INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
The investigation of the semantic component of short-term memory in sentence comprehension

Shernoff, Eve Heather, M.A.
Rice University, 1994
RICE UNIVERSITY

THE INVESTIGATION OF THE SEMANTIC COMPONENT OF SHORT-TERM MEMORY IN SENTENCE COMPREHENSION

by

EVE HEATHER SHERNOFF

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE MASTER OF ARTS

APPROVED, THESIS COMMITTEE

Randi C. Martin, Professor, Chair
Department of Psychology

James R. Pomerantz, Professor
Department of Psychology

David Lane, Associate Professor
Department of Psychology

Houston, Texas

May, 1994
ABSTRACT

The Investigation of the Semantic Component of Short-Term Memory in Sentence Comprehension

by

Eve H. Shernoff

Auditorily presented information is processed on an on-line basis to the extent possible (Marslen-Wilson and Tyler, 1980). Using sentences where immediate, on-line processing was not possible, this study investigated whether such sentences created a greater short-term memory burden for subjects. The types of sentences in which immediate integration was not possible had either a list of adjectives preceding a noun or a list of nouns preceding a verb. These were compared with sentences which had lists of adjectives following the noun or lists of nouns following the verbs, allowing for immediate integration. Two experiments, one using word detection and the other using a sentence anomaly judgment found support for the hypothesis that the inability to immediately integrate word meanings increases memory burden and impairs comprehension. However, the effects of the memory burden did not correlate with standard measures of short-term memory.
Acknowledgements

I would like to express my sincere appreciation to all those who helped me with this thesis. First and foremost, I want to thank Dr. Randi Martin for her constant guidance. She has continually provided me with advice and support throughout this project. I would also like to thank the members of my committee, Dr. James Pomerantz and Dr. David Lane, for their expert advice and assistance.

In addition, I would like to thank Dr. Cristina Romani for her help and guidance as well as for the hard work she put into this project. Finally, I would like to thank Lisa Sanders, Trisha Burton and Kim Doe for helping to collect data and test subjects.
# Table of Contents

1. Components of Sentence Comprehension  
   A. Phonological Component  
   B. Syntactic Component  
   C. Semantic Component  
2. Integration  
3. Chunking studies  
4. The Immediacy of Processing Theory  
5. Experiment 1 - Word Detection  
   A. Method  
      1. Subjects  
      2. Materials and Apparatus  
      3. Procedure  
   B. Results  
   C. Discussion  
6. Experiment 2 - Anomaly Judgment  
   A. Method - Anomaly Judgment  
      1. Materials and Procedure  
   B. Method - Memory Span Tasks  
      1. Materials and Procedure  
   C. Results - Anomaly Judgment  
   D. Results - Memory Span Tasks  
   E. Discussion  
7. General Discussion  
8. References
LIST OF TABLES

Table 1. Mean reaction times by subjects for the interaction of anomalous/sensible, before/after and distance - Exp. 1..............31
Table 2. Error rates for sensible/anomalous, before/after, and distance..........................................................34
Table 3. Mean reaction times by subjects for the interaction of anomalous/sensible, before/after and distance - Exp. 2..............44
Table 4. Mean reaction times by subjects for the interaction of anomalous/sensible, before/after and phrasetype - Exp. 2..............46
Table 5. Error rates for before/after, and distance.................................47
LIST OF FIGURES

Figure 1. Experiment 1 - Overall analysis: Interaction of distance and anomalous/sensible separately for before and after..................32
Figure 2. Experiment 2 - Predicted results..................................................38
Figure 3. Experiment 2 - Predicted and obtained results.......................52
LIST OF APPENDICES

Appendix A. Significant interactions from Experiment 1......................61
Appendix B. Significant interactions from Experiment 2......................62
Appendix C. Correlation table for subjects' performance on
             Experiment 2 with memory span tasks................................63-65
Components of Sentence Comprehension

Overview: Short term memory is assumed to play a crucial role in sentence comprehension. Although there are different components of short term memory this thesis focuses on the role that short-term memory plays in the retention of semantic information. This issue was investigated in two experiments in the context of auditory comprehension using normal subjects. Sentences were created that were similar in overall meaning, but had different syntactic structures. The syntactic structures differed in such a way that in some sentences the meaning of the lexical items could immediately be integrated while for other sentences immediate integration was impossible. The purpose of this syntactic manipulation was to try to determine if the sentences where integration was delayed would create a memory burden that affected comprehension.

The literature review below considers previous studies on the role of short-term memory in sentence processing. Although most such studies have focused on phonological short-term memory, a few have considered the role of syntactic and semantic short-term retention.

Background: Three factors involved in auditory sentence comprehension include the creation of phonological, syntactic and semantic representations of the input. The first stage requires the
transformation of the acoustic information into a phonological form. This in turn enables access to the lexicon. After the lexical entries are located the syntactic and semantic features are available for analysis. The outcome of the parsing and the semantic interpretation of the sentence is the propositional representation. According to Martin (1990), short-term memory deficits, which might occur for phonological, semantic or syntactic information could affect sentence comprehension in different ways. Each factor and its consequences will be discussed separately.

Various studies have looked at the role of phonological short-term memory using a dual-task method in which subjects articulate irrelevant verbal material over and over while carrying out a sentence comprehension task. The rationale behind these paradigms is that the subject will not be able to subvocalize text material while reading if the reader's speech tract is engaged in another task. It is therefore assumed that a phonological code of the sentence will not be created or stored in short-term memory. One possible consequence of this is that preventing phonological coding might impair comprehension by preventing access to word meaning. Another possibility is that prevention of phonological coding will disrupt retention of word order and consequently disrupt syntactic processing. Although numerous studies have been carried out, the results of these studies have been inconclusive. Kleiman (1975) and Levy (1978) concluded from their studies that phonological recoding plays a role in maintaining various parts of a sentence in working
memory although, it is not used for accessing word meaning. Waters, Komoda, and Arbuckle (1985) claim that the impairment found in sentence comprehension in the previous studies using articulatory suppression may not necessarily be due to prevention of phonological coding, but rather due to the drain of general mental resources from dual task methodology. In their experiments they had subjects read and comprehend prose passages while doing a concurrent task. They found that when they adjusted for difficulty of secondary task there was no specific interference that could be attributed to disruption of phonological coding.

A study by Martin, Wogalter and Forlano (1988) also challenges the claim made by Kleiman and Levy. This study sought to determine whether there would be a detrimental effect on reading comprehension from unattended speech. Unattended speech has been shown to interfere with short-term memory tasks, regardless of whether the speech is meaningful to the listener (Colle & Welsh, 1976, Salame & Baddeley, 1982). Unlike the short-term memory tasks, Martin et al. failed to find an interference effect on reading comprehension for meaningless speech background. This finding argues against a role for phonological short-term memory in reading comprehension.

Another interesting way to study the role of phonological short term memory is with stroke patients who have restricted short term memory due to brain damage. Because these patients already have a phonological short-term memory deficit, there is no need to use the
dual task method. Martin (1990) provides a review of the effects of a phonological deficit on sentence comprehension. Three possible roles for phonological memory in sentence processing are assessed through examination of a brain damaged patient, EA. She demonstrates a reduced memory span that which can be attributed to a phonological retention deficit.

The first role of phonological memory is that it may be crucial in computing a syntactic analysis of the sentence. Incoming information may be held in a phonological form while syntactic and semantic analyses are carried out on the earlier portion of the sentence (Martin, 1987). Evidence against this possibility comes from EA and other patients with phonological deficits such as HB (Howard, 1989) and BO (Waters, Caplan, & Hildebrandt, 1991). HB and BO could perform within the normal range on sentence-picture matching tasks, as well as other comprehension tasks, that were dependent upon syntactic analyses (Martin, 1993).

The second possible role of phonological memory is that it can serve as a backup representation in such cases where an ambiguity in a sentence has been interpreted incorrectly. Although this makes sense intuitively, EA's performance on comprehension of auditorily presented garden-path sentences does not support this hypothesis. A garden-path sentence is defined as a sentence in which the reader misinterprets an ambiguity in the sentence and must reanalyze it in order to understand its correct meaning. An example of a garden-path sentence would be, "Can the laughter, it wasn't funny." For a
garden-path sentence, it would seem that you would have to hold onto some verbatim form of the sentence, since you often have to back-track and reanalyze the sentence. EA scored within the normal range on this test, which she would not have done if she needed to retain a phonological analysis of the sentence. These results do not support the contention that phonological memory serves as a backup devise in aiding the subject in a reanalysis of a sentence in which an ambiguity is incorrectly interpreted.

A third role for phonological memory is that it can serve as a supplemental resource to syntactic and semantic memory records that can assist in verbatim repetition of sentences (Martin, 1990). Repetition tests done with EA show that she has trouble repeating many sentences, although there were some she repeated quite accurately. Her performance generally supports the contention that phonological short-term memory is need for repetition. It should be pointed out that although EA's repetition is poor her comprehension is good. She can retain the gist of a sentence even though she cannot repeat it back verbatim.

The second stage in which capacity limitations might cause a breakdown in sentence comprehension is in the retention of a syntactic representation. Although phonological retention does not appear to be important for syntactic processing, there may be short-term memory capacities specific to retaining syntactic structures (Kolk and Van Grunsven, 1985, Frazier and Fodor, 1978). Some work by Frazier (1985) supports the notion that the amount of syntactic
structure that must be created and maintained as each word is processed affects comprehension. She summarizes the evidence suggesting that the greater the “local nonterminal” count, the more difficult comprehension. The local nonterminal count refers to the number of syntactic nodes (i.e., noun phrase, verb phrase, prepositional phrase) that must be created within a short span of words. The idea that aphasic patients might have restricted capacity for retaining syntactic structures is supported by Ostrin and Schwartz (1986). They demonstrated that for some agrammatic patients, loss of syntactic structure occurs independently of semantic influences. In other words, they found that the memory trace for syntactic structure of a sentence is almost completely degraded, leaving the patients with a "skeletal frame" (pg. 342). However, the degree of degradation is not influenced by semantic factors such as plausibility and reversibility. The subject is left with a minimal amount of syntactic information and from this information the sentence must be reconstructed. Plausibility can guide this reanalysis, but it is not the key factor.

The third factor important in sentence comprehension is the ability to retain semantic information. Support for the idea that sentence comprehension is also dependent upon the contributions of a semantic component of short term memory is found in Martin, Shelton, and Yaffee (1994). They present data suggesting that the reduced memory span of some aphasic patients can be partially due to a reduction in their ability to retain semantic information.
Shelton, Martin and Yaffee (1992) present a patient, AB, who appears to have a semantic short-term memory deficit. The claim that AB's short-term memory deficit is semantic rather than phonological in nature, is based on his performance on short term memory tests as well as on sentence comprehension tests. Although AB has exhibited a reduced memory span similar to EA's (i.e. 2-3 items), he has shown normal effects for phonologically related variables, including a phonological similarity effect, word length effect and a recency effect (Martin, 1987). EA's performance was better with the visually presented words than with auditorily presented items, as is characteristic of a person with a phonological deficit. AB, however, showed the normal pattern, performing better on the auditorily presented words. Although AB has a greater phonological memory capacity, he shows worse performance than EA on retention of concrete lists (Shelton, Martin, and Yaffee, 1992). Other evidence for AB's semantic deficit comes from his performance on nonword list retention. He recalled nonword lists (2-3 items) at the same level as word lists. Unlike EA and normal subjects, AB shows little advantage for the retention of words over nonwords. AB's performance on a category probe task also illustrates his deficit in the retention of semantic information. In this task a list of words is first presented followed by a single probe word. The subject's task is to say whether the probe word appeared in the original list. AB's performance was significantly worse than EA's, despite his superior phonological memory. In fact, AB's performance was so
poor, he only scored 80% on one item lists (Martin, Shelton, & Yaffee, 1994). AB's performance on sentence comprehension and repetition tasks also provides evidence for his semantic short-term memory deficit. Although AB performs better than EA on sentence repetition, on a comprehension task in which a simple attribute question was asked such as, "Which is quiet a concert or a library?" AB performance was below chance. EA, however, scored 100% on this task. In order to ensure that AB understood the sentences, he was asked "Is a library quiet?" or "Is a concert quiet?" He was able to answer these sentences correctly. The results of this task and other tasks, such as the token task (De Renzi and Vignolo, 1962) "suggest that his [AB's] difficulty in retaining semantic information applies not only to the retention of random word lists but also to the retention of more meaningful semantic information" (Martin, Shelton, & Yaffee, 1994).

Based on the different profiles of these two patients, Martin et al. argues that there are two separable components to short term memory. The first component is involved in the retention of phonological information, which is disrupted in EA. The other component is the retention of semantic information, which is disrupted in AB.

There is little data from normal subjects concerning the role of and the capacity for semantic retention in sentence comprehension. The question addressed in this study relates directly to the retention of the semantic interpretation of words in
a sentence and how sentence processing is affected by differential memory loads. The present studies investigated the effects on sentence processing when subjects are required to retain semantic information before being able to fully integrate it with the other information in the sentences. More specifically, we examined comprehension of sentences in which there are words whose meanings cannot be immediately elaborated or fully understood until information later in the sentence is encountered. The question we addressed is whether unintegrated semantic information would place a greater burden on short-term memory than does integrated semantic information. The next section provides a discussion of the concept of integration.

Integration

The term integration has often been used in language processing research, though a definition is not usually provided. A few examples will give a general sense of how various authors have used the term. Murphy (1990) uses the term integrate in the context of the schema model saying, "The noun modifier does not automatically pick out one slot of the head noun's schema...In the example, apple pie, when apple appears as a modifier, listeners must try to integrate it with the noun in the best way possible for that pair of words." (pg.263) Baddeley and Wilson (1988) also use the term regarding a situation in which they claim a phonological code
would be useful for sentence processing. They state, "temporary storage is necessary when an early word arrives and cannot be disambiguously assigned within the mental model until later information has been provided....It is necessary [at times] to hold and integrate information across several words." (pg.495) An example illustrating this is the self-embedded sentence, "The book the pencil is on is red." In this sentence it is necessary to hold onto the word 'book' until reaching the word 'red' before they can be integrated. "It is not sufficient to remember that a book and a pencil are involved, one must also carry the intervening verbal information if one is to map the adjective onto the appropriate noun." (pg.495)

For the purposes of this research, the term integration will be defined as the combination of the meaning of terms. This combination will replace the individual meanings with reference to an object or action that contains more information than was available from the individual terms. For example, "green grass" can be integrated into a concept of an object which is grass of the color green.

The question raised in this study is whether integrated versus unintegrated information places different demands on short-term memory. This question relates, at least marginally, to early studies on chunking and short-term memory for verbal material. The term 'chunking' is closely related to the concept of integration. However, chunking is used here to describe a process that aids in recall of information, while integration is a term used to describe a process
that facilitates comprehension. It is possible to chunk information without integrating it, as in the case of remembering a telephone number, however, it is not possible to integrate information without chunking it at the same time. Previous studies have shown that when chunking is possible, memory for sentences is improved. In one study, Miller and Selfridge (1950) created a body of word lists which ranged from random-word order lists to sentences that were both syntactically and semantically correct. As the word lists became more semantically and syntactically correct, retention improved. This finding was explained by the use of chunking. As the word strings increased in their approximation to an English sentence, subjects used their knowledge of English to facilitate their immediate memory by combining or chunking the individual lexical items into larger, more meaningful units. By combining the information into fewer units, the memory load is decreased, thus facilitating retention of the verbal material. In a similar study by Tulving and Patkau (1962), results showed that improvement of performance in recall of word lists was not due to the subjects' recalling more chunks. Instead, improved performance was due to the fact that as the list grew in its likeness to an English sentence, the subjects formed larger chunks of information, as opposed to just increasing the number of chunks they formed. Tejirian (1968) addressed the role semantic and syntactic information play in the retention of verbal information. He found that the semantic information in word lists did not play a role in retention until the
word lists reached a certain level of approximation to English. Syntactic information, however, aided retention even for random word lists. This suggests that syntactic structure is needed to know how words should be utilized.

These early studies suggest that the possibility of integration does decrease the memory load. This does not mean the memory load is reduced to zero. However, these studies did not address the word-by-word processing of a sentence to determine the point at which integration occurs. Some later research by Marslen-Wilson and Tyler (1980) indicated that semantic and syntactic analyses are computed, to the extent possible, from the very first word of the utterance. Based on this assumption, a second issue, investigated in Experiment 1 and 2 of this master's thesis, is raised. These experiments address the issue of how sentence processing is affected when the information in the sentence does not immediately lend itself to chunking and integration with the other information in the sentence. Our studies contrasted sentence constructions where complete semantic interpretations cannot be made as each word is encountered with sentences where more immediate, on-line interpretations can be made. However, before considering situations in which interpretations must be delayed, it is important to review evidence that illustrates immediate processing in sentence comprehension when interpretations do not have to be delayed.
The Immediacy of Processing Theory

Marslen-Wilson and Tyler's (1980) immediacy of processing theory claims that from the very first word of an utterance, syntactic and semantic analyses are applied to the input and interpreted to the extent possible in the context of the current discourse. Their view contrasts with a "syntax-first" approach which assumes that syntactic structure is assigned first to a phrase followed by a semantic and pragmatic interpretation (Frazier, 1978; Frazier and Fodor, 1978). This theory proposes that subjects compute an analysis of a sentence based primarily on the syntactic information that is available. This syntactic analysis is computed in the simplest way possible such that the number of structural components into which the sentence is broken is kept at a minimum. This is termed minimal attachment. A late closure strategy is used to make an attachment judgment in a case when there are two possibilities for minimal attachment. What this means is that attachments to preceding items will be more likely than attachments to subsequent items. Both minimal attachment and late closure are a result of adopting the first analysis available. This model does not exclude the role of content. It holds that syntax guidance alone is not sufficient to account for all attachment decisions. Content is viewed as not having the ability to guide initial attachment, but it can serve to reject or accept an initial attachment (Rayner, et al. 1983).
The immediacy of processing theory proposed by Marslen-Wilson and Tyler (1980) contrasts the syntax-first model by claiming that the listener tries to fully interpret the input as he hears it and that the recognition of each word, from the beginning of an utterance, is directly influenced by the contextual environment in which that word is occurring. In other words, syntactic and semantic analyses of the sentence are carried out on a word-by-word, on-line basis, to the extent that it is possible. In order to support this hypothesis, their first experiment investigated how different knowledge sources affect the time to recognize individual words. The second question they addressed was whether or not there is a time ordering in the system that controls the availability of semantic and syntactic information across an utterance and, if so, at what stage in the analysis of the utterance do the different types of processing information become available.

A word monitoring task was used in which subjects pressed a button when the target word occurred in the sentence. There were three kinds of sentences; normal prose, syntactic prose, which had an acceptable syntactic structure but no coherent meaning, and random word order. There were also three different types of monitoring tasks. The first was identity monitoring, in which subjects monitored for the identical word that was presented. Second, there was rhyme monitoring, in which subjects monitored for a word that rhymed with the probe word, and third was category
monitoring, in which the subject monitored for a word that was a member of a specific category.

The results of the first experiment addressed the first question regarding how different knowledge sources affect the time to recognize individual words. There was a main effect on word recognition for prose condition, showing that reaction times were fastest for the normal prose condition, where both syntactic and semantic information are present. Reaction times were also facilitated in the syntactic sentences relative to the random word condition. There was also a main effect of the monitoring task where reaction time increased from identical to rhyme to category monitoring. As predicted, in the normal prose condition there was no difference between the Rhyme and Category monitoring conditions. According to the on-line interactive approach, word-identification involves interaction between the acoustic-phonetic input and the syntactic and semantic context in which the word is occurring. Therefore, by the time a word has been identified there should be no bias in the availability of phonological information over semantic information about that word. Thus, as demonstrated by the results, semantic attribute-matching decisions in the category monitoring task was not delayed relative to phonological attribute matching in the rhyme monitoring task.

Analysis of word position is important with regard to the second question of whether or not there is a time ordering in the system. The results indicate that in the normal prose condition
reaction times decreased as the target word occurred later in the sentence for all three types of monitoring tasks. In the syntactic prose condition reaction times decreased in identical and rhyme conditions, but not in the category monitoring task. In the random word order condition there was no effect of word position. The results showing that category monitoring was faster in the normal prose than syntactic prose, in early list positions, indicates that semantic information was available immediately and seems to have an equally strong influence throughout the sentence. Hence, their results argue against the notion that semantic interpretation is delayed relative to syntactic analysis.

Other evidence supporting the immediacy of processing theory is derived from a study by Marslen-Wilson, Brown, and Tyler (1988). This study will be described in some detail as the methodology used in this thesis is close to that used by Marslen-Wilson et al. The main issue raised in this study regards the information that is available to the listener when a specific lexical item is accessed. In this experiment, both anomalous and sensible sentences were used. The anomalous sentences were constructed from sensible sentences by changing a single word. These anomalous sentences contained either a pragmatic, syntactic or semantic violation, with respect to the verb. These different types of violations should only affect word monitoring if the relevant type of information indicating the anomaly was computed at the point that the word was perceived. In addition to the three types of violations, a lead-in
sentence was added which gave either a natural or unnatural linkage to the continuation sentence. The lead-in sentence was included to determine if its use would enhance the effects of pragmatic plausibility. An example of a complete set begins with the two possibilities for lead in sentences, which are connected to either a sentence in the normal, pragmatic, semantic, or syntactic condition.

**Lead in sentences:**

1. The nurses walk to their work each morning.
2. Christmas falls on a Friday this year.

**Conditions:**

**Normal:**
1. They pass the beach on their way to the hospital.

**Pragmatic:**
2. They measure the beach on their way to the hospital.

**Semantic:**
3. They chew the beach on their way to the hospital.

**Syntactic:**
4. They yawn the beach on their way to the hospital.

Reaction time to noun targets following the verb was used as the dependent measure. Time to detect the target in the anomalous sentences was compared to the time to detect the target in the acceptable, normal sentences. Longer monitoring times for the pragmatic condition would be expected if subjects were computing a pragmatic, non-lexical interpretation at this point in the sentence. This would be the case only if lexical and non-lexical representations associated with the verb and pragmatic plausibility have already started to affect an interpretation of the sentence.
Longer monitoring times in the semantic condition compared to normal sentences would be expected if the semantic restrictions associated with the argument frames of the verb have become available. Given that the semantic violation sentences are also pragmatically plausible, a greater disruption in monitoring might be expected for the semantic anomalies. The violations in the syntactic condition should affect sentence processing the most since they violate syntactic rules as well as semantic and pragmatic rules. The results of this experiment showed that pragmatic, semantic and syntactic latencies were all significantly slower than normal sentences. As hypothesized, the reaction times for the sentences with syntactic violations were the longest. These RT's were significantly slower than the RT's for the sentences with semantic and pragmatic violations. Although the RT's for the pragmatic and semantic violations were not significantly different using a min F', the difference was significant for both subjects and items on the Neuman-Keuls test. Therefore, the results for the pragmatic and semantic violations were interpreted according to the latter analysis. These results are consistent with the hypothesis that the sentences with the semantic violation will be more disruptive than the pragmatic violation since the semantic anomaly results in both lexical and non-lexical disruptions. "These results are consistent with approaches to language comprehension which stress its continuous and incremental nature and which emphasize
the role of lexically derived thematic roles in constructing these higher level representations." (pg.15)

Other studies supporting the immediacy of processing theory can be found in the 'ambiguity' literature. Foss (1970) and Foss and Jenkins (1973) used a phoneme detection task to investigate the issue of whether or not, in normal sentence processing, more than one grammatical interpretation of an ambiguous input sentence is recovered. As with the word detection paradigm, reaction time was used as an index of sentence processing. The results showed that reaction times were faster for unambiguous sentences than for ambiguous sentences, indicating that processing of the sentence or integration of semantic and syntactic information is momentarily slowed down by the presence of the ambiguity. This lends support to the immediacy of processing theory, because it implies subjects are attempting to exact an interpretation of each word as it is encountered. This on-line analysis, which is impeded in the case of an ambiguous sentence, is therefore slowed down, as is reflected in the reaction times of the subjects.

Carpenter and Daneman (1981) also present a reading study that supports the hypothesis that interpretations of the input are not deferred until some later stage, regardless of the level of processing. In order to test this they used eye-monitoring data. The assumption behind this process is called the eye-mind assumption which states that the eye remains fixated on a word as long as the word is being processed (Just and Carpenter, 1980). In the
Carpenter and Daneman (1981) study, they presented subjects with a two sentence passage. The first sentence biased the reader to incorrectly interpret an ambiguous word in the second sentence. An example of this is, "Tomorrow was the annual, one-day fishing contest and fisherman would invade the place. Some of the best bass guitarists in the country would come to this spot." The results showed that subjects spent an abnormally long time on the word 'guitarists', which is the word disambiguating the word 'bass.' These results indicate that subjects are interpreting the text as they read it and not waiting until a point later on in the sentence.

One key issue, not addressed by Marslen-Wilson and Tyler, is the situation previously mentioned, in which the listener is unable to immediately integrate the incoming words of a sentence. The initial, incomplete semantic representation of each word must be held in short-term memory until later information can aid in forming a single unit. For example, in the sentence, "The dry, red, chilled wine made the dinner a success," the initial semantic interpretation of the three adjectives must be held in short term memory until the word 'wine' is reached, at which time the unit as a whole (the three adjectives and the noun) can be integrated and understood. In other words, it is not until the noun is reached that the complete correct interpretation of the adjectives can be realized, at which point they are integrated with the noun to form a single representation. Integration of the lexical items would not only be expected to decrease the memory load based on chunking
studies, but it is also prerequisite for forming a correct semantic representation of the words in the sentences.

**EXPERIMENT 1 - WORD DETECTION**

The question raised in this study addresses the role that the semantic component of short term memory plays in sentence comprehension. This issue was addressed using an on-line probe task, called word detection, which was modeled after that used by Marslen-Wilson and Tyler (1980) and Tyler (1988). This task was designed to measure sentence processing by having subjects monitor for a specific word in a sentence. In this task, the subject first hears a single probe word which is then followed by a sentence. The subjects' task is to press a key on the computer keyboard as soon as the probe word is heard in the sentence.

In this first experiment, two types of sentences were used, noun phrase sentences and adjective phrase sentences. These sentences were not expected to interact differently with the other variables. The noun phrase sentences contained a verb and either 1, 2, or 3 nouns. The sentences were designed in pairs such that for every sensible sentence there was also an anomalous sentence. For example, if the sensible sentence was, "Tokens, coins and keys jingled *quietly* in her pocket," then the anomalous counterpart was "Gum, coins and keys jingled *quietly* in her pocket." The anomaly always occurred between the verb and the noun farthest from that
The target word, which in this example is the word 'quietly', always appeared after the verb. For the anomalous sentences, this meant the target word appeared after the word which completed the anomaly. The other type of sentences that were constructed were termed adjective phrase sentences. These sentences contained a noun which was modified by either 1, 2, or 3 adjectives. These sentences were also designed in anomalous/sensible pairs. For example, if the sensible sentence was "She brought the rusty, old, red pail to the beach," the anomalous sentence would be, "She brought the rusty, old, red swimsuit to the beach." The anomaly in these sentences occurred between the noun and the adjective farthest away from it. In other words, the anomaly exists between the words 'rusty' and 'swimsuit'. The target word, which in this example is the word 'to,' appears right after the noun. In the case of the anomaly sentences, it is the word that follows the completion of the anomaly. The rationale behind this manipulation is that the presence of the anomaly in the sentence will slow down the subjects' comprehension process. Therefore, the subjects' response time to the probe word for the anomalous sentences will be slower in comparison to their response time to the probe word in the matched sensible sentence.

Two other variables were included. The first variable involves the ordering of the words in the sentence. It is called the before/after manipulation. For adjective sentences in the before condition, the adjectives precede the noun (i.e., "She brought the
risky, old, red pail to the beach.") In the before condition for the noun phrase sentences, the nouns precede the verb (i.e., "Tokens coins and keys jingled quietly in her pocket.") In the after condition, the adjectives follow the noun in the adjective phrase sentences (i.e., "The pail was old, red, and rusty but she brought it to the beach anyway.") and the nouns follow the verb in the noun phrase sentences (i.e., "She jingled tokens coins and keys quietly as she walked.") The idea behind this manipulation is that if words do need to be held in semantic short-term memory until they can be integrated, then the subjects should have more difficulty detecting the anomaly in the before condition than in the after condition. This is based on the idea that for the adjective phrase sentences in the before condition, the adjectives cannot be integrated or attached to the noun that they are modifying until it is processed. In the before condition for the noun phrase sentences the same rationale is applied. The meaning of the nouns preceding the verb cannot be fully elaborated and integrated since the subject does not know the role the nouns play (i.e. agent, object) until the verb is encountered. This inability to fully integrate and elaborate the meaning of these words immediately places a greater memory load on the subject. It is hypothesized that there should be an interaction between the sensible/anomalous and before/after variables. In the before condition, where there is more of a memory burden placed on the subject, the anomaly will be less noticeable and therefore have less of an inhibitory affect on sentence processing. However, in the after
condition, the anomaly should slow down the subjects speed of processing since the subject will not have as great a memory load due to the opportunity to integrate all the information as it is encountered.

In addition to these variables, a third variable was manipulated. This variable, termed distance, refers to the number of nouns in the noun phrase sentences and the number of adjectives in the adjective phrase sentences. There are three levels of distance. An example of distance 1 in the before, adjective phrase condition would be, “She bought the soft pillow at the department store,” distance 2 would be “She bought the soft, new pillow at the department store” and distance 3, “She bought the soft, new, fluffy pillow at the department store.” These three levels of distance also apply in the after condition. This variable was used to exaggerate the differences of the before and after manipulation. As stated previously, it is expected that the 'before' condition will place a greater memory load on the subject than the 'after' condition does. It is expected that this before/after manipulation will interact with the three levels of distance. In other words, as distance increases in the before condition, the memory burden will increase. However, in the after condition, where the subject can integrate each word as it is heard, there should be little, if any effect of distance. These two variables should also interact with the sensible/anomalous manipulation resulting in a three-way interaction. In other words, for the before condition, the presence
of the anomaly should slow down processing time relative to the sensible sentences the most at distance 1, less at distance 2, and little, if any, at distance 3. However, in the after condition, the presence of the anomaly should slow down sentence processing equally at all three levels of distance.

Method

Subjects: Seventy-two Rice University undergraduate students participated as subjects in this experiment. All subjects were required to be native speakers of English. Each subject participated in a one-hour sessions.

Materials and Apparatus: The experimental stimuli consisted of a total of 576 sentences. These sentences were divided into adjective phrase and noun phrase sentences. In the adjective phrase condition there was an equal number of sentences which had a string of 1, 2, or 3 adjectives that either preceded or followed the noun in the sentence. In the noun phrase condition, there was an equal number of sentences that had a string of 1, 2, or 3 nouns preceding or following the verb in the sentence. When the string of adjectives or nouns is preceding the noun or verb respectively, it is called the 'before' condition, whereas when they are following the noun or verb, it is called the 'after' condition. The variable 'distance' refers to the number of adjectives preceding or following the noun in the
adjective phrase sentences and the number of nouns preceding or following the verb in the noun phrase sentences. There is a sensible and anomalous version of each experimental sentence. These pairs are matched on distance, before/after, and phrase type. The design of these sentences can be more clearly seen through an example.

The sentences are organized into blocks of 6 sentences. For every block of 6 sentences in the "after" condition a matched block of "before" sentences was created. The meaning of the sentences in the "before" and "after" conditions are the same, but the syntax of the sentences is different. Within each block there are six versions of the same sentence. For example a block of 6 sentences with the adjectives preceding the noun looks like this:

**Adjectives Before noun**

1. She brought her rusty pail to the beach.
   She brought her rusty swimsuit to the beach.

2. She brought her rusty, old pail to the beach.
   She brought her rusty, old swimsuit to the beach.

3. She brought her rusty, old, red pail to the beach.
   She brought her rusty, old, red swimsuit to the beach.

These same 6 sentences would then be made into the corresponding block of "after" sentences by taking the adjectives and placing them after the noun in the following format;
Adjectives After Noun

1. The pail was rusty, **but** she brought it to the beach anyway.
The swimsuit was rusty, **but** she brought it to the beach anyway.

2. The pail was old and rusty, **but** she brought it to the beach anyway.
The swimsuit was old and rusty, **but** she brought it to the beach anyway.

3. The pail was old, red, and rusty, **but** she brought it to the beach anyway.
The swimsuit was old, red, and rusty, **but** she brought it to the beach anyway.

The target word, **to**, directly follows the noun being modified in each of the 6 sentences in the before condition. In the after condition, the target word, **but**, directly follows the last adjective that is modifying the noun. It is important to note that the target word is always the same within a block of 6 sentences, but its corresponding "after" block might have a different target word. Each subject heard only one of the 12 sentences that were made up from the matched blocks of "before" and "after" sentences. Since there were 48 groups of 12 sentences, each subject heard 48 experimental sentences, which were divided into 24 adjective phrase sentences and 24 noun phrase sentences.

In addition to the 576 experimental sentences there were 144 filler sentences. Of these filler sentences 48 of them were modeled
after the noun phrase sentences and 48 were modeled after the
adjective phrase sentences. The remaining 48 were relative clause
sentences. The filler sentences were created so that they varied in
distance as well as in the before/after manipulation. The difference
between the filler sentences and the experimental sentences was
that for the filler sentences, the target word appeared in different
places in the sentence. These sentences were designed this way so
that the position of the target word would not be predictable. For
24 of the 144 filler sentences, the probe word did not appear in the
sentence at all. In other words, these sentences did not have a
target word that matched the probe word. Fifteen questions were
also constructed that correspond to specific filler sentences. These
questions were asked immediately after the corresponding sentence
was heard. The purpose of these questions was to make sure the
subjects were comprehending the sentences as they went along and
not simply monitoring for the probe word.

The materials were designed to try and eliminate some
possible methodological problems that were not necessarily
accounted for in Marslen-Wilson and Tyler (1980). First of all, the
target word followed the word that completed the anomaly instead
of being the word that completed the anomaly. This was done so
that the subjects could not anticipate where the target would be.
This also gave subjects a little more time to process the anomaly
thus making the effect of the anomaly on the reaction time to the
target word more salient. Another manipulation designed to reduce
predictability of the target was the inclusion of filler sentences in which the target appeared in different places. In addition, out of the twelve variations of each sentence, subjects only heard one of those twelve. This was done so that the subjects did not hear more than one variation of an experimental sentence. There were 24 blocks of adjective experimental sentences and 24 blocks of noun experimental sentences. As stated before, in each block there were 12 variations of the experimental sentence. Each subject heard one sentence from each block totalling 48 experimental sentences. Combined with the 144 filler sentences, each subject heard 192 sentences.

Ten sentences were included to serve as a practice session. In the practice session there were questions relating to 3 of the 10 sentences. Prior to the practice session a short set of instructions was read aloud to the subject by the experimenter. The sentences were recorded by a female voice at 11hz. The target words and the sentences were played through a Macintosh IIcx. The experiment was presented through a PsychLab experiment generation program (Gum and Bub, 1988) created for the Macintosh.

Procedure: The subjects were seated in front of the Macintosh IIcx computer and given instructions explaining that they would see a black dot on the screen indicating the beginning of a trial. After the black dot disappeared they heard a word followed by a sentence. The subject was then told there would be two tasks. First, the subject
was instructed to listen for the word in the sentence and to press the 'M' key as soon as the target word was heard. If the target word was not heard in the sentence, the subject was instructed to press the 'C' key at the end of the sentence. The keys were reversed for left handed subjects. After the sentence was completed a dash appeared on the screen indicating the end of the trial. The return button was pressed in order to proceed to the next trial. Secondly, the subjects were instructed that they would periodically be asked a question about a sentence they had just heard. They were told this was to make sure they were comprehending the sentences as they went along. Subjects were also instructed to find a comfortable position for their hands during the practice trial so that they would not move them around during the experimental session. They were instructed that they must hold one index finger over each response key, M or C.

Each subject participated in a single one-hour session. Each subject was tested individually with the experimenter present in the room.

Results

Subjects responded accurately to the questions that they were asked regarding specific sentences.

An ANOVA (phrasetype x anomalous/sensible x before/after x distance) using reaction time as the dependent measure was
performed. This analysis was done once using subjects as the random effect and once using sentences (items) as the random effect. Reaction times above 1200 msec and below 100 msec were thrown out. Based on this only 1% of the data was discarded.

Of major interest is the three-way interaction between anomalous/sensible, before/after, and distance, which was marginally significant for subjects, $F(2, 142) = 2.85, p < .06$, MSe = 28823 as well as for items, $F(2,91) = 2.48, p < .08$, MSe = 13889. The results for subjects, although only marginally significant, were further analyzed. These results are presented in Table 1 and in Figure 1.

Table 1

Mean reaction times by subjects for the interaction of anomalous/sensible, before/after, and distance - Experiment 1

<table>
<thead>
<tr>
<th>Before Condition</th>
<th>Distance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Anomalous</td>
<td>450</td>
<td>402</td>
</tr>
<tr>
<td>Sensible</td>
<td>420</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>After Condition</th>
<th>Distance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Anomalous</td>
<td>437</td>
<td>418</td>
</tr>
<tr>
<td>Sensible</td>
<td>428</td>
<td>402</td>
</tr>
</tbody>
</table>
Figure 1: Experiment 1 - Overall Analysis: Interaction between distance and anomalous/sensible variables separately for the before and after conditions
The two-way interactions between sensible/anomalous and distance were examined separately for the before and after conditions. In the before condition this interaction was significant, $F(2,142) = 3.81$, $p < .02$, MSe = 22,974. There was a significant main effect of distance, $F(1, 71) = .001$, MSe = 73,146 but not of the sensible/anomalous variable. In order to carry out specific contrasts in the before condition, a difference score was calculated by subtracting the sensible score from the anomalous score for each level of distance. The first contrast compared these difference scores between distance 1 and the average of distance 2 and 3. A significant difference was found indicating that sentence processing was affected more by the presence of the anomaly in the sentence at distance 1 than at the averages of distance 2 and 3, $t = 2.48$, $p < .05$. A second comparison was done looking at the difference between distance 2 and distance 3. No significant difference was found between these two levels of distance in the before condition, $t = 1.15$, $p > .10$.

For the after condition, an ANOVA was performed to determine if there was a differential effect on anomaly across the different levels of distance. Based on the hypothesis, it was expected that the presence of the anomaly should slow down sentence processing at all three levels of distance due to immediate integration of lexical information. This hypothesis was supported by the significant main effect of anomaly $F(1,71) = 6.08$, $p < .02$, 
MSe = 25,953, as well as by the non-significant interaction between the anomalous/sensible x distance variables, $F (2,142) = 0.01$, p > .10, MSe = 231. The mean of the anomalous sentences was 429 msec, while the mean for the sensible sentences equalled 416 msec.

Appendix A shows all significant interactions, none of which were particularly interesting. They all included the variable 'phrasetype' which was not of interest since the overall four-way interaction between phrasetype x anomalous/sensible x before/after x distance was not significant. The error analysis, as shown in Table 2, also did not reveal anything of particular significance.

Table 2

Error rates for sensible/anomalous by before/after by distance

<table>
<thead>
<tr>
<th></th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Before Condition:</td>
<td></td>
</tr>
<tr>
<td>Sensible</td>
<td>.00</td>
</tr>
<tr>
<td>Anomalous</td>
<td>.004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>After Condition:</td>
<td></td>
</tr>
<tr>
<td>Sensible</td>
<td>.00</td>
</tr>
<tr>
<td>Anomalous</td>
<td>.00</td>
</tr>
</tbody>
</table>
Discussion

The goal of this experiment was to examine the influence of memory load on sentence processing. The hypothesis, as stated before, predicts that when information is not immediately integratable, the subjects' speed or ease of processing will be affected by the amount of information that must be retained until integration of that information is possible. However, when information can immediately be integrated into the sentence, sentence processing should not be affected by the amount of information in the sentence. These hypotheses were supported to some extent by the data. Although the results were only marginally significant, the trends of the data were in the expected directions. One possible reason these results failed to reach significance could be due to the task. In the word detection task, it is not necessary that the subject process the sentence for meaning. The task can be viewed almost as a dual task in which the subject must monitor for a specific word as well as comprehend the sentence. Perhaps the second task of comprehending the sentence is not receiving the attention it needs. As mentioned earlier, subjects were given questions with the hopes of ensuring that they would be forced to comprehend the sentences they heard and not simply monitor for the target word. However, it is not clear if that was the case. For the most part, the questions were answered merely by repeating the sentence or part of it. It is possible that subjects processed the
sentence phonologically when doing the word detection task and only computed a semantic representation when a question was asked. If the sentence is not comprehended on a semantic level during word detection, this could result in a decrease in the size of the effects. It is possible that this lack of semantic processing is responsible for the marginally significant results discussed below.

The second experiment investigates the same issue of sentence comprehension using an anomaly judgment paradigm. Subjects listened to an auditorily presented sentence and made a decision as to whether or not it was anomalous. In this paradigm comprehension of the sentence is required in order to perform the task.

EXPERIMENT 2 - SENTENCE ANOMALY JUDGMENT

This experiment was designed to investigate the role of short-term memory in sentence comprehension using a slightly different paradigm. The subjects were required to determine if an auditorily presented sentence was anomalous or sensible. Subjects were instructed to respond by hitting a designated key on the keyboard as soon as they detected the sentence was anomalous. If the sentence was sensible, they made their response at the end of the sentence. In this experiment, the same variables (i.e. noun/adjective phrase, anomalous/sensible, before/after, and distance) were manipulated. The rationale behind this experiment is the same as in Experiment 1.
Subjects should be unable to immediately integrate the string of adjectives when they precede the noun, as in the case of the adjective phrase sentences in the 'before' condition (e.g. "She brought her rusty, old, red pail to the beach"). Ease of integration in the 'before' condition will further be affected by the number of adjectives that precede the noun being modified. However, for the sentences in the 'after' condition, e.g. "The pail was old, red, and rusty but she brought it to the beach anyway," it is assumed that the meanings of the adjectives that follow the noun can immediately be integrated into the already existing semantic structure, and therefore do not place a memory burden on the subject. Since there is no memory burden placed on the subject, the manipulation of 'distance', the number of adjectives in the sentence, should not affect speed of processing. The same rationale applies to the noun phrase sentences. Based on this rationale, an interaction between the anomalous/sensible, before/after, and distance variables is expected.

Although the two experiments are investigating the same issue, the data from this experiment will look different from the data in Experiment 1. First of all, because subjects had to wait until the end of the sentence to make their response if the sentence was sensible, the reaction times for the sensible sentences will be much longer than for the anomalous sentences. Second, difference scores are not appropriate to use since the point of the decision is not the same in the anomalous and sensible sentences. Third, in the
before condition, both reaction times and error rates for the anomalous sentences should increase as distance increases. This is due to the increase in the memory load as the amount of unintegrated information in the sentence increases. The increase in the memory load causes the anomalies in the sentence to be less salient, thus resulting in more errors. For the anomalous sentences in the after condition, reaction times should either stay the same or perhaps decrease depending on whether there is a speed accuracy trade-off with the error rate. The predictions for the anomalous sentences are shown in Figure 2. Reaction times to the sensible sentences should remain stable across the levels of distance for both the before and after conditions.

![Graph showing reaction time vs. distance](image)

**Figure 2.** Experiment 2 - Predicted results
Just and Carpenter (1992) address the issue of a reduced short-term memory capacity from a slightly different angle. They view working memory not only as a storehouse for items and partial results of an analysis, but also as being used for carrying out processing. Much of their work has examined individual differences in working memory capacity and their effects on sentence comprehension. The measure of working memory capacity that they have adopted is "sentence span." This measure is obtained by having subjects read a set of sentences aloud and having them attempt to retain the last word of each sentence. The sentence span is determined by the maximum number of final words the subjects is able to recall. Just and Carpenter have reported that this measure correlates more reliably with comprehension measures than traditional memory span measures. They argue that this is because the sentence span measure reflects processing capacity and storage capacity while traditional span tasks reflect only storage capacity. In addition they claim that the memory capacities of individuals vary a great deal, which in turn affects their language comprehension.

This issue was investigated in our second experiment by calculating the effect of the memory burden on comprehension for each subject and then correlating it with three different measures of memory span. If the Just and Carpenter sentence span tasks is sensitive to semantic capacity as well as to processing, then based on the data from AB we would expect a correlation with the
subject's data from Experiment 2. However, there should not be a correlation with digit span, since digit span is primarily tapping phonological short-term memory. The suppression task was included to see whether it would be similar to sentence span task. Since both tasks are looking at retention of information when rehearsal of that information is not possible, one might hypothesize that they were both measures of capacity, as well as processing. If the Just and Carpenter hypothesis is correct, then the effect of the memory burden on sentence processing should be less for those subjects with a greater working memory capacity.

Method

Subjects: Sixty subjects from Rice University undergraduate participated in this experiment. All subjects were required to be native speakers of English. Each subject participated in a one-hour session. Fifty-seven subjects returned for the short-term memory tasks.

Materials-Anomaly Judgment: The 576 experimental sentences that were used in Experiment 1 were used in this experiment. In addition, 208 filler sentences were included. Of these 208 filler sentences, 72 were modeled after the adjective phrase sentences, 72 after the noun phrase sentences, and 64 were relative clause sentences. In this experiment, there was an equal number of
anomalous and sensible sentences. Reaction times (RTs) were measured in milliseconds from the onset of the first word that was heard after the completion of the anomaly. For example, "The soda, the rice and the chicken heated quickly on the stove." In this case the reaction time was measured from the onset of 'quickly'. RTs in the sensible sentences were measured from the corresponding word, e.g. "The soup, the rice and the chicken heated quickly on the stove."

**Procedure:** Subjects were instructed to place one index finger on a key labeled 'yes' and the other index finger on a key labeled 'no'. They keys were set up so the dominant hand was always used for the 'no' response. The subjects were instructed to press the key labeled 'no' as soon as the sentence no longer made sense. They were told to press the 'yes' key if the sentence was completed and no anomaly was detected. After each trial, the subject pressed the 'return' key in order to go on to the next trial. Each subject heard only one of the 12 sentences that were made up from the matched blocks of "before" and "after" sentences. Since there were 48 groups of 12 sentences, each subject heard 48 experimental sentences, which were divided into 24 adjective phrase sentences and 24 noun phrase sentences. Each subject participated in only one session in which they heard only 48 experimental sentences and 208 filler sentences. Each subject heard the same filler sentences. Each subject heard the same 10 practice sentences before their experimental session began. The experiment was presented through a PsychLab experiment
generation program (Gum and Bub, 1988) created for the Macintosh. The sentences were randomized before each presentation. Each session lasted about 45 minutes.

Materials and Procedure - Memory Span tasks: Subjects participated in three different memory span tasks; digit span, reading span, and word span under suppression. The digit span task from the WAIS-R (Wechsler, 1981) was used. Subjects were instructed to simply repeat the digits the experimenter said. Subjects were given two trials at each list length with list lengths ranging from 3-9 items. The task was discontinued when the subject missed both trials of the same length. Only materials from the forward direction repetition task were used.

In the reading span task, subjects read sentences that were typed individually on index cards. They were instructed to remember the last word from each sentence in the order in which the sentences were read. So, if they read three sentences, they would be asked to verbally recall three words. The trials ranged from 2-6 sentences, with three trials at each length. Testing was discontinued when subjects missed two of the three trials at a particular length.

In the word span task, subjects saw one word at a time on a monitor. Ten one syllable words were used which were matched on length and frequency. The subject's task was to remember the order in which the words appeared. However, while they were doing this,
they were also repeating the word "COLA" over and over. A verbal response was not required of the subjects. Instead, a sheet of paper was taped on to the keyboard such that each of the ten words in the corpus of words they could choose from corresponded to a number on the keyboard. Therefore, subjects had to find the word on the paper and press its corresponding number. All subjects were tested on list lengths of 3-5 items. The information was presented on a Macintosh Plus computer.

Results - Anomaly Judgments

An ANOVA (anomalous/sensible x before/after x distance) using reaction time as the dependent measure was performed. This analysis was done once using subjects as the random effect and once using sentences (items) as the random effect. Reaction times that were above 1200 msec and below 100 msec were thrown out. Based on this criterion 1% of the data were discarded.

The three-way interaction between anomalous/sensible, before/after, and distance was significant by subjects, $F(2,118) = 3.89$, $p< .02$, MSE = 78,4209 as well as by items, $F(2,85) = 3.39$, $p < .04$, MSE = 28,3091. For the sensible sentences there was no significant interaction between the before/after and distance variables. For the anomalous sentences, the interaction between before/after and distance was significant, $F(2,118) = 9.39$, $p< .0002$, MSE = 1,728,904. Table 2 shows the data for the anomalous sentences. This interaction was then examined further.
Table 2
Mean reaction times by subjects for the interaction of anomalous/sensible, before/after, and distance - Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Anomalous</td>
<td>1121</td>
</tr>
<tr>
<td>Sensible</td>
<td>1796</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Anomalous</td>
<td>1156</td>
</tr>
<tr>
<td>Sensible</td>
<td>2019</td>
</tr>
</tbody>
</table>

First, tests were done to determine if there was a significant difference between the before and after conditions at each level of distance. At distance 1, there was no significant difference, however at distance 2 and distance 3 the reaction times in the after condition were significantly faster than the reaction times of the before condition, $E (1,59) = 8.74, \ p < .01, \ MSe = 1,610,577, \ E (1,59) = 19.42, \ p < .01, \ MSe = 5,453,740$ respectively. Two-way comparisons were then carried out. The comparisons of the before and after conditions at distance 1 to distance 2 and distance 1 to distance 3 were significant, $E (1,59) = 8.19, \ p < .01, \ MSe = 1,185,275 \ E (1,59) =$
18.14, \( p < .01 \), MSe = 3,419,220 respectively. However, the comparison between distance 2 and 3 was not significant.

Analyses were done looking at the before and after sentences separately. For the sentences in the before condition, there were no significant differences between reaction times at any of the three levels of distance. However, for the sentences in the after condition, all three comparisons were significant. Distance 1 was significantly different from distance 2 and distance 3, \( F(1,59) = 10.54 \), \( p < .01 \), MSe = 1,701,556, \( F(1,59) = 29.23 \), \( p < .01 \), MSe = 5,481,082 respectively and distance 2 was significantly different from distance 3, \( F(1,59) = 5.41 \), \( p < .02 \), MSe = 1,076,078.

Appendix B lists all significant interactions. Although some of the two-way interactions were significant, they are not of interest. For example, the two-way interactions with the sensible/anomalous variable are not particularly interesting because the decision made by the subject did not occur at the same point in the sentence. In the anomalous sentences, the decision occurred as soon as the anomaly occurred, whereas for the sensible sentences, the decision occurred at the end of the sentence. Since there was no significant four-way interaction the noun phrase and adjective phrase sentences were not analyzed separately. However, a closer look at the significant three-way interaction between the phrasetype \( x \) before/after \( x \) and sensible/anomalous variables, as shown in Table 3, shows that the pattern of results for the noun phrase sentences is closer to that of the hypothesized pattern of results.
Table 3
Mean reaction times by subjects for the interaction of anomalous/sensible, before/after, and phrasetype - Experiment 2

**Adjective Phrase**

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensible</td>
<td>1681</td>
<td>2231</td>
</tr>
<tr>
<td>Anomalous</td>
<td>989</td>
<td>996</td>
</tr>
</tbody>
</table>

**Noun Phrase**

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensible</td>
<td>1947</td>
<td>1816</td>
</tr>
<tr>
<td>Anomalous</td>
<td>1310</td>
<td>996</td>
</tr>
</tbody>
</table>

Although is unclear as to the reason for this difference, it cannot be attributed to task difficulty since the error analysis did not reveal a significant difference between these variables.

It should be noted that the error analysis included only anomalous sentences because the few errors made on the sensible sentences occurred only on those sentences that had extreme reaction times and were thus discarded from the analysis. The analysis shows a main effect of the before/after condition $F(1,59) = 8.52, p < .01, \text{MSE} = 0.45$ with more errors being made in the before condition (13%) versus in the after condition (8%). There was also a main effect of distance, $F(2,118) = 4.57, p < .01, \text{MSE} = 0.33$ which
shows that as distance increased the number of errors also increased. This data is shown in Table 4.

Table 4
**Error rates for before/after by distance - Experiment 2**

<table>
<thead>
<tr>
<th>Distance</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>.08</td>
<td>.13</td>
<td>.17</td>
</tr>
<tr>
<td>After</td>
<td>.06</td>
<td>.09</td>
<td>.08</td>
</tr>
<tr>
<td>Difference</td>
<td>.02</td>
<td>.04</td>
<td>.09</td>
</tr>
</tbody>
</table>

Another interesting finding with regard to the error rates is that there were approximately twice as many errors at distance 2 and 3 for the 'before' condition than there were in the 'after' condition. This indicates that the task was harder for the subjects in the before condition at these levels of distance, presumably due to the memory burden. Although it appears that the difference in error rates between before and after is larger at distance 2 and 3 than it is at distance 1, the interaction failed to reach significance, \( F(2, 118) = 1.80, p = .17, \text{MSe} = 0.15 \).

The anomalous filler sentences were also analyzed in order to determine if the point of judgement in the sentence influence subjects' reaction times. The noun phrase and adjective phrase sentences were analyzed together. The sentences were coded
according to whether they were 'before' or 'after' sentences. These sentences fit the same criterion established for the experimental sentences. For example, the sentence, "The old, striped, dirty pants were talking to the shirts" is considered a 'before' sentence since the adjectives precede the noun. An example of sentence in the 'after' condition is, "The symphony was pleasant, melodic and pretty so we booed loudly," since the adjectives follow the noun. The difference between these sentences and the experimental sentences is that the word that completes the anomaly can occur at any place in the sentence. An analysis was done separately for each subject in which reaction time was regressed over word/target position. The variable word/target position was calculated for each filler sentence by counting the position of the word that occurred after the completion of the anomaly. After the slope for each subject was calculated separately for the before and after conditions a t-test was done to see if the slopes differed from zero. Both analyses were significant. In the after condition $t = -9.01, p < .001$ while in the before condition $t = -3.72, p < .01$. The slope for the after condition is $b = -59.5$ and for the before condition, $b = -24.6$. This indicates that as the point of judgement occurs later in the sentence, the subject responds faster. An overall analysis was done for the relative clause filler sentences since they were not divided into before and after conditions. This analysis also supported the findings showing that as the point of judgement occurred later in the sentence, reaction times decreased, $t= -12.89, p< .001, b = 84.10$. 
Results - Memory Span Tasks

The effect of the memory burden on comprehension was calculated for each subject. Only anomalous sentences were included in this analysis. This was done by first finding the slope for the before and after conditions for each subject and then subtracting these values. A second analysis subtracted the mean of the after condition from the mean of the before condition. Similar analyses were carried out on the error data. These differences were then correlated with the three different measures of memory. Surprisingly, the only correlation that was significant was the negative correlation between the errors and the subjects reaction time scores, $r = -0.28, p < 0.03$. In other words, as reaction times decreased errors increased. These correlations can be seen in Appendix C.

Although some of the other correlations, diffrt* digitspan, $r = 0.25, p < 0.06$ and diffrt*sentspan $r = 0.25, p < 0.06$, were close to significance, they were in the wrong direction. Diffrt refers to the measure in which the mean of the after condition was subtracted from the mean of the before condition for each subject. In other words, an increase in the difference in reaction times between the before and after conditions indicates that the subjects is affected by the memory burden. This should be correlated with a smaller span/capacity as measured by the memory tasks reported above. However, this is not the case as evidenced by these positive correlations.
Discussion

The results of Experiment 2 provide evidence that the structural and semantic variables that were manipulated in the experimental sentences created a memory burden for the subject, which in turn affected comprehension of those sentences. This result is supported by the error data as well as the reaction time data. The error data shows a main effect of error, as well as distance. More errors were made in the before condition, where the subject has a greater memory burden than in the after condition. Twice as many errors were made at levels of distance 2 and 3 in the before condition than in the after condition.

The hypothesized outcome for the reaction time data was that in the after condition, reaction times should be the same at all three levels of distance. This is based on the idea that immediate integration can occur in the after condition regardless of the numbers of lexical items in the sentence. Therefore, there should be no effect of the manipulation of distance. In addition, these reaction times should be faster or equal to the reaction times for the before sentences at distance 1. The reaction times for the before condition, however, should increase as distance increases. This prediction is based on the idea that at distance 1 the anomaly is salient, making the anomaly judgment easy. However, in the before condition, as the distance increases between the two words that create the anomaly, the memory burden also increases, making the anomaly less salient. This in turn makes the decision more
difficult. This pattern, however, was not found. Instead, the reaction times for the before condition were the same across the levels of distance, while the reaction times for the after condition decreased. Figure 3 shows the predicted and obtained results.

The difference between the predicted and obtained results can be partially explained by the results of the filler analysis as well as results found in Marslen-Wilson and Tyler (1980). These results show that the decision of the subject is facilitated when the point of the decision or response occurs later in the sentence. This can be explained by the fact the more semantic and syntactic information the listener has, the easier it is to anticipate where the decision will be made. Therefore, this facilitation must be taken into account when looking at the reaction time data since it is confounded with the distance manipulation in both the before and after condition. Once this is accounted for, the results become somewhat more compatible with those that were hypothesized. However, this explanation does not completely account for difference between the predicted and obtained results since the decrease in reaction time found for the filler sentences is less than that found for the experimental sentences. It is possible that the decrease in reaction time can be partially explained by a speed-accuracy trade-off, as evidenced by the significant negative correlation between reaction time and error rate. The problem with this explanation, however, is that this trade-off shows up only for the anomalous sentences and not for the sensible sentences.
**Obtained Results**

**Predicted Results**

![Graph showing reaction time vs distance for obtained results and predicted results.](image)

**Figure 3:** Experiment 2 - Predicted results as compared to obtained results
Although it is difficult to completely account for the decrease in reaction time across the levels of distance, it is still clear that the pattern of both the reaction time data and the error data support the idea that unintegrated information results in a memory burden for the subject that in turn affects comprehension of that material.

It is unclear as to why the results of the correlation between the memory span tasks and Experiment 2 were not significant. The error data for AB, the patient with a semantic short-term memory deficit, showed that in the before condition errors increased as distance increased however, in the after condition they stayed the same across the levels of distance. Based on this, we expected that subjects who were more affected by the memory burden in Experiment 2 would have a smaller processing capacity and span than those subjects who were not as affected by the memory burden. Not only did the results fail to support this hypothesis, but we also failed to find a significant correlation between the different span tasks we used. One reason why there was no correlation could perhaps be explained by saying that the different tasks are tapping different aspects or even different levels of complexity of short-term memory. Since none of the memory tasks actually required the subject to use semantic information to perform the task, it could be that the anomaly judgment task is tapping semantic memory, while the other tasks, including the sentence span task, are not. It is also possible that, although the range of memory span was large, there
were not enough subjects included at each of the different levels of memory span.

General Discussion

The purpose of this study was to investigate the role of the semantic component of short-term memory in sentence comprehension. The interest regarding this component stems from patient data, which indicates that in addition to a phonological component of short-term memory, there is also a semantic component (Martin, Shelton and Yaffee, 1994). The investigation of this semantic component is of interest since it presumably affects comprehension of a variety of tasks. The question of whether or not we could experimentally manipulate the appropriate variables in order to gain insight into this issue was addressed in Experiment 1 - Word Detection and Experiment 2 - Anomaly Judgements. In order to successfully investigate the issue the task involved must require the subject to semantically process the sentence. The experimental manipulations were effective at differentially affecting response times and error rates indicating that we had influenced the memory of the subject. The inability to immediately integrate lexical information slowed down sentence processing, while immediate integration facilitated it. It is interesting that a seemingly small amount of information can tax our short-term memory to such an extent. The fact that the unintegrated information in the sentence can place a greater memory burden on the subject than the
integrated information is both interesting and a little surprising. It was unclear to us originally whether the sentence manipulations would be successful in creating different memory burdens for the normal subjects since 3-4 words might well be within their capacity. Although in patients such as AB, the damage to this component results in profound impairments on tasks that require the use of semantic short-term memory, it is clear that normal subjects can also be impaired, although to a lesser extent when faced with similar tasks.

In general, the immediacy of processing theory (Marslen-Wilson and Tyler, 1980), which claims that the listener immediately processes and integrates information to the extent possible, is supported by the difference in results between the before and after conditions. Both the reaction time and error rate data indicate that when lexical information is not able to be immediately integrated, memory is taxed and comprehension suffers. However, when the lexical items can be immediately integrated in the sentence, the subjects' memory is less burden thus facilitating comprehension.
References


Appendix A: Significant Interactions from Experiment 1 - Word Detection

Distance * Phrasetype - \( p < .04 \)

Bef/Aft * Phrasetype - \( p < .001 \)

Anom/Sens * Bef/Aft * Phrasetype - \( p < .06 \)
Appendix B: Significant Interactions from Experiment 2 - Anomaly Judgment

Anom/Sens * Bef/Aft \( p < .001 \)

Anom/Sens * Phrasetype \( p < .001 \)

Bef/Aft * Phrasetype \( p < .001 \)

Bef/Aft * Distance \( p < .01 \)

Anom/Sens * Bef/Aft * Phrasetype \( p < .001 \)

Anom/Sens * Bef/Aft * Distance \( p < .02 \)
Appendix C

Correlation table for subjects performance on Experiment 2 with memory span tasks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sloperta</td>
<td>-160.0</td>
<td>215.0</td>
</tr>
<tr>
<td>Slopertb</td>
<td>22.33</td>
<td>216.5</td>
</tr>
<tr>
<td>Differr</td>
<td>0.049</td>
<td>0.135</td>
</tr>
<tr>
<td>Diffrt</td>
<td>155.8</td>
<td>339.3</td>
</tr>
<tr>
<td>Digitspan</td>
<td>7.86</td>
<td>1.06</td>
</tr>
<tr>
<td>Sentspan</td>
<td>3.21</td>
<td>0.977</td>
</tr>
<tr>
<td>Wordspan</td>
<td>3.15</td>
<td>1.56</td>
</tr>
</tbody>
</table>
Appendix C continued

- The upper value represents the correlation
- The lower number represents the significance level

<table>
<thead>
<tr>
<th></th>
<th>Sloperta</th>
<th>Slopertb</th>
<th>Slopertdiff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sloperta</td>
<td>1.00</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.96</td>
<td>0.45</td>
</tr>
<tr>
<td>Slopertb</td>
<td>0.01</td>
<td>1.00</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.0</td>
<td>0.66</td>
</tr>
<tr>
<td>Slopertdiff</td>
<td>0.10</td>
<td>-0.06</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>0.66</td>
<td>0.0</td>
</tr>
<tr>
<td>Differr</td>
<td>0.07</td>
<td>-0.08</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>0.60</td>
<td>0.56</td>
<td>0.16</td>
</tr>
<tr>
<td>Diffrt</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.8</td>
<td>0.99</td>
</tr>
<tr>
<td>Digitspan</td>
<td>0.27</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>0.62</td>
<td>0.96</td>
</tr>
<tr>
<td>Sentspan</td>
<td>-0.10</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>Wordspan</td>
<td>0.03</td>
<td>0.14</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>0.30</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Appendix C continued

<table>
<thead>
<tr>
<th></th>
<th>Differ</th>
<th>Diff rt</th>
<th>Digitspan</th>
<th>Sentspan</th>
<th>Wordspan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sloperta</td>
<td>0.07</td>
<td>-0.02</td>
<td>0.27</td>
<td>-0.10</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>0.60</td>
<td>0.90</td>
<td>0.09</td>
<td>0.45</td>
<td>0.85</td>
</tr>
<tr>
<td>Slopertb</td>
<td>-0.08</td>
<td>-0.03</td>
<td>0.07</td>
<td>-0.02</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>0.56</td>
<td>0.8</td>
<td>0.62</td>
<td>0.90</td>
<td>0.30</td>
</tr>
<tr>
<td>Slopertdif</td>
<td>0.18</td>
<td>-0.001</td>
<td>0.01</td>
<td>-0.01</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>0.16</td>
<td>0.99</td>
<td>0.96</td>
<td>0.95</td>
<td>0.60</td>
</tr>
<tr>
<td>Differ</td>
<td>1.00</td>
<td>-0.28</td>
<td>-0.16</td>
<td>0.01</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.03</td>
<td>0.23</td>
<td>0.93</td>
<td>0.51</td>
</tr>
<tr>
<td>Diff rt</td>
<td>-0.28</td>
<td>1.00</td>
<td>0.25</td>
<td>0.25</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>0.0</td>
<td>0.06</td>
<td>0.06</td>
<td>0.59</td>
</tr>
<tr>
<td>Digitspan</td>
<td>-0.16</td>
<td>0.25</td>
<td>1.00</td>
<td>-0.03</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
<td>0.06</td>
<td>0.0</td>
<td>0.77</td>
<td>0.64</td>
</tr>
<tr>
<td>Sentspan</td>
<td>0.11</td>
<td>0.25</td>
<td>-0.04</td>
<td>1.00</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>0.06</td>
<td>0.76</td>
<td>0.0</td>
<td>0.14</td>
</tr>
<tr>
<td>Wordspan</td>
<td>-0.09</td>
<td>0.08</td>
<td>-0.07</td>
<td>0.20</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.59</td>
<td>0.64</td>
<td>0.14</td>
<td>0.0</td>
</tr>
</tbody>
</table>