothesized that if the integration phenomena are habitual and dominant (Franks and Bransford's description), then subjects should display the same type of integration effects under these circumstances as in any other. The results of Katz and Grenewald show that subjects clearly do not choose as old any item above L-1 length. The subjects were very unlikely to select any L-2 through L-4 items. Katz uses this data to argue that subjects were not integrating the material.

Here again, there are alternative conclusions possible. It is
significant that even though subjects were only selecting $L-1$ sentences as old, these choices were all false recognitions since all test sentences were new. There must be some basis for this willingness to choose a new sentence as having been previously seen. Perhaps the subjects had formed an integrated semantic representation. But why then, if there is some kind of integrated memory, did the subjects not select any of the longer test items which should more resemble that constructed whole.

One possible answer is that some sort of selector mechanism (Underwood & Schultz, 1960) or response pooling was responsible. The subjects may well have been cognizant that no timulus sentence longer than $L-1$ had been presented; therefore, they might have used that rule to eliminate all recognition foils larger than $L-1$. It would have been interesting to see the results of a control group with $L-1$ study items and similar sentences, but in which subjects were asked to respond on the basis of meaning and not form. Had these subjects demonstrated an ability to recognize (select) sentences which represented the themes involved, Katz's contention would be much more compelling.

Throughout this series of experiments, Katz posited a guessing strategy to account for the typical Bransford recognition pattern. However, the need for this rather complex explanation depends on the acceptance of a set of suspect assumptions and on the use of his weak methodology in which all recognition items were new. If old and new sentences are tested in recognition, and this has been done (Bransford & Franks, 1973; Franks and Bransford, 1972, 1974),
the integration effects remain and the viability of Katz's explanation does not.

In fairness to Katz, he has at least contributed several interesting questions in pondering the limits of the integration effects. His point that little is known about what kind of experiences (sentences) a subject must have (observed) in order to evidence an integration, is a challenging observation. What type of sentences must be observed and what is the time course of this process are the questions addressed in the present experiments. The materials and paradigm used in these studies are quite similar to Bransford's work. This was done deliberately in the hope that the results of these experiments would clarify the nature of the processes and help establish a reasonably understood paradigm for even more sophisticated tests of sentence memory. The theory is generally that comprehension of sentences involves prior knowledge, intention, context, and task demands in combination with the input materials to control discourse processing (Fredericksen, 1976a, 1976b, 1976c). A comprehender is thought to construct a semantic interpretation as a function of such variables. The experiments were an attempt to help characterize the nature of these processes.
EXPERIMENT 1

This experiment attempted to evaluate the effects of experience on sentence memory by studying the idea derived sentences in a modified continuous recognition paradigm. The study and test materials were highly similar to Bransford and Franks (1971, 1972) materials in order to facilitate comparisons of the results. The methodology and procedures adopted resulted from a consideration of two important points: (i) the hypothesis of a strong constructive theory about the encoding of a thematic input; and (ii) a possible frequency of semantic-unit (component) confounding in previous Bransford and Franks research. Each concern is briefly reviewed before the methodology is detailed.

Consider the question of how or when a schema is developed, and what consequences this has for subsequent related inputs. Essentially the issue concerns the effect of context or experience on memory for thematic material. Single, unrelated sentences such as studied by Johnson (1965) or Shepard (1967) appear to be treated as independent "objects" in memory and are well remembered. However, constructive theory suggests that as a reader experiences multiple sentences on a theme, he or she seems to form a schema to represent that theme. The comprehender is then unlikely to retain any individual (episodic) representations of subsequent related inputs. The current experiment manipulates how much experience or
context a subject has had with a theme before he or she is asked to recognize some sentence belonging to that theme. It is predicted that a subject's ability to distinguish new or old thematic items will decrease as a function of experience with that theme. There are two issues involved: how much experience is necessary to develop a schema, and when subsequent input is thought to be incorporated into that schema. This experiment allows a test of the necessary experience with a context to evidence a loss of episodic memory (i.e., an inability to distinguish new and old related inputs). This experiment provides this evidence with a procedure based on a strong constructive theory's (the Bransford-Franks-McCarrell view) assumption that even the initial encoding or representation is inseparable from the appropriate schema.

Previous research such as Sachs (1967, 1974) demonstrated that the available memory for sentences in prose shows quite definite changes over time plus experience. The current experiment is distinct in holding the retention interval for a sentence constant and somewhat brief (approximately 12 seconds) while testing for the effects of several levels of prior experience. Again, the major variable of interest is how many related (same theme) sentences a reader has attempted to learn before he receives a recognition test.

The second major consideration involves a possible confounding in the original Bransford and Franks articles. In general, they reported that there were no reliable differences in subjects' ability to distinguish between old versus new (but
thematically appropriate) items. Also, and most important here, Bransford and Franks typically reported a linear effect in that longer thematic items were the most probably chosen as old, even if never before viewed. There is a possible confounding in the methodology of these studies involving the frequency of the semantic-unit components (see earlier review). Previous studies have not structured the acquisition materials so that the frequency of occurrence of each semantic unit (literally, the information contained in each L-1 sentence) was counterbalanced across length and type of recognition sentence. It is possible that the integration evidence reported is at least in part due to uncontrolled frequency effects. For example, if subjects were making their judgments on the basis of frequency of each test item's semantic units, and the thematic items (especially the longer new or old items) were most likely to contain the most frequent units, the overall results would falsely appear to support an integration hypothesis.

The present study attempted to avoid this confounding by examining only memory for L-2 sentences, and counterbalancing so all tested semantic components have occurred equally often. Thus any changes in the available memory of L-2 items can be attributed to the effects of experience with the theme(s) being used. L-2 sentences then are the critical test items. The TBR items prior to all L-2 test sentences have been counterbalanced for unit frequency, that is, all tested pieces of semantic information have occurred equally often for all types of test sentences (new, old, or noncase).
It is prohibitive in terms of complexity and length to do such counterbalancing for all lengths \((L-1\) and \(L-3\)) of test item. However, in order to avoid any obvious strategy or over-emphasis on these \(L-2\) items as a class, subjects were also tested equally often on \(L-1\) and \(L-3\) items. Although these recognition trials were included to disguise the critical nature of the \(L-2\) sentences, these data were eventually analyzed to test for a length of test sentence effect.

According to the present design, a subject received a recognition trial for some TBR \(L-2\) item after having viewed either 1, 2, 3, 4, or 5 related (same theme) TBR sentences. The recognition test for a TBR item occurred at a constant lag (number of intervening events) for all levels of testing. During each clearly marked test trial, a subject was asked if a sentence was new or old, and how confident he or she was of that decision. There were three types of test sentences: (1) old--items previously presented for study; (2) new--items derived from the same linguistic ideas as the TBR sentences, but that have never been seen; and (3) noncases--sentences constructed from semantically (thematically) inappropriate combinations of previously seen words. Type of test sentence and level of experience were completely crossed. Again, the methodology called for an equal number of \(L-1\) and \(L-3\) tests, but no attempt was made to control the frequency of the component propositions or units.

In summary, the major question asked was does the experience of highly related material influence the memorial representation
established for subsequent sentences? The prediction derived from a constructive theory is that a reader is more likely to display integrated representations after having studied multiple inputs on some theme. Therefore, as a reader views more related inputs, and thereby has more probably (or possibly more firmly) constructed a holistic representation, he or she should be less likely to display episodic memory. It was thus predicted that a reader's ability to distinguish new and old items should decrease across levels of experience. Extrapolating from previous studies, it was expected that memory for new \( L^2 \) items should closely resemble old \( L^2 \) items at higher levels of experience. The experiment attempted to define the amount of experience necessary to observe these types of integration effects. Such knowledge is interesting in itself and is extremely useful in developing subsequent tests of the nature of the schematic representation hypothesized for sentences.

**Method**

**Subjects**

The subjects were 36 undergraduates enrolled in introductory psychology courses at the Ohio State University. They participated in order to receive partial credit in those courses.

**Design**

The study was a complete \( 2 \times 3 \times 5 \times 3 \) factorial design. The factors were (1) list: with two identical forms being used; (2) type of recognition test item: old, new, and noncase; (3) the level of experience with an idea or theme of a recognition test:
that is, the number (1, 2, 3, 4, or 5) of related sentences studied prior to a recognition test; and (4) order of test blocks: three equivalent orders of the test sentence materials were developed.

Materials

All materials consisted of 15 sets of thematically related sentences constructed in the following manner: (1) each idea set was derived from one of the 15 complex sentences shown in Table 1; (2) each of these complex sentences was constructed to convey all the semantic relationships (units) specified by four simple, declarative sentences. These simple sentences and all derivatives of them had been arbitrarily selected, with no special regard for their theoretical status according to extant linguistic theories. The most important aspect of the complex sentences was that they could be meaningfully and easily analyzed into specifiable derivatives (subset items).

Each of the 15 sets of materials began with the complex $L^{-1}$ or idea sentence and the $4$ $L^{-1}$ sentences which were embedded within it. Also there were at least four $L^{-2}$ sentences formed by embedding two of the most simple $L^{-1}$ items and at least three $L^{-3}$ sentences which represent a combination of three of the $L^{-1}$ items for that idea. A sample of the subset sentences for one idea is presented in Table 2. These sets of sentences provided all the TBR and thematic recognition sentences for the identical experimental lists. The lists differed only in theme assignments, that is, which theme was used to test particular levels of the other variables. Noncase item were constructed as needed by
Table 1.
The Complex or L—1 Sentences

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The fat man sitting at the table in the restaurant ate the greasy hamburger.</td>
</tr>
<tr>
<td>2</td>
<td>That night, the college students drank the cold beer in the downtown bars.</td>
</tr>
<tr>
<td>3</td>
<td>The elderly man walked up the mountain smoking his pipe.</td>
</tr>
<tr>
<td>4</td>
<td>The little girl running to her parents fell down the stairs.</td>
</tr>
<tr>
<td>5</td>
<td>The scared cat running from the barking collie jumped on the couch.</td>
</tr>
<tr>
<td>6</td>
<td>The warm breeze blowing from the sea stirred the cool evening air.</td>
</tr>
<tr>
<td>7</td>
<td>The handsome prince dreamed of saving a beautiful damsel in distress.</td>
</tr>
<tr>
<td>8</td>
<td>The tired boy walked quietly though the large park in the center of the city.</td>
</tr>
<tr>
<td>9</td>
<td>The clowns at the circus made all the orphan children laugh.</td>
</tr>
<tr>
<td>10</td>
<td>Last year's fire destroyed most of the large homes in the valley.</td>
</tr>
<tr>
<td>11</td>
<td>The rocks which slid down the mountain crushed the peasant's hut beside the forest.</td>
</tr>
<tr>
<td>12</td>
<td>The ants in the cupboard finished the sweet jelly on the shelf.</td>
</tr>
<tr>
<td>13</td>
<td>The old woman resting on the couch read the story in the newspaper.</td>
</tr>
<tr>
<td>14</td>
<td>The rusty car pulling the trailer climbed the steep hill.</td>
</tr>
<tr>
<td>15</td>
<td>The tall tree in the front yard shaded the man watching television.</td>
</tr>
</tbody>
</table>
### Table 2
A Sample Set of Related Sentences for One $L^{-4}$

<table>
<thead>
<tr>
<th>$L^{-4}$</th>
<th>The old woman resting on the couch read the story in the newspaper.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L^{-3}$</td>
<td>The woman resting on the couch read the story in the newspaper.</td>
</tr>
<tr>
<td></td>
<td>The old woman resting on the couch read the story.</td>
</tr>
<tr>
<td></td>
<td>The old woman read the story in the newspaper.</td>
</tr>
<tr>
<td>$L^{-2}$</td>
<td>The woman resting on the couch read the story.</td>
</tr>
<tr>
<td></td>
<td>The woman read the story in the newspaper.</td>
</tr>
<tr>
<td></td>
<td>The old woman rested on the couch.</td>
</tr>
<tr>
<td></td>
<td>The old woman read the story.</td>
</tr>
<tr>
<td>$L^{-1}$</td>
<td>The woman was resting on the couch.</td>
</tr>
<tr>
<td></td>
<td>The woman was reading the story.</td>
</tr>
<tr>
<td></td>
<td>The story was in the newspaper.</td>
</tr>
<tr>
<td></td>
<td>The woman was old.</td>
</tr>
</tbody>
</table>
combining L-1 sentences across ideas in a way that the resultant sentences has sensible meaning but were inappropriate to any theme of the study.

Procedure

Subjects were tested in groups of 5 to 10. The instructions, TBR, and test materials were presented in written form. When each group assembled, a participant was handed a packet which contained all experimental materials. A page of instructions was read aloud as subjects followed. The emphasis of these directions was to stress accuracy in recognition and to answer all procedural questions. It was explained that some sentences would be highly similar, but the precise form and meaning of each TBR sentence should be retained as well as possible.

Each subject's packet also contained a booklet consisting of 3 sets of approximately 50 related sentences. Each of these sets contained approximately 35 TBR items and 15 recognition tests. The test trials were distinctly marked by a series of asterisks which appeared above and below a test item. Subjects were directed that they did not have to memorize test items, but only had to indicate whether or not they had previously studied the item (new/old) and their confidence on a five point scale. The general instructions defined and exemplified both of these indices. All recognition tests pertained only to the TBR items within the same set or block.

Across all three sets, a subject thus saw 45 recognition tests which allowed his/her memory to be tested for every type
and length of recognition item at all levels of experience. The theme used for each of the tests and the ordering of all tests was randomized in several ways described below.

List Construction: Counterbalancing

Each subject viewed a list consisting of three different sets of sentences. As mentioned earlier, these materials allowed all type, level of experience, and length of test sentences to be tested for every subject. It was important to adequately counterbalance several different factors to test all the variables which were manipulated. The sections which follow explain the nature and ordering of the recognition tests, as well as the TBR items which intervened between these tests. Several constrained randomization procedures were used to counterbalance the order of tests, the theme used for each type of test, and the unit frequency of TBR materials for all critical L-2 tests.

List. In order to identify and balance any theme effects, two different but equivalent lists were constructed. The themes shown in Table 1 were randomly assigned to one of three sets except that (a) each set contained five ideas judged by this author to be semantically distinct; and (b) each idea was used only once per subject.

Also, to avoid any order of sets effect, three different orders of the sentence blocks were established, and subjects randomly assigned to an ordering. Further a practice set of 25 TBR-like related sentences was given to all subjects initially to control for warm-up effects.
TBR Items. Each set of a list contained 35 TBR items composed of 7 sentences from each of 5 themes within that set. The seven TBR items from each theme consisted of one $L^{-1}$, two $L^{-3}$, two $L^{-2}$, and two $L^{-1}$ sentences, the order of which was randomized with the demands imposed by the scheduled recognition tests. This sequencing was also restricted in that: (a) no two sentences from the same theme occurred in immediate succession; (b) $N$ sentences from each theme were presented before any theme had $N+1$ sentences; (c) if an old recognition test of theme $R$ was specified on some trial $K$, then the $K-3$ item was identical to the recognition test sentence; and (d) if a recognition test for a new item was called for, the previous theme items must have specified all the semantic information in that new item (though of course, not in the same form).

Recognition Tests. Each set or block of sentences contained 15 recognition tests. All recognition tests followed the target by two intervening presentations (either test or TBR items) or approximately 12 seconds. The 15 tests within any set were randomly arranged with the constraints that (1) each theme was tested three times; (2) all levels of experience were tested three times; (3) all types of test items (new, old, and noncase) were tested five times; (4) each length of test item ($L^{-1}$, $L^{-2}$, and $L^{-3}$) was tested as equally often as possible; and (5) all $L^{-2}$ recognition tests involved tests of equally frequent semantic unit components. As a function of these restrictions, the nature and order of the required tests was counterbalanced within blocks. All such tests were responded by writing in the test booklets.
Results

Initially a $2 \times 3 \times 5 \times 3$ factorial analysis of variance (ANOVA) was performed on the recognition rating data for all L-2 test items. The factors in this analysis were (L) List—with two levels, lists A and B; (O) Order—with three levels, each representing a different order of the three sets of test sentences; (E) Experience—with five levels, literally the number of related theme sentences studied before a recognition test of an L-2 sentence; and (T) Type of recognition test sentence—with three levels, old, new, and noncase. A complete source table is available for this ANOVA in Appendix A.

Before describing the major results of this analysis, a brief description of the dependent variable being analyzed may be in order. The recognition rating index was introduced by Bransford and Franks (1972). It represents a combination of the probability of a correct recognition with the confidence subjects have in that labeling. A response of new becomes a score from 1-5 with 1 representing most confident in that new decision; a response of old is treated as a 6-10 score with 10 meaning most confident the item is old. Thus a subject's response of "new-2" or "old-3" is transformed to a recognition rating of 4 and 8 respectively. The higher the recognition rating (to 10), the more confident a subject is that an item is old. Correspondingly, the lower (to 1) a recognition rating, the more confident a subject is that an item is new.

This initial $2 \times 3 \times 5 \times 3$ analysis of L-2 items alone
indicated significant main effects of experience, $F(4,120) = 31.25$, $p < .01$, type of recognition test sentence, $F(2,60) = 502.43$, $p < .01$, and a significant experience X type interaction, $F(8,240) = 33.61$, $p < .01$. Neither list nor order were found to have significant effects, and consequently the tabled means provided (see Table 3) are summed across both of these variables. None of the other possible interactions displayed any significant effects.

Clark (1973) has argued that data from experiments using language materials should be analyzed treating the verbal stimuli as a random variable. The validity of using this type of analysis for this design is questionable (Clark, Cohen, Smith, & Kepple, 1976; Forster & Dickinson, 1976; Moeser, 1976; and Wilke & Church, 1976). However, an analysis of the data using Clark's proposed quasi-F ratio (see also Winer, 1971) found identical results—a significant effect of experience, $F'(4,4) = 161.2$, $p < .01$, type of recognition test sentence, $F''(2,2) = 1612$, $p < .01$, and the interaction of experience X type, $F'''(8,8) = 84$, $p < .01$.

The mean recognition rating for the five levels of the experience main effect were 3.92, 4.86, 5.16, 5.64, and 6.71 for Levels 1 through 5 respectively. An orthogonal polynomial trend analysis indicated a significant linear trend ($df = 1,120$, $p < .01$). The data thereby lend themselves to interpretation in support of both a strong and a weak version of constructive theory. The mean recognition rating of Level 1 of experience is lowest, and a reliable trend (increase) is found across the subsequent levels of experience. Apparently, subjects are more accurate in judgments at the
Table 3
Mean Recognition Ratings in 2 X 3 X 5 X 3 ANOVA

<table>
<thead>
<tr>
<th>List Means</th>
<th>List A 5.24</th>
<th>List B 5.28</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order Means</td>
<td>Order 1 5.17</td>
<td>Order 2 5.21</td>
</tr>
<tr>
<td>Experience Means</td>
<td>Level 1 3.92</td>
<td>Level 2 4.86</td>
</tr>
<tr>
<td>Type Means</td>
<td>Old 8.65</td>
<td>New 5.87</td>
</tr>
<tr>
<td>Experience X Type</td>
<td>Old 8.22</td>
<td>New 1.89</td>
</tr>
<tr>
<td></td>
<td>Level 2 9.47</td>
<td>New 4.11</td>
</tr>
<tr>
<td></td>
<td>Level 3 8.14</td>
<td>New 6.17</td>
</tr>
<tr>
<td></td>
<td>Level 4 8.14</td>
<td>New 7.69</td>
</tr>
<tr>
<td></td>
<td>Level 5 9.28</td>
<td>New 9.50</td>
</tr>
</tbody>
</table>
lower levels of experience, and get more confused with new/old thematic decisions with more experience.

A Newman-Keuls test was used to clarify the main effect found for type of recognition test sentence. The mean recognition rating for old, new, and noncase type items were 8.76, 5.87, and 1.26 respectively. The Newman-Keuls showed all types to be significantly different from each other. Note also that the experience X type interaction is displayed in Figure 1. The corresponding means are provided in Table 3. Again, the data provide support for the effect of experience with a theme even when the retention interval is rather short (12 seconds).

A second major analysis was performed to investigate the importance of length of a recognition test sentence. Essentially the previously discussed data were re-analyzed along with a fifth factor, length (H) so that subjects' responses to L-3 and L-1 sentences were also examined. A complete source table for the 2 X 3 X 5 X 3 X 3 ANOVA which resulted in displayed in Appendix A. The results are somewhat complex and lengthy. To simplify this presentation, first the main effects will be discussed in some detail, and then the numerous significant interactions are described.

The main effects found by this analysis were experience ($F(1,120) = 36.21, p<.01$), type of recognition test sentence ($F(2,60) = 763.28, p<.01$), and length of the recognition sentence ($F(2,60) = 55.23, p<.01$). Neither list nor order had any reliable effects. Again Quasi-$F$ ratios were developed (see
Figure 1
Mean Recognition Rating by Level of Experience
previous arguments concerning their appropriateness) which indicated the same pattern of results: significant effects (all $p<.01$) for experience, $F''(4,4) = 11.8$, type, $F''(2,2) = 155.8$, and length, $F'' = 301.6$.

Newman-Keuls comparisons performed on each of these main effects helped define these data. The means for each of these main effects are provided in Table 4. Each variable is considered in turn. First, for experience, the only significant difference ($p<.05$) was between Level 1 and Level 5—a pattern which appears to suggest that multiple experiences with thematic inputs were important to observe constructive effects across all sentence lengths. Again, one must be concerned that the $L-3$ and $L-1$ response data are confounded to some extent by component frequencies. The Newman-Keuls performed on the type of test sentence data showed that old, new, and noncase were significantly different from each other. Finally, the length analysis indicated the only significant difference (comparison) to be $L-3$ from $L-1$.

For present purposes, the most important results derive from the experience and length effects, data which generally support constructive predictions.

There were several important interactions (the means of which are presented in Tables 4 and 5) which should be mentioned. This ANOVA showed the following interactions to be reliable (all $p<.01$): list by experience, $F(4,120) = 3.05$; list by type, $F(2,60) = 4.80$; experience by length, $F(8,240) = 8.48$; type by length, $F(4,120) = 17.56$; experience X type X length, $F(16,400) = 13.88$; and list X
Table 4
Main Effect and Some Interaction Means in 2 X 3 X 5 X 3 X 3 ANOVA

<table>
<thead>
<tr>
<th>Experience Means</th>
<th>Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>4.39</td>
<td>5.13</td>
<td>5.24</td>
<td>5.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type Means</th>
<th></th>
<th>Old</th>
<th>New</th>
<th>Noncase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>7.81</td>
<td>5.76</td>
<td>1.26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Length Means</th>
<th></th>
<th>L-3</th>
<th>L-2</th>
<th>L-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5.50</td>
<td>5.08</td>
<td>4.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List X Experience</th>
<th>Level</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>List A</td>
<td></td>
<td>4.44</td>
<td>4.33</td>
<td>5.06</td>
</tr>
<tr>
<td>List B</td>
<td></td>
<td>3.57</td>
<td>4.45</td>
<td>5.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List X Type</th>
<th>Type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Old</td>
<td>New</td>
<td>Noncase</td>
<td></td>
</tr>
<tr>
<td>List A</td>
<td></td>
<td>8.13</td>
<td>5.55</td>
<td>1.28</td>
</tr>
<tr>
<td>List B</td>
<td></td>
<td>7.48</td>
<td>5.97</td>
<td>1.24</td>
</tr>
</tbody>
</table>
Table 5
Interaction Means from 2 X 3 X 5 X 3 X 3 ANOVA

<table>
<thead>
<tr>
<th>Experience X Type</th>
<th>Level of Experience</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>6.79</td>
<td>7.87</td>
<td>7.49</td>
<td>7.94</td>
<td>8.93</td>
<td></td>
</tr>
<tr>
<td>New</td>
<td>3.64</td>
<td>4.22</td>
<td>6.72</td>
<td>6.59</td>
<td>7.63</td>
<td></td>
</tr>
<tr>
<td>Noncase</td>
<td>1.57</td>
<td>1.07</td>
<td>1.19</td>
<td>1.19</td>
<td>1.27</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience X Length</th>
<th>Level of Experience</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-3</td>
<td>5.33</td>
<td>5.46</td>
<td>5.31</td>
<td>5.78</td>
<td>5.63</td>
<td></td>
</tr>
<tr>
<td>L-2</td>
<td>4.00</td>
<td>4.35</td>
<td>4.91</td>
<td>5.52</td>
<td>5.63</td>
<td></td>
</tr>
<tr>
<td>L-1</td>
<td>2.69</td>
<td>3.35</td>
<td>5.18</td>
<td>4.43</td>
<td>5.56</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience X Type X Length</th>
<th>Level of Experience</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-3</td>
<td>6.39</td>
<td>8.78</td>
<td>8.06</td>
<td>8.06</td>
<td>9.06</td>
<td></td>
</tr>
<tr>
<td>L-2</td>
<td>8.22</td>
<td>8.22</td>
<td>7.83</td>
<td>7.94</td>
<td>9.00</td>
<td></td>
</tr>
<tr>
<td>L-1</td>
<td>5.78</td>
<td>6.61</td>
<td>6.58</td>
<td>7.83</td>
<td>8.72</td>
<td></td>
</tr>
<tr>
<td>New</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-3</td>
<td>8.16</td>
<td>6.39</td>
<td>6.50</td>
<td>8.17</td>
<td>6.67</td>
<td></td>
</tr>
<tr>
<td>L-2</td>
<td>1.61</td>
<td>3.83</td>
<td>5.78</td>
<td>7.50</td>
<td>9.56</td>
<td></td>
</tr>
<tr>
<td>L-1</td>
<td>1.16</td>
<td>2.41</td>
<td>7.89</td>
<td>4.11</td>
<td>6.67</td>
<td></td>
</tr>
<tr>
<td>Noncase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-3</td>
<td>1.44</td>
<td>1.22</td>
<td>1.39</td>
<td>1.11</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>L-2</td>
<td>1.16</td>
<td>1.00</td>
<td>1.11</td>
<td>1.11</td>
<td>1.33</td>
<td></td>
</tr>
<tr>
<td>L-1</td>
<td>1.11</td>
<td>1.00</td>
<td>1.06</td>
<td>1.33</td>
<td>1.31</td>
<td></td>
</tr>
</tbody>
</table>
order X experience X type X length, \( F(32, 480) = 1.84 \). The cell means for all these interactions are displayed in Tables 4 and 5. The pattern of responses to new items across experience is especially interesting when contrasted with old and noncase items. Any further discussion or interpretation of this pattern is postponed until presentation of the second experiment.
EXPERIMENT 2

This study was designed to examine the importance of using the same words in the semantically related sentences of a Bransford and Franks (1971) paradigm. Previous studies demonstrating integration effects, i.e., a loss of episodic encoding, have consistently used different combinations (surface structures) of the same words to form all related sentences. The question involved here is whether these integration phenomena are dependent upon the same words, or are they generalizable to sentences which contain different words, but still share the same meaning (semantic referents).

There have been some previous studies which have attempted to describe the nature of memory for words in sentences. As reviewed earlier, Johnson-Laird, Robbins, and Velicogna (1974) and Johnson et al. (1974) suggested that the memory representations are of a deeper character than word events. However, Brewer (1975) and Peterson and McIntyre (1975) have found that people are somewhat able to discriminate synonyms from the words presented in study sentences. Such data suggest that a comprehender may have a type of response pooling or a selector mechanism (Underwood & Shultz, 1962) concerning the occurrence of words. The present study investigated whether sentences which use different words to describe a possible event would evidence integration effects.

A linguistic object hypothesis of sentence memory must suggest
that semantically related sentences with different word events
should be more discernible than typical Bransford and Franks'
materials. In fact, an associative model such as HAM (Anderson
& Bower, 1972) predicts that sentences with more different words
should not be very confusable. Thus sentences such as "The elderly
man walked up the mountain" and "The old gentleman climbed the
steep hill" should not be readily integrated. The general logic
of this position, something true of most linguistic object
theories, is that because each sentence is stored in its entirety,
and since each sentence is composed of different key words, there
should be records available to allow discriminations of the items
presented.

A constructive, semantic encoding however, predicts that all
thematic events which share referents should become integrated
in memory. The constructive process is thought to involve ref­
erence and semantic relationships cued by different words, not
just the words themselves. In a very real sense, a sentence is
believed not to exist in memory at all (Weimer, 197; Rommetveit,
197; Smith, 1975). Instead, a sentence is used only as a set
of clues or instructions to modify the structure of memory in order
to convey the deep, underlying components that suggest meaning. A
comprehender's grasp of meaning of sentences and language in general
is supposed to be a constructive act. Smith (1975) in particular
emphasized that meaning is in the person and not in the sentence—it
is constructed from a set of cues in linguistic form. To borrow a
term introduced by Perfetti (1972), thematization should occur
across the sentences in question here. Perfetti noted that a paragraph or group of sentences that identified a certain Dr. Jones and relates that he is a surgeon, is thematizing a referent which could be realized by the expressions Dr. Jones, the surgeon, the doctor, he, and other noun phrases which identify that referent. Constructive theory predicts that as long as the meaning of thematic material is maintained, the sentences should be integrated in memory.

In summary, the following experiment was designed to test the effect of using sentences which have the same approximate meaning, while lessening the amount of word-event redundancy. The procedure is the more standard two-stage recognition paradigm, similar to that used by Bransford and Franks. Following a constructive materials rationale, it was predicted that subjects would display very similar integration effects with this type of material.

In order to accomplish this, subjects were presented with related sentences for two ideas which shared a common theme. Subjects studied sentences for four such pairs of ideas. For example, they were presented subset sentences of various lengths for the ideas "That night, the college students drank the cold beer in the downtown bars." and "That evening, the underclassmen finished the frosted ale in the town's taverns." Then in recognition, subjects were asked to make new/old decisions about both new and old related items, noncases, and something termed a "hybrid" test sentence. Hybrids were constructed by combining L-1 sentences of the two idea sets which constituted a theme. For
example, a \( L-2 \) hybrid for the theme described above might be "That evening, the underclassmen drank the beer." The subjects' responses to these hybrids were critical evidence of integration. If subjects have formed a schema to represent a theme (combined idea sets), then hybrids should be responded to as would old or new related items. If however, hybrid items are treated as noncases, then subjects are not integrating sentences which only share meaning but use different words to express that meaning.

**Method**

**Subjects**

The subjects were 50 undergraduates at the State University of New York College of Arts and Sciences at Cortland, New York who were concurrently enrolled in an introductory psychology course. They participated in order to receive partial credit in that course.

**Design**

The study was a \( 2 \times 4 \times 3 \) factorial design. The factors were: (1) list: \( A \) and \( B \); (2) type of recognition test item: new, old, non-case, and hybrid; and (3) length of the recognition test item: \( L-2 \), \( L-3 \), and \( L-4 \).

**Materials**

Materials consisted of a number of English sentences constructed in the following manner. First, two sets of 4 complex sentences from the first experiment were chosen. The idea sentences were labeled \( A, B, C, D, E, F, G, \) and \( H \). Next, these eight idea sentences were matched with eight new sentences which shared meaning, but not words. That is, the eight idea sentences \( A' \) through
were constructed so that there were eight themes (cognitive episodes) with two idea sentences describing each theme. Theme Set 1 for example, concerned the demolition of a small dwelling. The two idea sentences involved were $A$—"The rocks which slid down the mountain crushed the peasant's hut beside the forest;" and $A'$—"The avalanche which swept down the cliff destroyed the villager's cottage at the edge of the woods." The sixteen idea sentences are shown in Table 6.

Next, each of these 16 complex sentences (representing 8 theme sets) were broken down into a complete set of subset sentences (an idea set) which consisted of one $L^{-4}$, three $L^{-3}$, four $L^{-2}$, and four $L^{-1}$ related sentences (see Experiment 1 for further details about an idea set). Thus for each theme, there were $2^4$ related sentences, 12 belonging to each idea sentence of that theme.

Finally, sets of noncase and hybrid items for each theme were constructed. Noncase items were defined in Experiment 1 also. Here again they consist of combinations of $L^{-1}$ sentences across theme sets. Noncase items of $L^{-4}$, $L^{-3}$, and $L^{-2}$ were developed. Hybrid items were constructed by combining $L^{-1}$ sentences within a theme set, that is, across the two ideas belonging to one theme set. For example, an $L^{-1}$ of idea $A$ might be combined with an $L^{-1}$ of $A'$ to form a hybrid $L^{-2}$ for Theme Set $A$. Such a sentence might read, "The rocks crushed the villager's cottage." Hybrid sentences of length $L^{-2}$, $L^{-3}$, and $L^{-4}$ were constructed.

Procedure
Table 6
Theme Sets for Experiment 2's Materials

<table>
<thead>
<tr>
<th>Theme 1:</th>
<th>Theme 2:</th>
<th>Theme 3:</th>
<th>Theme 4:</th>
<th>Theme 5:</th>
<th>Theme 6:</th>
<th>Theme 7:</th>
<th>Theme 8:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A -- The rocks which slid down the mountain crushed the peasant's hut beside the forest.</td>
<td>B -- The ants in the cupboard finished the sweet jelly on the shelf.</td>
<td>C -- That night, the college students drank the cold beer in the downtown bars.</td>
<td>D -- The elderly man walked up the mountain smoking his pipe.</td>
<td>E -- The little girl running to her parents fell down the stairs.</td>
<td>F -- The scared cat fleeing from the barking collie jumped on the couch.</td>
<td>G -- The warm breeze blowing from the sea stirred the cool evening air.</td>
<td>H -- The handsome prince dreamed of saving a beautiful damsel in distress.</td>
</tr>
<tr>
<td>A' -- The avalanche which swept off the cliff destroyed the villager's cottage at the edge of the woods.</td>
<td>B' -- The small insects in the cabinet ate all the tasty jam on the ledge.</td>
<td>C' -- That evening, the underclassmen finished the frosted ales in the town's taverns.</td>
<td>D' -- The old gentleman climbed the steep hill puffing his tobacco.</td>
<td>E' -- The young child running to his Mom and Dad tripped down the steps.</td>
<td>F' -- The frightened kitty running from the loud dog leaped on the sofa.</td>
<td>G' -- The mild wind off the ocean rippled the chilly dusk calm.</td>
<td>H' -- The dashing young hero longed to rescue a lovely princess in danger.</td>
</tr>
</tbody>
</table>
**Acquisition.** Subjects were tested in groups of 5 to 10. It was explained that a list of sentences would be read aloud, and they were to remember as many of the sentences as possible. Each subject heard one of two equivalent lists. The lists differ only in which four theme sets were used to derive all TBR and test sentences. Within a list there were 12 sentences related to each theme set as follows: (i) 6 sentences from one idea of the theme of the following lengths: two $L-3$, two $L-2$, and two $L-1$; (ii) 6 sentences from the matching idea sentence of that theme of the following lengths: one $L-4$, one $L-3$, two $L-2$, and two $L-1$. In summary then, subjects viewed 48 sentences in acquisition, 12 each of 4 theme sets or 6 from each of the 8 idea sentences involved.

All 48 recognition sentences were read aloud approximately one every 5 seconds. The items were randomly presented with the stipulation that: (i) no two consecutive sentences pertained to any one theme set; and (ii) no one theme had $N+1$ sentences presented before each had $N$.

**Recognition.** Subsequent to the acquisition phase, subjects were presented with an appropriate 48 item recognition list in the same manner as acquisition. Each test sentence was read aloud, and the subjects had approximately seven seconds to respond new or old, and to indicate their confidence in that decision.

The recognition list consisted of 12 items pertaining to each of the 4 theme sets studied in acquisition. Each of these 12 item sets were broken down as follows. First, three of the sentences were old items from each theme set. These items were one $L-4$, one
L-3, and one L-2 item. Decisions as to which idea to use for any particular test were done on a random basis as possible. Secondly, a new L-4, L-3, and L-2 sentence were presented from each theme set. Thirdly, three hybrid items of matching length were included for each theme set. And finally, three similar noncase items were included for each theme set. All 48 test items were randomly presented. This series of tests allows a subject to be tested with L-4, L-3, and L-2 length items for new, old, noncase, and hybrid items.

Results

A 2 X 3 X 4 factorial analysis of variance was performed on both the recognition rating data and the number of correct recognitions. Each of these analyses will be explained in turn. For both, the factors are: (L) -- list with two levels, A and B; (T) -- type of recognition test sentence, with four levels: old, new, noncase, and hybrid; and (H) -- length of recognition test sentence with three levels, L-4, L-3, and L-2.

A source table for the ANOVA on recognition rating scores (a detailed description of this index is available in the first experiment) is presented in Table 11 in Appendix A. The ANOVA indicated a significant effect for type of test sentence, $F(3,144) = 829, p<.01$, and a significant type X length interaction, $F(6,288) = 9.76, p<.01$. Neither list, length of test sentence, nor any of the other interactions were significant.

Again, a quasi-ANOVA analysis as proposed by Clark (see earlier discussion) was performed, and resulted in the same general pattern
of results. There was a single significant main effect for type, $F'(3,3) = 79, p < .01$, and a lone significant type X length interaction, $F'(6,6) = 20, p < .01$.

Newman-Keuls comparisons on the main effect of type showed significant differences between noncase and all others, as well as between new and both hybrid and old (all $p < .05$). The mean recognition rating for each type of test sentence (minimum = 4, maximum = 40) was 4.78, 27.56, 29.14, and 29.39 for noncase, new, hybrid, and old types respectively.

The pattern of results seem to suggest subjects were both reasonably accurate and confident in their labeling of old and noncase items, especially the later. However, new and hybrid items were consistently confused as old. The significantly different comparison between new and hybrid type items, with hybrids receiving the higher ratings, is especially interesting. It lends evidence to the notion that subjects are reliably more apt to treat a hybrid as old than a new test item. It would appear that hybrid items are being well accepted as TBR materials—that subjects have formed a schema representing a theme set, not just an idea set.

The significant type X length interaction is displayed in graphic form in Figure 2. No theoretical explanation for this pattern is apparent. The individual cell means and standard deviations for all combinations of test, type, and length are available in Table 7.

A second major analysis of variance was performed on the
Figure 2

Mean Number of Correct Recognitions by Sentence Type
### Table 7
Mean Recognition Ratings for List, Type, and Length of Test Sentence

<table>
<thead>
<tr>
<th>Length of Test Sentence</th>
<th>(L^{-4})</th>
<th>(L^{-3})</th>
<th>(L^{-2})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>List A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>26.9 (5.6)</td>
<td>31.5 (5.9)</td>
<td>28.1 (5.0)</td>
</tr>
<tr>
<td>New</td>
<td>30.0 (5.0)</td>
<td>25.3 (4.0)</td>
<td>26.9 (5.5)</td>
</tr>
<tr>
<td>Noncase</td>
<td>5.2 (2.8)</td>
<td>4.8 (1.6)</td>
<td>5.3 (3.0)</td>
</tr>
<tr>
<td>Hybrid</td>
<td>29.9 (5.9)</td>
<td>29.7 (6.7)</td>
<td>29.0 (4.6)</td>
</tr>
<tr>
<td><strong>List B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>29.1 (7.4)</td>
<td>32.2 (3.7)</td>
<td>28.6 (5.9)</td>
</tr>
<tr>
<td>New</td>
<td>30.4 (5.4)</td>
<td>25.2 (5.6)</td>
<td>27.5 (5.7)</td>
</tr>
<tr>
<td>Noncase</td>
<td>4.3 (0.8)</td>
<td>4.2 (0.4)</td>
<td>5.0 (2.4)</td>
</tr>
<tr>
<td>Hybrid</td>
<td>28.3 (5.9)</td>
<td>29.4 (5.5)</td>
<td>28.6 (7.7)</td>
</tr>
</tbody>
</table>

1The minimum possible score is 4, and the maximum 40. The figures in parentheses are standard deviations for the associated means.
the number of correct recognitions. A source table for this ANOVA is presented in Table 12 in Appendix A. The dependent variable in this analysis represents how many times (out of four) a subject correctly recognized each combination of type and length of test sentence. It is important to realize that a correct recognition involves different responses for the different type of test sentences. That is, a correct response for any old item is "old," while a correct response to all hybrid, new, and noncase items is "new." Any tendency on the part of subjects to mislabel new and hybrid sentences could be construed as evidence for a constructive memorial representation.

This number of correct responses' analysis indicated a main effect for type of test sentence, $F(3,144) = 4.54$, $p < .01$, length of test sentence, $F(2,96) = 11$, $p < .01$, and the type X length interaction, $F(6,288) = 5.25$, $p < .01$. Each of these effects is addressed separately below. Neither list nor any of the other interactions were found to be significant ($p > .05$).

A Newman-Keuls analysis was subsequently performed on the type of test sentence main effect. The mean number of correct recognitions per type (maximum possible was 4) were 3.9, 3.0, 1.1, and 0.9 for noncase, old, new, and hybrid type items respectively. The Newman-Keuls indicated that all paired-comparisons were significantly different ($p < .05$). That is, each type of test item was found to be reliably different from all other types. The data seem easily interpretable as evidence for a constructive hypothesis. Subjects appear to be very accurate in dealing with noncase and old items, but to be
confused by both new and hybrid items. While subjects do seem to
have a preference for old above both hybrid and new related items,
both of the later are reliably treated as different from noncases.
The low mean number of correct recognitions for new and nybrid items
indicates that subjects reacted to these types as if they believed
them to be TBR materials.

The main effect for length of test sentences was also further
analyzed using a Newman-Keuls test. The mean number of correct
responses for each length of test item were 2.2, 2.5, and 2.1 for
lengths $L^-4$, $L^-3$, and $L^-2$ respectively. The only significant
differences were found to be between $L^-3$ length and both $L^-4$ and
$L^-2$ ($p<.01$). The linear effect suggested by the Monte Carlo analysis
of Bransford and Franks was not evidenced here.

The length X type interaction is displayed in Figure 3. The
most noticeable feature of this data is the rise in the number of
correct recognitions associated with both new and old $L^-3$ sentences.
No intuitive explanation seems obvious. The means and standard
deviations for all combinations of list, type, and length are pre­
sented in Table 8.

Once more the quasi-$F$ ratios suggested by Clark (1973) were
developed for this data. This procedure found essentially identical
results, a significant difference between lengths, $F''(2,2) = 96$,
$p<.05$, type of recognition test sentence, $F''(3,3) = 940$, $p<.01$,
and a significant interaction between type and length, $F''(6,6) =
32$, $p<.01$.

Finally a signal detection analysis was performed on the
Figure 3

Recognition Rating of All Types and Lengths of Test Sentences

KEY
- - - Old
- - - New
- - - Hybrid
- - - Noncase
Table 8

The Mean Number of Correct Recognitions for All Combinations of Tests

<table>
<thead>
<tr>
<th>Length of Test Sentences</th>
<th>L-4</th>
<th>L-3</th>
<th>L-2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>List A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>2.7 (0.9)</td>
<td>3.4 (0.8)</td>
<td>2.7 (0.8)</td>
</tr>
<tr>
<td>New</td>
<td>0.7 (0.7)</td>
<td>1.5 (0.6)</td>
<td>1.1 (0.8)</td>
</tr>
<tr>
<td>Noncase</td>
<td>4.0 (0)</td>
<td>4.0 (0)</td>
<td>4.0 (0)</td>
</tr>
<tr>
<td>Hybrid</td>
<td>0.9 (0.9)</td>
<td>0.9 (0.8)</td>
<td>0.9 (0.8)</td>
</tr>
</tbody>
</table>

| **List B**               |     |     |     |
| Old                      | 3.0 (0.9) | 3.5 (0.6) | 2.9 (0.9) |
| New                      | 0.8 (0.8) | 1.5 (0.8) | 1.2 (0.7) |
| Noncase                  | 4.0 (0)   | 4.0 (0)   | 3.9 (0.2) |
| Hybrid                   | 0.9 (0.9) | 1.0 (0.7) | 1.1 (1.0) |

1Note that a correct recognition involves different responses to different types of test sentences. That is, a correct response to an old item is "old," a correct response to a hybrid, new, or noncase is "new."
subjects' recognition data. The primary argument for its use has been the provision for a discrimination index (d') which is independent of any response bias factors. Bias here refers to the fact that, independent of the memorial representation available, subjects may not use all responses equally often. Any possible response bias should not be confused with d', thus possibly providing a more sensitive measure of memory. In this situation, d' is thought to provide a clear method of contrasting the memory trace strength available for old, new, and hybrid items when each is compared with noncases (noise). The d' values derived for old, new, and hybrid items were 3.03, 2.90, and 3.03 respectively. Therefore this index also provides evidence that old and hybrid items are being treated in a very similar fashion—being judged alike in these recognition tests, with the news very similarly held, but less emphatically than either old or hybrids. Again, the data seem to support the constructive hypothesis. Any further discussion is postponed until the final discussion.
CONCLUSIONS

The general cognitive model of sentence memory proposed earlier suggested that both the linguistic and constructive theories of sentence memory provide reasonable explanation of some aspects of the cognitive processes involved. To briefly summarize, the linguistic or linguistic object theory would predict each sentence is understood, and retained as an independent trace or process. Each studied or read sentence would thus result in some type of independent memorial record. This individual representation would thereby allow the recognition or recall of quite specific and accurate information about each studied input item, or at least, few if any combinatory confusions. The linguistic object theory suggests something similar to what Tulving (1972) has termed episodic information should be available for to-be-remembered sentences. Several research investigations in the past few years have studied memory for unrelated sentences, and provided substantial support for the general linguistic object model or theory.

The constructive notion of sentence memory posits that the memorial encoding of sentences, and perhaps even the initial input development of meaning involves much more than episodic or single sentence information. Some versions of this constructive position might well argue that a single sentence's information is never available alone in any real time sense. The general constructive theory suggests that a sentence is at least quickly parsed and processed
to some semantic level, using information derived from the currently attended item as well as contextual information and knowledge about the world. Thus, a sentence is thought not to be processed or stored in isolation, but combined or understood as a part of a schematic representation in accordance with other cognitive (possibly alinguistic) information. A constructive theory thus predicts a reader or comprehender would not be able to well retain individual sentence detail over time. A considerable amount, if not all, of an individual sentence record is predicted to be lost in the construction of a representative schema, especially with (though not limited to) semantically related sentences. Research previously discussed such as Bartlett (1932), Sachs (1967), and many others provides support. The constructive theory emphasizes then, a deeper level of processing, not unlike Tulving's (1972) semantic memory in which specific items are integrated (when possible) with the multi-dimensional structure of long-term memory.

The combination, cognitive theory proposed earlier, suggests that both the linguistic object and constructive theories accurately portray a portion of the processing when a reader comprehends sentence materials. This more general cognitive notion purports that both episodic and semantic information are available at different times or phases of the understanding and memory process. Under particular conditions of time and experience (possibly intentional strategy as well), both the linguistic object and constructive theories are useful in accounting for behavior. This combination theory acknowledges the concept of situational task demands (Neisser,
1977) and emphasizes the complex and sophisticated cognitive apparatus available to the simplest of readers.

For present purposes, what has been studied is one type of general sentence materials in two different paradigms to support and decipher some aspects of this combined cognitive approach to sentence memory. The overall model contends a linguistic object type of encoding, displaying episodic knowledge of the input, should be available at an initial or perhaps shallow level of memory load or experience. However, it is hypothesized that typical reading strategies and involvement with multiple inputs of thematic material will lead to the development of a deeper or semantic (possibly schematic) representation. With such multiple inputs, it was hypothesized that a loss of episodic information would occur.

Experiment 1 appears to substantiate this notion, that both episodic and semantic encodings may be evidenced with linguistic idea type materials (Bransford and Franks, 1971), depending on the number of related sentences experienced. Subjects were able to display very accurate sentence recognition as long as only a single sentence from a theme had been studied. However, as more related sentences were studied, subjects made progressively more confusions in recognition. Subjects began to make reliably more confusions between new and old items. The type of error necessitated by an acceptance of a noncase was very, very uncommon. Also the significant level of experience by type of test sentence interaction appears to result from the recognition rating of new test items increasing sharply as experience with the theme increased. The
overall conclusion from this experience supports the cognitive model. Initially subjects are able to make quite accurate new-old decisions, however, that discrimination is quickly reduced with additional thematic experience. The more thematic items studied seemed to result in an increased likelihood that a new but related (schema acceptable) will be judged as old, even with this study's relatively short 12-second retention interval.

Experiment 2 suggests the memorial representation of a constructive process is a schema or plan or network which may not be limited to a single word set, but emphasize some sort of referential information. Subjects were presented with sets of items derived from thematic ideas which shared referent action but not major words (verbs or nouns). That is, sentences were studied which described the same events, but used two distinct word sets. In recognition, subjects were presented with hybrid test items, sentences which mixed these two previously distinct word sets relating to the same referents. The major question concerned whether hybrid test sentences would be confused as old since they did combine unique word sets sharing a deeper meaning. The data, in general, suggest subjects are unable to discern that never before seen hybrid combinations have not been seen before. Subjects appear to accept these hybrid combinations of two referentially related word sets as old. It should be noted that the linear effect of test sentence length often shown with these type of sentence materials was not evidenced here.

It does appear that the shared referents of the two theme
sentences in each idea set result in a shared or combined idea or schema, with the level of encoding possibly not word specific. Therefore test sentences which are viable by this more basic (deeper) representation, are accepted as previously presented items. A schema in this situation, might not just be confined to a single word set, but a combination of word sets which describe the same or similar events or referents. The nature of this schema is not yet available, but this data provides a suggestion that some representation of the referents, perhaps semantic features or some equivalent, may be an important aspect of this memory coding. Certainly it could be argued such a memorial process in terms of reference would be more efficient in processing language materials.
Table 9

Source Table for 2 x 3 x 5 x 3 ANOVA on Recognition Rating for L-2s

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>List (L)</td>
<td>0.18</td>
<td>1,30</td>
<td>0.18</td>
<td>.03</td>
</tr>
<tr>
<td>Order (O)</td>
<td>5.52</td>
<td>2,30</td>
<td>2.76</td>
<td>.44</td>
</tr>
<tr>
<td>L X O</td>
<td>2.95</td>
<td>2,30</td>
<td>1.47</td>
<td>.24</td>
</tr>
<tr>
<td>Subjects/L X O</td>
<td>186.11</td>
<td>30</td>
<td>6.20</td>
<td></td>
</tr>
<tr>
<td>Experience (E)</td>
<td>454.04</td>
<td>4,120</td>
<td>113.51</td>
<td>31.51*</td>
</tr>
<tr>
<td>L X E</td>
<td>2.81</td>
<td>4,120</td>
<td>.70</td>
<td>.19</td>
</tr>
<tr>
<td>O X E</td>
<td>7.59</td>
<td>8,120</td>
<td>.95</td>
<td>.26</td>
</tr>
<tr>
<td>L X O X E</td>
<td>12.72</td>
<td>8,120</td>
<td>1.59</td>
<td>.44</td>
</tr>
<tr>
<td>E X Ss/L X O</td>
<td>435.78</td>
<td>120</td>
<td>3.63</td>
<td></td>
</tr>
<tr>
<td>Type (T)</td>
<td>5022.45</td>
<td>2,60</td>
<td>2511.22</td>
<td>502.43*</td>
</tr>
<tr>
<td>L X T</td>
<td>3.11</td>
<td>2,60</td>
<td>1.56</td>
<td>.31</td>
</tr>
<tr>
<td>O X T</td>
<td>10.76</td>
<td>4,60</td>
<td>2.69</td>
<td>.54</td>
</tr>
<tr>
<td>L X O X T</td>
<td>3.12</td>
<td>4,60</td>
<td>0.78</td>
<td>.16</td>
</tr>
<tr>
<td>T X Ss/L X O</td>
<td>299.89</td>
<td>60</td>
<td>5.00</td>
<td></td>
</tr>
<tr>
<td>E X T</td>
<td>899.39</td>
<td>8,240</td>
<td>112.42</td>
<td>33.61*</td>
</tr>
<tr>
<td>L X E X T</td>
<td>10.61</td>
<td>8,240</td>
<td>1.33</td>
<td>.40</td>
</tr>
<tr>
<td>O X E X T</td>
<td>44.29</td>
<td>16,240</td>
<td>2.77</td>
<td>.83</td>
</tr>
<tr>
<td>L X O X E X E X T</td>
<td>37.49</td>
<td>16,240</td>
<td>2.34</td>
<td>.76</td>
</tr>
<tr>
<td>E X T X Ss/L X O</td>
<td>802.89</td>
<td>240</td>
<td>3.35</td>
<td></td>
</tr>
</tbody>
</table>

*p < .01
Table 10

Source Table for the 2 X 3 X 5 X 3 X 3 ANOVA on Recognition Rating

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>List (L)</td>
<td>3.2</td>
<td>1</td>
<td>3.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Order (O)</td>
<td>1.6</td>
<td>2</td>
<td>0.8</td>
<td>0.1</td>
</tr>
<tr>
<td>L X O</td>
<td>1.1</td>
<td>2</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>Ss/L X O</td>
<td>331.5</td>
<td>30</td>
<td>11.1</td>
<td></td>
</tr>
<tr>
<td>Experience (E)</td>
<td>747.2</td>
<td>4</td>
<td>186.8</td>
<td>36.2*</td>
</tr>
<tr>
<td>L X E</td>
<td>63.0</td>
<td>4</td>
<td>15.8</td>
<td>3.0</td>
</tr>
<tr>
<td>O X E</td>
<td>39.3</td>
<td>8</td>
<td>4.9</td>
<td>0.9</td>
</tr>
<tr>
<td>L X O X E</td>
<td>56.1</td>
<td>8</td>
<td>7.0</td>
<td>1.4</td>
</tr>
<tr>
<td>E X Ss/L X O</td>
<td>619.0</td>
<td>120</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>Type (T)</td>
<td>12123.1</td>
<td>2</td>
<td>6061.5</td>
<td>763.3*</td>
</tr>
<tr>
<td>L X T</td>
<td>77.8</td>
<td>2</td>
<td>38.9</td>
<td>4.9*</td>
</tr>
<tr>
<td>O X T</td>
<td>30.4</td>
<td>4</td>
<td>7.6</td>
<td>1.0</td>
</tr>
<tr>
<td>L X O X T</td>
<td>12.2</td>
<td>4</td>
<td>3.0</td>
<td>0.4</td>
</tr>
<tr>
<td>T X Ss/L X O</td>
<td>476.5</td>
<td>60</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Length (H)</td>
<td>446.4</td>
<td>2</td>
<td>223.2</td>
<td>55.2*</td>
</tr>
<tr>
<td>L X H</td>
<td>1.5</td>
<td>2</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>O X H</td>
<td>9.4</td>
<td>4</td>
<td>2.3</td>
<td>0.6</td>
</tr>
<tr>
<td>L X O X H</td>
<td>26.4</td>
<td>4</td>
<td>6.6</td>
<td>1.6</td>
</tr>
<tr>
<td>E X Ss/L X O</td>
<td>242.5</td>
<td>60</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>E X T</td>
<td>816.7</td>
<td>8</td>
<td>102.1</td>
<td>22.1*</td>
</tr>
<tr>
<td>L X E X T</td>
<td>40.3</td>
<td>8</td>
<td>5.0</td>
<td>1.1</td>
</tr>
<tr>
<td>O X E X T</td>
<td>30.2</td>
<td>16</td>
<td>1.9</td>
<td>0.4</td>
</tr>
<tr>
<td>L X O X E X T</td>
<td>50.0</td>
<td>16</td>
<td>3.1</td>
<td>0.7</td>
</tr>
<tr>
<td>E X T X Ss/L X O</td>
<td>1106.4</td>
<td>240</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>E X L</td>
<td>370.6</td>
<td>8</td>
<td>46.3</td>
<td>8.5*</td>
</tr>
<tr>
<td>L X E X H</td>
<td>41.1</td>
<td>8</td>
<td>5.1</td>
<td>0.9</td>
</tr>
<tr>
<td>O X E X H</td>
<td>29.3</td>
<td>16</td>
<td>1.9</td>
<td>0.3</td>
</tr>
<tr>
<td>L X O X E X H</td>
<td>45.9</td>
<td>16</td>
<td>2.9</td>
<td>0.5</td>
</tr>
<tr>
<td>E X H X Ss/L X O</td>
<td>1310.6</td>
<td>240</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>T X H</td>
<td>361.8</td>
<td>4</td>
<td>90.4</td>
<td>17.6*</td>
</tr>
<tr>
<td>L X T X H</td>
<td>45.5</td>
<td>4</td>
<td>11.4</td>
<td>2.2</td>
</tr>
<tr>
<td>O X T X H</td>
<td>27.1</td>
<td>8</td>
<td>3.4</td>
<td>0.6</td>
</tr>
<tr>
<td>L X O X T X H</td>
<td>36.5</td>
<td>8</td>
<td>4.6</td>
<td>0.9</td>
</tr>
<tr>
<td>T X H X Ss/L X O</td>
<td>618.3</td>
<td>60</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>E X T X H</td>
<td>1121.6</td>
<td>16</td>
<td>70.1</td>
<td>13.9*</td>
</tr>
<tr>
<td>L X E X T X H</td>
<td>98.0</td>
<td>16</td>
<td>6.1</td>
<td>1.2</td>
</tr>
<tr>
<td>O X E X T X H</td>
<td>130.1</td>
<td>32</td>
<td>4.1</td>
<td>0.8</td>
</tr>
<tr>
<td>L X O X E X T X H</td>
<td>297.1</td>
<td>32</td>
<td>9.3</td>
<td>1.8*</td>
</tr>
<tr>
<td>E X T X H X Ss/L X O</td>
<td>2424.8</td>
<td>480</td>
<td>5.1</td>
<td></td>
</tr>
</tbody>
</table>

*p < .01
### Table 11

Experiment 2: 2 X 3 X 4 ANOVA on Recognition Ratings

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>List (L)</td>
<td>0.1</td>
<td>1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Subjects Within List</td>
<td>146.3</td>
<td>48</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>Type (T)</td>
<td>2584.5</td>
<td>3</td>
<td>861.5</td>
<td>829.0*</td>
</tr>
<tr>
<td>List X Type</td>
<td>3.4</td>
<td>3</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Ss X T Within List</td>
<td>149.6</td>
<td>144</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Length (H)</td>
<td>1.7</td>
<td>2</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>List X Length</td>
<td>0.0</td>
<td>2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Ss X H Within List</td>
<td>104.9</td>
<td>96</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Type X Length</td>
<td>42.9</td>
<td>6</td>
<td>7.1</td>
<td>9.8</td>
</tr>
<tr>
<td>List X Type X Length</td>
<td>1.6</td>
<td>6</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Ss X T X H/L</td>
<td>210.3</td>
<td>288</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>

* p<.01
Table 12

Experiment 2: $2 \times 3 \times 1^*$ ANOVA on Total No. of Correct Recognitions

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>List ($L$)</td>
<td>0.04</td>
<td>1</td>
<td>0.04</td>
<td>2.5</td>
</tr>
<tr>
<td>Subjects within List</td>
<td>0.8</td>
<td>48</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Type ($T$)</td>
<td>39.6</td>
<td>3</td>
<td>13.2</td>
<td>454.4*</td>
</tr>
<tr>
<td>List X Type</td>
<td>0.02</td>
<td>3</td>
<td>0.01</td>
<td>0.3</td>
</tr>
<tr>
<td>Subjects X $T/L$</td>
<td>4.2</td>
<td>144</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Length ($H$)</td>
<td>0.05</td>
<td>2</td>
<td>0.03</td>
<td>11.8*</td>
</tr>
<tr>
<td>List X Length</td>
<td>0.00</td>
<td>2</td>
<td>0.00</td>
<td>0.07</td>
</tr>
<tr>
<td>Subjects X $H/L$</td>
<td>2.05</td>
<td>96</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Type X Length</td>
<td>0.6</td>
<td>6</td>
<td>0.1</td>
<td>5.3*</td>
</tr>
<tr>
<td>$L \times T \times H$</td>
<td>0.01</td>
<td>6</td>
<td>0.00</td>
<td>0.09</td>
</tr>
<tr>
<td>Ss X $T \times H/L$</td>
<td>5.3</td>
<td>288</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

* $p<.01$
Bibliography


Harris, R.J. Memory and comprehension of implications and inferences of complex sentences. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 626-637.


Kintsch, W. The representation of meaning in memory. Hillsdale,


Sachs, J. Memory in reading and listening to discourse. Memory and Cognition, 1974, 2, 95-100.


Simmons, R.C. Semantic networks: Their computation and use for understanding English sentences. In R. Shank and K. Colby (Eds.), Computer models of thought and language. San Francisco:


