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 5” 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.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
RICE UNIVERSITY

THE CONTRIBUTIONS OF SEMANTIC AND PHONOLOGICAL REPRESENTATIONS IN VERBAL RECALL: A STUDY USING SENTENCE REPETITION

by

MICHAEL A. KATZ

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

APPROVED, THESIS COMMITTEE

Randi C. Martin, Director
Elma W. Schneider Professor of Psychology

Jeffrey S. Bowers, Assistant Professor
Psychology

David M. Lane, Associate Professor
Psychology

Houston, Texas

May, 1998
ABSTRACT

The Contributions of Semantic and Phonological Representations in Verbal Recall: A Study Using Sentence Repetition

by

Michael A. Katz

Five experiments investigated whether phonological representations play a role in verbatim sentence repetition. Experiments 1 and 2 revealed that words related both semantically and phonologically to sentence words were more likely to be intruded into sentences than were words related only semantically. Experiment 4, the most stringently controlled, also yielded more intrusions for semantically and phonologically related lures, with a larger effect for abstract than concrete stimuli. Experiment 5 demonstrated that words related only phonologically to sentence words could be intruded into sentences as well. It is argued that veridical sentence recall taps phonological as well as semantic memory representations for accurate performance; a view inconsistent with the “Semantic Regeneration Hypothesis” offered by Potter and Lombardi (1990). Further, it is argued that the semantic and phonological representations act in an interactive fashion and support each other during sentence recall.
Acknowledgments

I would like to express my sincere appreciation to all who have assisted me during the completion of this thesis. I would like to give special thanks to my advisor Dr. Randi Martin for continuous guidance and support throughout the duration of this project. I would also like to thank Dr. Jeffrey Bowers and Dr. David Lane who served as committee members. I am also grateful to Markus Damian who provided much assistance in the early stages of this project. Finally, I wish to express my sincere gratitude to my family and friends whose support has been invaluable.
# TABLE OF CONTENTS

Abstract .......................................................................................................................... p. ii

Acknowledgments .......................................................................................................... p. iii

Table of Contents .......................................................................................................... p. iv

List of Tables ................................................................................................................ p. vii

Three Accounts of Immediate Verbatim Sentence Repetition .................................. p. 1

The Semantic Regeneration Hypothesis ..................................................................... p. 10

**EXPERIMENT 1** ........................................................................................................ p. 14

  Method ......................................................................................................................... p. 14

  Participants ................................................................................................................ p. 14

  Materials .................................................................................................................... p. 14

  Stimulus Ratings ...................................................................................................... p. 15

  Design and Procedure .............................................................................................. p. 15

  Analysis ...................................................................................................................... p. 16

Results and Discussion ............................................................................................... p. 17

**EXPERIMENT 2** ........................................................................................................ p. 21

  Method ......................................................................................................................... p. 22

  Participants ................................................................................................................ p. 22

  Materials .................................................................................................................... p. 22

  Design and Procedure .............................................................................................. p. 22

Results and Discussion ............................................................................................... p. 22
EXPERIMENT 3 ................................................................. p. 23
Method ........................................................................ p. 24
Participants ............................................................... p. 24
Materials ................................................................. p. 24
Design and Procedure .................................................. p. 24
Results and Discussion ................................................ p. 24
Cloze Comparison of Potter and Lombardi (1990) and Experiments 1 and 2 ........ p. 25
EXPERIMENT 4 ............................................................. p. 28
Method ........................................................................ p. 29
Participants ............................................................... p. 29
Materials ................................................................. p. 29
Stimulus Ratings ........................................................ p. 30
Design and Procedure .................................................. p. 33
Results and Discussion ................................................ p. 33
Supplementary Analysis of Serial Position Effects .............................................. p. 38
EXPERIMENT 5 ............................................................. p. 43
Method ........................................................................ p. 44
Participants ............................................................... p. 44
Materials ................................................................. p. 44
Design and Procedure .................................................. p. 44
Results and Discussion ................................................ p. 45
GENERAL DISCUSSION .................................................... p. 49
The Role of a Phonological Representation in Verbatim Recall ........................................p. 49

The Multiple Capacity View of Short-Term Memory (STM)...........................................p. 53

The Mechanism of Verbatim Sentence Repetition: Two Frameworks ..............................p. 58

Conclusions ......................................................................................................................p. 65

Footnotes .........................................................................................................................p. 66

References .......................................................................................................................p. 69
List of Tables

Table 1  Proportion of Possible Trials in which a Substitution Occurred by Condition for Potter and Lombardi (1990) and Experiments 1, 2, 3, and 5

Table 2  Cloze Ratings for Potter & Lombardi (1990) and Current Experiments 1 and 2

Table 3  Cloze Ratings (%) for Stimuli used in Experiment 4

Table 4  Results of Experiment 4

Table 5  Substitution Rates (%) by Condition as a Function of Location Group for ABSTRACT Targets

Table 6  Substitution Rates (%) by Condition as a Function of Location Group for CONCRETE Targets

Table 7  Mean Phonological Substitution Rates (%) as a Function of Location Group
Three Accounts of Immediate Verbatim Sentence Repetition

The seemingly straightforward process of immediate verbatim sentence repetition has garnered some controversy in recent years. Three theoretical accounts exist to explain such an ability and they can be classified as the "Phonological," "Conceptual," and "Combination" views. The Phonological account, which has historically been the most widely accepted, argues that immediate sentence repetition is performed using a verbatim phonological record of the sentence. It must be noted that although many different views exist under the broad "Phonological" view heading, all agree that some kind of non-semantic surface representation of the sentence is responsible for verbatim repetition. Much previous research is equivocal in terms of the specific nature of the surface representation involved; the representation may be a pre-lexical sound-based code or a lexical-phonological record which preserves word order (see Butterworth, Shallice, & Watson, 1990, for evidence that the phonological record contains order information). This issue was explored further in the current study. The Conceptual view states that a meaning-based representation of the sentence is used for repetition in lieu of a surface representation of any kind. Finally, the Combination account argues that both conceptual and phonological information are used (either additively or interactively) to support immediate verbatim repetition of sentences. Specifically, this account assumes that listeners develop a conceptual representation and use this in conjunction with surviving phonological information to repeat a sentence.

Initial credibility for the Phonological account comes from list recall studies
which clearly demonstrate a role for a phonological code. Experimental phenomena such as the phonological similarity effect (e.g. Conrad & Hull, 1964) and the word length effect (e.g. Baddeley, Thomson, & Buchanan, 1975) strongly support the existence of a phonological code which supports immediate word list recall. Evidence from the realm of sentence repetition comes from studies by Jarvella and others (e.g. Jarvella, 1971; Glanzer, Dorfman, & Kaplan, 1981; Anderson & Paulson, 1977; Butterworth, Shallice, & Watson, 1990). Jarvella (1971) found that running memory span was approximately two sentences (12 words for “long” condition in Experiment 2) although verbatim recall had a span of approximately seven words (the most recent clause). A similar measure of span was found by Wingfield and Butterworth (1984) using a task which required subjects to listen to an auditorily presented passage of text, stop the recording at self-selected intervals, and recall the immediately preceding text. Using this method, they found that the average segment length chosen by the subjects (for recall) was 10.6 words (although the accuracy of recall was 80%).

A measure of verbatim span of approximately two sentences was found in a study by Glanzer et al. (1981) in which paragraphs were presented auditorily with interruptions which required the subjects to recall a certain one of the last few sentences verbatim. The finding that two recently presented sentences could be recalled verbatim can be construed as strong support for the Phonological position. That is, such accurate recall implies storage of the precise surface form of the sentence instead of a more deeply processed semantic representation. Although other evidence (presented later) seems to suggest that recall of more words in sentences than in lists supports the existence of a conceptual
representation used during sentence recall and not a phonological one, confusion can be avoided if one focuses on different aspects of the result. Glanzer et al.'s (1981) data, for example, can be construed as evidence for the phonological position if one focuses on the perfect accuracy of recall whereas it can serve as evidence for the conceptual position if one focuses on the great amount of information that could be correctly recalled.

In another study by Glanzer and colleagues (Glanzer, Fischer, & Dorfman 1984), subjects were presented with paragraphs to read with self-paced presentation of sentences and time to read each sentence was recorded. Following paragraphs were comprehension questions which subjects answered orally. In one condition, however, paragraphs contained interleaved unrelated factual statements with corresponding comprehension questions presented later (in conjunction with comprehension questions corresponding to the paragraph content). It was found that presenting unrelated factual statements between sentences of a paragraph slowed reading time of paragraph sentences but did not affect reading times for the factual statements themselves. They argued that the interrupting task interfered with carryover information used to comprehend organized text. More relevant to the present issue, however, they also found that presentation of a word before each paragraph sentence which was consistent with the theme of the paragraph did not compensate for the slowed reading times. Glanzer et al. concluded that it was the disruption of verbatim information that slowed reading times and not disruption of thematic or conceptual information.

They supported this argument by presenting paragraphs with interruptions halfway through, at which point a segment (a few sentences) from a different paragraph was
presented, then followed by “old” information, and finally the remainder of the paragraph. This “old” information consisted of one of the three following conditions: Nothing (“no old information” condition); repetition of a sentence seen immediately prior to the interruption; or repetition of an earlier seen general “thematic” sentence (a sentence thematically consistent with the main idea of the paragraph). In a previous experiment, they demonstrated a significant reaction time benefit for reading new sentences of the paragraph (presented after the “old” information) if the two sentences before the interruption were repeated first (as opposed to no repetition). In the current experiment, no such reaction time benefit was obtained by presenting a thematically consistent sentence after the interruption and before the new sentences supporting the notion that verbatim representations of previous sentences are required. Glanzer et al. argued that since subjects have been shown to have two sentences available for verbatim recall, it is likely that this information is used, and they suggest that its role lies in the process of reading comprehension:

“The function of short-term storage is to facilitate the linkage of successive sentences of the text. The linkage is based on anaphoric and sequential relations. Storage in verbatim form gives readers a major advantage in handling the anaphoric and sequential demands of each new sentence as it is read...If the subject has only highly processed information available in short-term storage, then a pronoun or connective that appears in the next read sentence may not be easily interpretable. The highly processed information may not contain the specific information needed” (p. 484).
As can be inferred from their argument, Glanzer et al. (1984) placed little emphasis on a conceptual record ("highly processed information") although they undoubtedly believed that verbatim information is not the only information carried over during reading comprehension. Furthermore, the verbatim representation they described referred to a representation of the sentence that contained information about its surface structure as well as all of its words. Although phonology was not tested in their experiments, per se, Glanzer et al.’s notion of a verbatim representation is consistent with the definition provided in the introduction.

Unfortunately, many criticisms can be raised against the conclusions of Glanzer and colleagues. According to Gernsbacher (1990), after the interruption during which sentences from a new paragraph are presented, a new "structure" is being built for the new paragraph which leads to poorer accessibility of the structure of the original paragraph information, resulting in general slowing. The fact that less slowing occurred following presentation of verbatim versus thematic sentences could simply be due to the differential degrees of difficulty of reinstating old information that was presented prior to the interruption. The verbatim sentence allowed a rapid reinstatement of previous information whereas the thematic sentence would require more interpretive analysis to do so due to the lack of overt connective information (e.g. antecedents to pronouns). Therefore, there is no need to postulate the use of a "verbatim" representation in the Glanzer task; any representation which contained the necessary sequential information (for paragraph comprehension) would be sufficient.

The second approach to immediate sentence recall emphasizes the role of a
conceptual record. Most theorists would advocate the existence of a conceptual record utilized in sentence processing tasks and the need for a conceptual representation during sentence repetition has been confirmed by studies demonstrating exact repetition for sentences greatly exceeding memory span (e.g. Wingfield & Butterworth, 1984); an ability which is presumably due to the contribution of semantic as well as syntactic structure (Tejirian, 1968). In addition, other support comes from studies which suggest that a semantic representation of a heard sentence is formed immediately (e.g. Cole & Jakimik, 1980; Marslen-Wilson & Tyler, 1980; Marslen-Wilson, 1975; see Von Eckardt & Potter, 1985 for additional references).\(^1\)\(^2\) Von Eckardt and Potter (1985) have also provided evidence which contradicts the once popular notion that a semantic representation of a heard sentence is derived from a phonological representation; instead, they argued that a conceptual representation is formed in parallel with or even in advance of a (presumably post-acoustic) phonological representation. Von Eckardt and Potter assigned a more prominent role to the conceptual representation than other theorists and argued that the quickly formed semantic representation of a sentence was primarily responsible for immediate memory performance.

Their experiment was based on Caplan’s (1972) probe recognition procedure. In Caplan’s experiments, subjects were presented auditorily with a two-clause sentence followed 150 msec later by a spoken or written probe word and the task was to determine whether or not the probe word had appeared in the sentence. It was found that responses to words that had appeared in the second clause were 67 msec faster than responses to words from the first clause, even though distance from the test word to the end of the
sentence was matched across conditions. In Von Eckardt and Potter’s experiment, subjects were presented with a probe matching task in which a word or picture was presented and the subject was required to determine if the word or the word corresponding to the picture had been presented in a sentence. They found that for both words and pictures, the time to respond to the probe was greater when the corresponding word in the sentence had been presented in the first as opposed to the second clause. Therefore, they explained Caplan’s (1972) finding of faster probe responses to words in the second clause (as opposed to first clause) of a sentence as reflecting a greater availability of semantic information for the second clause. Previously, this result had been interpreted as reflecting highly available semantic information in the first clause with highly available surface (verbatim) information in the second clause. However, if such an account were true, assuming the comparison of the picture to the sentence were based on semantics, one would have expected faster responses to picture probes when the corresponding word was in the first clause of the sentence (as opposed to the second). Therefore, Von Eckardt and Potter argued for a much more prominent role for conceptual information in performance of immediate memory tasks.

The positions reviewed thus far have placed emphasis on either a phonological code or a conceptual code during verbatim repetition. However, a third position on this issue might be that semantic and phonological information are both important and that they support each other during verbatim repetition. Indeed, this Combination or “dual-code” account of verbatim sentence repetition has been supported by evidence from brain damaged patients. Caramazza, Basili, Koller, and Berndt (1981) argued that the poor
repetition performance found in conduction aphasia was due to a deficit affecting auditory-verbal STM. More recently, motivated by a multiple capacity view of short-term memory which incorporates semantic as well as other representations, Martin, Shelton, and Yaffee (1994) studied a patient EA who was determined to have an impaired phonological short-term memory (STM) capacity in conjunction with an intact semantic STM capacity. On a task of sentence repetition, EA made errors which preserved the overall semantic representation of the sentence while neglecting its precise wording. In fact, on a verbatim sentence repetition task, she was correct on 0% of syntactically simple sentences and 3% of syntactically complex sentences (averaged over all sentence types for each complexity group). However, meaning was preserved on 80% of adverbial clause and 60% of conjoined structures; the mean percentage of responses classified as paraphrasal errors collapsed across all complex structures was 50%. Another patient (AB) was found to have the reverse pattern of impairments (impaired semantic STM with intact phonological STM). He was as high as 68% correct on some simple syntactic sentence structures (for verbatim repetition). Notably, EA’s error types were mainly semantic substitutions while those of AB were mostly omission errors. EA’s errors were thought to be due to her intact conceptual representation of the sentence unaided by a phonological record, and on the basis of such evidence, Martin et al. argued that a phonological record was indeed necessary for normal performance on the verbatim sentence repetition task. Specifically, they argued that the conceptual representation and the surviving phonological record of the sentence are integrated in some fashion during repetition.
Saffran and Marin (1975) presented the case of a patient who, like EA, demonstrated difficulty in retaining phonological information. This patient paraphrased sentences he could not correctly repeat and this result is consistent with the notion that verbatim repetition cannot rely solely on a conceptual representation. In addition, further neuropsychological evidence is provided by studies of patients presenting a pattern of impaired repetition in the face of intact comprehension (e.g. Butterworth et al., 1986), again suggesting that adequate processing and retention of semantic information in the absence of phonological information is insufficient to support verbatim repetition. As an interesting aside, the converse impairment does not necessarily hold. Martin and Saffran (1990), for instance, found that a patient with transcortical sensory aphasia showed the ability to repeat sentences she could not understand (although she was poor at repeating word lists). Although not the focus of this paper, this serves as evidence of a syntactic representation held in short-term memory which can be beneficial to repetition performance.

Inconsistent with the notion that a conceptual representation is essential for verbatim recall, some evidence suggests that purely phonological substitutions are possible in sentence repetition. Wingfield, Tun, and Rosen (1995), in their analysis of age differences for veridical (verbatim) recall reported that although rare, responses existed in which phonologically related word substitutions were made. For example, one subject recalled “a humble frame of mind” as “a humble freight of mind.” Furthermore, nonsense word responses were sometimes made which also preserved some phonological aspects of the target word (e.g. “typhoid” --> “pipeoids”). It is important to note that for
both of these types of error, the semantic representation of the original stimulus was corrupted, suggesting the powerful influence of the phonological code. However, one could reasonably argue that subjects in the Wingfield et al. study may have simply misperceived the input. Therefore, more powerful evidence of the role of the phonological code is needed.

The Semantic Regeneration Hypothesis

Recently, Potter and Lombardi (1990), using a dual-task RSVP paradigm, put forth a direct challenge to the combination account of verbatim sentence repetition. They demonstrated that “semantic intrusion” errors could occur during a verbatim sentence repetition task with normal subjects. In their experiment, all stimuli were presented using rapid serial visual presentation (RSVP), that is, presented one word at a time in the center of the screen with a brief presentation time for each word. First, subjects saw five apparently unrelated words, then a sentence, and finally a capitalized word (the probe). Subjects were told to first make a button press response to indicate whether the probe was present on the list of the first five words; then, they were asked to repeat the sentence verbatim. This difficult task (which used sentences of between 10 and 15 words) was performed surprisingly well by the subjects, with 86% of the non-target sentence words correctly recalled.

The experimental manipulation of interest was the effect of the word list on sentence recall. On some trials a word was included that was closely related semantically
to a word in the sentence. Potter and Lombardi argued that if verbatim sentence repetition was indeed performed by using a phonological record of the sentence, then a semantically related word (with unrelated phonology) would not be expected to replace the word in the sentence to which it was related during repetition. However, such semantic intrusions occurred on 22% of the experimental trials. This observation led them to argue that a phonological record is not used during verbatim sentence repetition; specifically, they made the argument that the ordered auditory representation that is important in list repetition is not used in sentence repetition. Instead, they argued that the task was performed using conceptual and lexical memory stores.

More specifically, the overall meaning of a presented sentence would be stored in conceptual memory while a list of all recently activated (seen) words would be stored (unordered) in lexical memory. (Although the notion of lexical memory is not explained in detail by Potter and Lombardi, they appear to advocate perceptual activation of entries in a mental lexicon. These entries are presumed to contain no phonological information.) During verbatim recall, the conceptual representation of the sentence would preferentially select from the lexical store words which corresponded to the overall semantic representation of the sentence. Potter and Lombardi term this process “semantic regeneration”: the exact wording of the sentence is “regenerated” from its conceptual representation. Therefore, if two words existed in lexical memory which satisfied the conceptual representation of the sentence, then either could be selected during recall, with only one of the words actually having been presented in the sentence. To cite an example from Potter and Lombardi, if the word “reply” were presented in a sentence to be recalled
and "answer" were presented in the list of the first five words, then one might expect on their account that, on some trials, "answer" would be recalled in the sentence (incorrectly). Indeed, this account of verbatim repetition has received acceptance both within and outside the field of psycholinguistics (e.g. Crowder, 1993; Reder & Kusbit, 1991).

The current set of experiments was carried out to investigate the semantic regeneration hypothesis. The neuropsychological evidence cited earlier suggested that both semantic and phonological information are important to sentence repetition. The data reported by Potter and Lombardi do not provide strong evidence against a role for a phonological code; they argued that the semantic substitutions should have been prevented if a phonological code was involved. However, if both phonological and semantic codes are involved in recall, then purely semantic substitutions might occur for words for which the phonological code was weak. However, in general, one would expect a greater likelihood of substitutions of words that are both phonologically and semantically related compared to words that are only semantically related. In addition, it is possible that purely phonological substitutions could occur under certain conditions.

Previous evidence has therefore suggested a role for both phonological and semantic information during verbatim repetition. This issue was addressed by postulating an additive (or perhaps interactive) role for the two systems during verbatim sentence repetition in which semantic and phonological short-term memory codes potentially support each other during normal performance (for a similar suggestion, see Wingfield & Lindfield, 1995; Wingfield, Tun, & Rosen, 1995). In fact, the notion of multiple memory
capacities used in conjunction to support repetition performance is not a novel idea (e.g., Butterworth et al., 1986). Butterworth et al. argued that:

"[The phonological short-term store] may...have a role in sentence recall, where reconstructive processes will make use of phonological material in STS as well as more abstract representations of syntax and semantics" (p. 731).

In addition, inspired by the work of Anderson and Paulson (1977) who demonstrated the existence of a residual verbatim representation of a sentence even after long delays, the possibility that the phonological representation used during verbatim repetition is more long-term than previous models suggested (e.g. Baddeley, 1986) was explored, which might indicate a more prominent role for phonology when it is used for normal language processes.

To preview the following experiments: Experiment 1 introduced a new manipulation to the Potter and Lombardi (1990) paradigm which investigated the role of phonology in sentence repetition. Specifically, it compared substitution rates for semantically related words to those both semantically and phonologically related. Based on arguments by Dell (1986), it was predicted that if a phonologically-based memory representation of the presented sentence exists, this method should be adequate to reveal its influence on recall. Dell argued that a semantic error in production would be more likely when a semantic target shared phonetic segments with a semantically related intruding item, particularly when the overlapping segments were word initial. Therefore, the words to be used in Experiment 1 which are classified as both semantically and phonologically related (generally) had the first two phonemes in common with the target.
Experiments 2 and 3 attempted to replicate the findings from Potter and Lombardi (1990) and reconcile these with ours, while Experiment 4 investigated the role of concreteness in the aforementioned effects. Finally, Experiment 5 sought further confirmation that a phonological record is used in performance of normal verbatim sentence repetition by showing that purely phonological substitutions could occur.

EXPERIMENT 1

Participants. 36 Rice University undergraduate students participated in the experiment for course credit.

Materials. Sentences consisting of 10 to 15 words were created \( (M = 12.2 \text{ words}) \) with one content word in each sentence being classified as the "target." A list of five nouns was created for each sentence with one of the words (the "lure" word) being classified as one of the following types: a word semantically related to the target word \( (S) \); a word phonologically related to the target word \( (P) \); a word both semantically and phonologically related to the target word \( (SP; \text{ e.g. } \text{“inspection” to the target “investigation”}) \); a word unrelated to the target word. The list of five words were constrained so that they were similar in length to the target and lure for the sentence but unrelated in meaning to the words of the sentences. For each condition (type of relation), 10 sentences were used. Due to the imbalance between the number of sentences with lures in the list \( (N = 30) \) versus the number of sentences without a lure in the list \( (N = 10) \), 20 filler sentences which did not contain related lures were added to the stimuli. The recall responses were not recorded for the filler items.
Stimulus Ratings. Several comparisons of the stimuli were carried out to rule out confounding factors affecting the recall of S vs. SP lures. The mean difference in frequency for the SP ($M = 69$) and S ($M = 62$) lures was 7, which was not significant, $t(39) = .33, p > .5$. A semantic judgment task was administered to 20 subjects in which they were asked to rate the semantic similarity between the targets and the lures when each word was presented in the stimulus sentence (to be) used in the experiment (presented in pairs for the rating task). There were 4 (levels of relatedness factor) x 5 (possible unrelated words for each sentence) = 20 rating sets. The semantic relatedness to each target word for its corresponding SP and S lures was analyzed (on a scale of 1 to 10) and the SP lures ($M = 8.875$) were found to be only slightly more related to the targets overall than the S lures ($M = 8.735$) (i.e. Mean[SP - S] = .14, $t(19) = .80, p > .4$, n.s.). The mean difference from the target word ($M = 8.1$) in number of letters (using absolute values) for the SP lure ($M = 7.5$) and S lure ($M = 6.4$) (across all stimuli) was 1 letter (i.e. Mean [|$T - SP$] - |$T - S$|] = -1), $t(39) = -2.46, p < .05$. If a significant difference is found between the occurrence of SP and S intrusions, an ANCOVA will be performed (with length controlled as a covariate) to determine if the length difference played a role in the effect.

Design and Procedure. For the experiments, the DMASTR program developed by Ken Forster and Jonathan Forster at the University of Arizona was implemented on an IBM compatible 486 computer using a similar experimental paradigm as Potter and Lombardi (1990), Experiment 2. Four randomized versions of the materials were used, counterbalancing for the presence of the particular types of lure: Semantic and Phonological, Semantic, Phonological, or Unrelated. When a lure was present, it
replaced one of the five words of the list. The probe word would also be present as one of the first five words on half the trials; the lure word was never the probe nor was it ever the last word of the first five. There were six practice trials without lures present on the list.

On a particular trial, using RSVP in which all stimuli are presented briefly in the center of the screen, the following events occurred: Once the subject pressed the spacebar, a mask of asterisks was presented for 300 msec followed by a 350 msec blank; five words were then presented for 250 msec each, followed by a 250 msec mask of asterisks; then a sentence was presented one word at a time for 200 msec per word; finally, following a mask of $520^4$ msec, a “probe” word was presented for 500 msec. The first word of the sentence was capitalized but no punctuation was used.

Subjects were told that the study was an investigation of memory for sentences while doing a secondary task simultaneously. They were instructed to respond to the probe word by pressing the shift keys (corresponding to “yes” and “no”) according to whether the probe word was present on the list, and then to recall the sentence aloud to the experimenter as accurately as possible. The recall response was recorded by the experimenter for later scoring.

**Analysis.** The primary goal of this experiment was to determine if SP lure words were more likely than S lure words to be substituted for the target words in the sentences. Therefore, for this and all other experiments, the raw number of each type of substitution (or “intrusion”) was converted into the proportion of possible trials in which an intrusion occurred. For instance, if 2 SP intrusions occurred out of 10 sentences in which an SP lure was presented, the proportion would be 20%. Then, the mean proportion of intrusions
was computed for each lure type and an ANOVA was performed to determine if any
differences existed among conditions.

In addition, due to the fact that spontaneous intrusions (in which the lure word
was not presented on the list of the first five words) may be more likely for certain lure
words or lure types due to uncontrolled factors, an analysis of the (SP - S) difference was
performed using the spontaneous rates as a baseline (i.e., [(SP - SponSP) - (S - SponS)]).
This way, if an advantage of SP over S lures is found, the cause of the effect as due to
some property of the SP lures other than their phonological relation to the target can be
ruled out. In a sense, this is a conservative analysis since a greater spontaneous SP than S
substitution rate could be due solely to their greater phonological similarity to the target.

Results and Discussion

If there is a phonological record of the sentence being used during verbatim recall,
then one would expect more frequent intrusions of the SP than S lures. The SP lures
would provide a better match to the targets because they contain consistent phonological
as well as semantic information. The results are presented in Table 1. The data refer to
proportion of errors of a particular type out of the possible trials on which that type of
error could occur. An asterisk indicates that a particular type of intrusion was not
investigated in a particular study. “Spon” refers to “spontaneous intrusions” in which the
lure word was not presented on that trial but an intrusion of the lure word occurred
anyway. “Other Target Errors” refer to the proportion of trials in which non-intrusion
recall errors were made to the target. Most such errors were omissions, but less
frequently words (including the target word) or phrases of the sentence would be replaced
in favor of a different or more simplified interpretation. On rare occasions, a semantically
Table 1

Proportion of Possible Trials in which a Substitution Occurred by Condition for Potter and Lombardi (1990) and Experiments 1, 2, 3, and 5

<table>
<thead>
<tr>
<th>Experiment</th>
<th>SP</th>
<th>S</th>
<th>P</th>
<th>Spon SP</th>
<th>Spon S</th>
<th>Spon P</th>
<th>Other Target Errors</th>
<th>Nontarget Words Recalled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potter &amp; Lombardi (1990) (N=16)</td>
<td>*</td>
<td>22</td>
<td>*</td>
<td>*</td>
<td>9</td>
<td>*</td>
<td>8</td>
<td>86</td>
</tr>
<tr>
<td>Exp. 1:</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>85</td>
</tr>
<tr>
<td>Hard Sentences (N=36)</td>
<td>(70)</td>
<td>(76)</td>
<td>(75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exp. 2:</td>
<td>17</td>
<td>4</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>85</td>
</tr>
<tr>
<td>Easy Sentences (N=20)</td>
<td>(68)</td>
<td>(72)</td>
<td>(79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exp 3:</td>
<td>*</td>
<td>16</td>
<td>*</td>
<td>*</td>
<td>10</td>
<td>*</td>
<td>14</td>
<td>83</td>
</tr>
<tr>
<td>Replication (N=20)</td>
<td></td>
<td></td>
<td></td>
<td>(71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exp 5:</td>
<td>*</td>
<td>*</td>
<td>7</td>
<td>*</td>
<td>*</td>
<td>6</td>
<td>17</td>
<td>81</td>
</tr>
<tr>
<td>Phonological Lures (N=30)</td>
<td></td>
<td></td>
<td></td>
<td>(70)</td>
<td></td>
<td></td>
<td>PES: 3</td>
<td>POS: 4</td>
</tr>
</tbody>
</table>

(or semantically and phonologically) related word which did not appear as a lure on any trial would be substituted for the target. "Nontarget Words Recalled" refers to the proportion of non-target words from all sentences in the experiment that were recalled correctly. Finally, the numbers in parentheses refer to the proportion of target words correctly recalled for that particular condition.
Data from three subjects was replaced (with data from new subjects) due to a greater than 40% error rate on the probe recognition (button pressing) task. The mean percentage of correct responses on the probe task for the retained subjects was 77%. As shown in Table 1, subjects performed the recall task quite well, recalling 85% of the words correctly (excluding targets). (F₁ and t₁ refer to analyses performed with subjects as the random factor while F₂ and t₂ refer to analyses with items as the random factor.) A repeated measures ANOVA was carried out (on the proportion data) with two factors: induced vs. spontaneous error and lure type for the SP, S, and P intrusions. This analysis revealed a main effect of type of lure, F₁(2, 70) = 84.99, MSE = 31.95, p < .001, F₂(2, 78) = 20.57, MSE = 149.16, p < .001, with the proportion of SP intrusions exceeding those of S and P intrusions, as well as a main effect of induced vs. spontaneous substitutions, F₁(1, 35) = 5.71, MSE = 38.08, p < .05, F₂(1, 39) = 8.14, MSE = 33.41, p < .01, as shown by a higher rate of induced intrusions. The interaction between lure and the induced/spontaneous factor was also significant, F₁(2, 70) = 3.78, MSE = 33.23, p < .05, F₂(2, 78) = 5.03, MSE = 32.35, p < .05, with the difference between induced and spontaneous lures much greater for the SP than for the S and P conditions.

The analysis of primary interest revealed that, as predicted, a significantly greater proportion of substitutions occurred for the SP lures (14% of possible trials) compared to the S lures (2% of possible trials), t₁(35) = 5.80, t₂(39) = 4.51, ps < .001, suggesting the influence of a phonological record of the sentence during verbatim recall. Also, planned comparisons revealed that a significantly greater proportion of substitutions occurred for the spontaneous SP lures (9% of possible trials) compared to the spontaneous S lures (1% of possible trials), t₁(35) = 7.82, p < .001, t₂(39) = 2.87, p < .01. This significant
advantage of SP over S intrusions even under conditions in which the lure words were not presented also suggests a phonological representation in verbatim repetition. In addition, there was a higher rate of intrusions for SP vs. Spontaneous SP lures, $t_6(35) = 2.25$, $p < .05$, $t_6(39) = 2.50$, $p < .05$, suggesting that the presence of the lure word in the list increased the likelihood of a substitution.

Although the comparison of SP vs. S revealed a higher probability of substitution for the former type, it was argued earlier that a more conservative analysis would take the rate of spontaneous substitutions into account. Therefore, an interaction contrast was computed using spontaneous substitution rates as baseline: $[(SP - SponSP) - (S - SponS)]$. A marginally significant result was found, $t_6(35) = 1.73$, $p = .09$. The failure of the interaction contrast to reach significance even though the overall interaction of lure type and induced/spontaneous factor was significant can be explained by the exclusion of the P-lure substitutions for this analysis.

As mentioned in the Stimulus Ratings section, there was a significant mean difference from the target words in number of letters (using absolute values) for the corresponding SP and S lures. This one letter length difference could have made it more likely that SP intrusions would occur due to the closer relation they possessed in number of letters to the target word. Therefore, an analysis of covariance (ANCOVA) was performed (with items as the random factor) to investigate the proportion of SP and S substitutions with the length difference controlled as a covariate. Specifically, the covariate was the absolute value of the length difference from the target word for each lure. With length as a covariate, there remained a significant advantage of (induced) SP over S substitutions, $F(1, 38) = 12.33$, $\text{MSE} = 126.33$, $p < .01$. 
Despite the fact that many potential confounds such as length and semantic relatedness were controlled in Experiment 1, other factors may still exist which influence the relative substitution rates of the different lure types. One such potential confound might be "cloze" rates, or the probability that a word not presented anywhere in the experiment would be inserted into a sentence. Such a possibility will be explored in Experiment 4. Also, no purely phonological substitutions occurred. This finding was obviously counter to predictions and will be examined in greater detail in Experiment 5. As shown in Table 1, a much smaller proportion of purely semantic substitutions was obtained relative to Potter and Lombardi. Although this finding does not invalidate a finding of greater SP than S intrusions, it does suggest a difference in materials that influence the likelihood of S errors. Experiment 2 was performed to explore the source of the discrepancy.

EXPERIMENT 2

The finding in Experiment 1 of few S-lure substitutions was investigated. It was hypothesized that the Experiment 1 sentences were more difficult and less predictable than those of Potter and Lombardi (1990), thus diminishing the role of the conceptual representation (of the sentence) for the verbatim repetition task; such a representation would be more difficult to form for unpredictable sentences. Therefore, as an example, the sentence, "Investigation of the body revealed two small insect bites on the arm" was replaced with "Investigation of the body was ordered to determine the cause of death." It was presumed that the latter sentence in the example would have a more predictable
ending than the former, therefore allowing a stronger overall meaning representation to be formed. Using these easier sentences, it was expected that it would be easier for a lure word to match the conceptual representation of a target word and thus more semantic intrusions would occur.

Method

Participants. 20 Rice University students participated for course credit. None had participated in Experiment 1.

Materials. The materials used were the same as in Experiment 1 with the exception that new sentences were created for the existing targets and lures (Mean number of words per sentence = 12). As stated above, these sentences were thought to be more predictable and less difficult to comprehend than those of Experiment 1.

Design and Procedure. Same as in Experiment 1.

Results and Discussion

Data from one subject was replaced by a new subject due to a greater than 40% error rate on the probe recognition task. The mean percentage correct on the probe task for the retained subjects was 76%. The data revealed a similar pattern to that found in Experiment 1. As shown in Table 1, subjects once again performed the recall task quite well, recalling 85% of the words correctly (excluding targets). The repeated measures ANOVA revealed a main effect of type of lure, $F_1(2, 38) = 49.73, \text{MSE} = 37.37, p < .001, F_2(2, 78) = 16.04, \text{MSE} = 224.90, p < .001$, with the proportion of SP lures once again exceeding the other types, as well as a main effect of the induced/spontaneous factor, $F_1(1, 19) = 10.97, \text{MSE} = 35.65, p < .01, F_2(1, 39) = 14.15, \text{MSE} = 62.29, p < .001$, as shown by a larger proportion of induced substitutions. The interaction between the
induced/spontaneous factor and type of lure was also significant, \( F_1(2, 38) = 4.31, \text{MSE} = 27.72, p < .05, F_2(2, 78) = 4.25, \text{MSE} = 58.06, p < .05 \), with the difference between the induced SP and spontaneous SP intrusions much larger than the induced vs. spontaneous differences for the other two conditions.

As in Experiment 1, a significantly greater proportion of substitutions occurred for the SP lures (17%) compared to the S lures (4%), \( t_1(19) = 4.95, p < .001, t_2(39) = 3.32, p < .01 \), suggesting the influence of a phonological record of the sentence during verbatim recall. Also, the advantage of spontaneous SP (9%) over spontaneous S (1%) substitutions was replicated as well, \( t_1(19) = 6.40, p < .001, t_2(39) = 3.35, p < .01 \).

Further, a higher proportion of substitutions was obtained for SP versus spontaneous SP lures as in Experiment 1, \( t_1(19) = 2.67, p < .05, t_2(39) = 2.87, p < .01 \). The interaction contrast performed for Experiment 1, \((\text{SP} - \text{SponSP}) - (\text{S} - \text{SponS})\), was also performed for the current experiment and yielded a non-significant result, \( t_1(19) = 1.74, p = .10, t_2(39) = 1.42, p > .10 \).

Contrary to the prediction, the rate of S lure intrusions increased only from 2% to 4% of possible trials, far below the 22% found by Potter and Lombardi (1990). Therefore, although Experiment 2 served as a valuable replication of Experiment 1, it was found that the rate of S lure intrusions did not approach that found by Potter and Lombardi (1990). Therefore, before proceeding, a replication of the findings of Potter and Lombardi, Experiment 2, was attempted.

EXPERIMENT 3
A much smaller proportion of semantic intrusions in Experiments 1 and 2 was
found relative to Potter and Lombardi (2% in Experiment 1 and 4% in Experiment 2 vs.
22% in Potter and Lombardi). Before a further attempt at determining the source of this
discrepancy, a replication of this result was attempted using their sentence materials,
which were provided in an Appendix.

Method

Participants. 20 Rice University students participated for course credit. Some had
participated in one or more of the above experiments but were not debriefed until all
testing was completed. None had participated in both experiments 1 and 2.

Materials. The stimuli were taken from Potter and Lombardi (1990), Appendix A.

Design and Procedure. Identical to that of Potter and Lombardi (1990),
Experiment 2 (with the same procedure as in the current Experiments 1 and 2).

Results and Discussion

Before considering the results of the replication, a brief review of the results from
Potter and Lombardi (1990), Experiment 2 is warranted. They found a substitution rate
for semantically related stimuli of 22% (of all possible trials), with 86% of the words
(excluding targets) being recalled correctly. Furthermore, they found a spontaneous S
rate of 9%.

The results from the replication can be found in Table 1. Data from one subject
was replaced due to a greater than 40% error rate on the button pressing task. The mean
percentage correct on the probe task for the retained subjects was 82%. Subjects showed
a similar pattern of performance, recalling 83% of the words correctly (excluding targets),
with a 16% rate of S intrusions, which is not unlike that found in Potter and Lombardi
(1990), but is much greater than that found in Experiments 1 and 2 (2% and 4%, respectively). In addition, a spontaneous $S$ substitution rate of 10% was found which is nearly identical to the 9% found by Potter and Lombardi (1990). A test of the difference (in the current data) between induced and spontaneous semantic substitution rates yielded a significant result, $t_1(19) = 2.22, p < .05$, $t_2(19) = 2.46, p < .05$.

Experiment 3 has clearly replicated the phenomenon of semantic substitutions found by Potter and Lombardi (1990). Although the rate of 16% was not quite as large as the 22% found previously, it was clearly greater than the rates found in Experiments 1 and 2. Another possible source of the difference is simply the naturalness of the lures when substituted in the sentences. Potter and Lombardi had chosen their semantic lures to be more probable words in the sentence than the target. Although the SP and $S$ lures used in the current experiments were also chosen on an intuitive basis to be more probable than the targets, the first two experiments were carried out without explicit measurement of their naturalness in the sentences. Moreover, even if they were more probable than the targets there was no guarantee that they were as probable in their respective sentences as were the Potter and Lombardi lures. To address this issue the current sentences and the Potter and Lombardi sentences were presented in a modified cloze task in which the sentences were presented minus the target. Subjects were asked to fill in the blank with the word that best fit the sentence. In this fashion, both the relative probability of $S$ lures in the current experiments versus Potter and Lombardi, and the relative probabilities of $S$ and SP lures could be assessed.
A cloze comparison was performed to examine what influence (if any) the naturalness of the targets and lures in the sentences played in the rate of lure intrusions from Potter and Lombardi (1990, Experiment 2) as well as the current studies (see Table 2). In other words, an analysis was performed to determine how likely substitution of particular words were with no lure or target words present. Thirty-four subjects were presented with the stimulus sentences from the experiments with the target words deleted and were told to simply fill in the blank with an appropriate word. This cloze task was unlike traditional cloze tasks, however, in that each sentence was presented in its entirety (minus the target word). Most cloze tasks present the sentence only up to the word before the target. However, predictability only to that point was not of interest since subjects in the experiment would be presented with the entire sentence. Instead, the interest was in the likelihood that subjects were to substitute a particular word when they perceived the rest of the sentence; it is certainly possible on some trials that subjects miss the target but perceive the rest.

Surprisingly, an analysis of the original Potter and Lombardi (1990) stimuli revealed that over 48% of the responses (to the sentence blanks) corresponded to the semantically related lures used by the original authors compared to 23% for targets. This 25% advantage for S lures could very well account for the 22% semantic intrusion rate found by Potter and Lombardi, Experiment 2.

In comparison, the cloze rates for the current semantic lures were only 11% for Experiment 1 and 13% for Experiment 2, which differed from their cloze rates for corresponding targets by only 4% and 6%, respectively. Thus, the relative absence of S substitutions in Experiments 1 and 2 could be due to this factor related to naturalness or
Table 2

**Cloze Ratings for Potter & Lombardi (1990) and Current Experiments 1 and 2**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Target</th>
<th>SP</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potter &amp; Lombardi (1990)</td>
<td>23.09</td>
<td>*</td>
<td>48.83</td>
</tr>
<tr>
<td>Exp. 1: Hard Sentences</td>
<td>6.76</td>
<td>34.26</td>
<td>10.71</td>
</tr>
<tr>
<td>Exp. 2: Easy Sentences</td>
<td>6.77</td>
<td>28.09</td>
<td>12.65</td>
</tr>
</tbody>
</table>

predictability in the sentence. Unfortunately, as revealed in Table 2, the cloze data for Experiments 1 and 2 revealed a large and significant advantage for SP lures over S lures in terms of cloze completion rates. Thus, even in the absence of any semantic or phonological information for the target, SP words were more likely completions. Consequently, the higher insertion rate for spontaneous SP than spontaneous S lures could be due to this naturalness factor rather than to phonology. Furthermore, given the failure to find a significant interaction contrast between the induced vs. spontaneous intrusions for SP vs. S lures in either experiment undermines the conclusion that the greater induced SP than S substitutions is due to their phonological properties.
Consequently, Experiment 4 was designed such that cloze rates were equated for the SP and S lures. In addition, the factor of concreteness vs. abstractness of lures and targets was manipulated.

EXPERIMENT 4

In addition to the cloze differences between the S-lures used in Experiments 1 and 2 and those of Potter and Lombardi, it was noted that target words in Experiments 1 and 2 tended to be more abstract than targets used by Potter and Lombardi, and it was hypothesized that this abstractness might have affected the frequency of S intrusions in Experiments 1 and 2. That is, the vague nature of abstract targets might make them less sensitive to conceptual analysis than concrete targets, and, therefore, a much weaker conceptual representation might be formed. Therefore, concreteness ratings were collected for Experiment 1 and 2 stimuli and compared to concreteness ratings from the Oxford Psycholinguistic Database (OPD; Quinlan, 1992) for Potter and Lombardi's (1990) Experiment 2 stimuli. 14 of the Experiment 1 and 2 targets had ratings in the OPD, and the correlation between ratings on the two scales was .899, which allowed confidence in comparing these numbers. Since the new rating scale ranged from 1 to 7 and the OPD ratings ranged from 100 to 700, a regression analysis was performed for these 14 stimulus targets to obtain a means of converting from one scale to the other. The regression yielded the following equation: New Scale = \(-2.07 + (.013)\)Oxford Scale.

Looking at the concreteness ratings of stimuli from Experiments 1 and 2, the mean rating (on a scale of 1 = low concreteness and 7 = high concreteness) was 3.22. Converted
ratings obtained for stimuli used in Potter & Lombardi (1990), Experiment 2 (from the OPD), however, revealed a mean rating of 4.64. Thus, the Potter and Lombardi materials appeared to be somewhat more concrete than those used in Experiments 1 and 2.

Thus, in this experiment target concreteness was specifically manipulated. For example, an abstract target might be “investigation” while a concrete target might be “house.” Such concrete and abstract target words were matched (as in Experiment 1) on variables of frequency, length, semantic relatedness of SP and S lures to targets, as well as a cloze measure of how likely target and lure words were to be inserted into the stimulus sentences when the list words were not presented. As in Experiments 1 and 2, S and SP lures were presented to determine if SP > S for both concrete and abstract targets.

Thus the goal was to determine the influence of concreteness on the rate of semantic intrusions. It was predicted that concrete targets would elicit semantic intrusions at a higher rate than abstract targets, thus contributing to the discrepancy between the results of Experiments 1 and 2 and the results of Potter and Lombardi (1990), Experiment 2. If the predicted results are borne out, this would suggest that abstract stimuli are especially reliant on the phonological record which Potter and Lombardi (1990) have suggested does not even exist.

Method

Participants. 36 Rice University students participated to satisfy a course requirement or for extra credit. None had participated in any prior experiments.

Materials. 60 new sentences were created with 30 being classified as “concrete” and 30 classified as “abstract”; concrete sentences conveyed a concrete message and included a concrete target word (e.g. “The trail mix contained various kinds of nuts and a
dried apricot"), while abstract sentences conveyed an abstract message and included an abstract target word (e.g. "Investigation of the crime by the police took weeks longer than expected"). (The mean number of words per sentence was 12.4). The abstract targets and lures consisted mainly of those used in Experiments 1 and 2 with few exceptions. All concrete targets and lures were created for this experiment only.

Stimulus Ratings. As stated in the introduction to Experiment 4, many factors which could potentially affect the outcome of the experiment were measured and controlled. Ratings of concreteness were collected from a group of 20 subjects who were asked to rate the concreteness of all the target words according to the instructions outlined in Toglia and Battig (1978, p. 3). On a scale from very abstract (1) to very concrete (7), target words selected to be used as abstract ranged from 1.30 to 5.35 (M = 3.13), while words selected to be used as concrete targets ranged from 5.40 to 6.90 (M = 6.47).

Although stimuli were created so that they would be matched on various factors, it was not possible for them to be matched exactly on all. Therefore, as in Experiment 1, comparisons of length, frequency, and semantic relatedness were carried out to rule out alternative explanations of potential findings. Since the only interest was in alternative accounts of findings which are in accordance with predictions, statistical analyses were only conducted for results which may bias the experiment toward those predictions. The mean difference in frequency for the abstract SP (M = 33) and S (M = 50) lures was -17 (that is, with S more frequent), and the mean difference for the concrete SP (M = 28) and S (M = 18) lures was 10, which was not significant, t(29) = 0.95, p = .35. Only the concrete frequency difference was tested statistically due to the fact that a higher proportion of SP versus S substitutions could potentially be caused by a higher mean
frequency for SP lures, but not a higher mean frequency for S lures.

For the concrete stimuli, the mean difference from the target word ($M = 5.73$) in number of letters (using absolute values) for the SP lure ($M = 5.23$) and S lure ($M = 5.07$) (across all stimuli) was -.57 letter (i.e. Mean $|l(T - SP)| - |l(T - S)| = -.57$, $t(29) = -1.61$, $p = .12$). For the abstract stimuli, the mean difference from the target word ($M = 8.27$) in number of letters (using absolute values) for the SP lure ($M = 7.73$) and S lure ($M = 6.33$) (across all stimuli) was -.93 letter (i.e. Mean $|l(T - SP)| - |l(T - S)| = -.93$, $t(29) = -2.05$, $p < .05$). As in Experiment 1, if there is a higher substitution rate for SP than S lures (for concrete or abstract stimuli), an ANCOVA will be performed with the length difference as a covariate to determine its role in the effect. Finally, the comparison of abstract (T-S) versus concrete (T-S) was made to rule out any length difference which might bias the results in favor of more frequent concrete (than abstract) semantic substitutions; it revealed a non-significant result, $t(29) = 0.95$, $p = .35$.

A semantic judgment task was administered to 20 subjects in which they were asked to rate the semantic similarity between the targets and the lures when each word was presented in the stimulus sentence (to be) used in the experiment (presented in pairs for the rating task). There were 4 (levels of relatedness factor) x 5 (possible unrelated words for each sentence) = 20 rating sets. The semantic relatedness to each target word for its corresponding SP and S lures was analyzed (on a scale of 1 to 10) and for the concrete stimuli, the SP lures ($M = 7.3$) were found to be only slightly more related to the targets overall than the S lures ($M = 7.1$) (i.e. Mean[SP - S] = .2), $t(29) = .66$, $p = .52$, n.s.). For the abstract stimuli, the SP lures ($M = 8.0$) were found to be slightly less related to the targets overall than the S lures ($M = 8.1$) (i.e. Mean[SP - S] = -.1).
In addition, a cloze measure was obtained; 20 naive subjects were administered the stimulus sentences to be used in Experiment 4 with blanks present in place of target words. They were told to simply fill in the blank with an appropriate word. As before, this cloze task was unlike traditional cloze tasks in that each sentence was presented in its entirety (minus the target word). To ensure that the cloze rates would be properly controlled, subjects were presented with more sentences than would be necessary for the experiment. The percentage of responses which matched the target word or lures used for each sentence was then analyzed.

Sentences (and their corresponding targets and lures) were then selected so that important potential confounds due to the cloze measure would be eliminated. Cloze comparisons with the final selected stimuli revealed that the percentage of responses which matched target words were equalized for concrete versus abstract sentences (See Table 3). Further, the percentages corresponding to the SP and S lures were greater for the abstract than concrete sentences. This is important because any obtained advantage in rate of lure intrusion for SP or (particularly) S lures for concrete over abstract stimuli could not be explained in terms of the naturalness of the lure word in the sentence. Finally, as can also be seen in Table 3, for both concrete and abstract stimuli, the naturalness of the SP lures in the sentences was either equal to or less than that of the S lures (to rule out the naturalness explanation of the advantage of primary interest, namely SP over S lures).

Therefore, the purpose of Experiment 4 was two-fold. One goal was to replicate the lure intrusion advantage for SP over S lures found in Experiments 1 and 2 (with cloze rates controlled), and a second goal was to determine the role of concreteness in the
Table 3

**Cloze Ratings (%) for Stimuli used in Experiment 4**

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Sentences</th>
<th>Concrete</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td></td>
<td>9.67</td>
<td>9.67</td>
</tr>
<tr>
<td>SP lure</td>
<td></td>
<td>8.33</td>
<td>11.17</td>
</tr>
<tr>
<td>S lure</td>
<td></td>
<td>8.33</td>
<td>12.00</td>
</tr>
</tbody>
</table>

... (some text not shown)...

Design and Procedure. The design is similar to that of Experiments 1 and 2 with the exception that a new factor (i.e. concreteness) was manipulated. There were three randomized counterbalanced sets with each containing all 30 concrete and 30 abstract sentences. Both factors of concreteness (concrete vs. abstract) and lure type (SP, S, U) were varied within subjects (resulting in 6 stimulus conditions). As in prior experiments, due to the imbalance between related (N = 40) and unrelated (N = 20) stimuli, 20 filler sentences were created. The procedure was identical to that of Experiments 1 and 2.

Results and Discussion

Data from eight subjects was replaced due to a greater than 40% error rate on the probe recognition task. The mean percentage correct on the probe task for the retained subjects was 79%. Overall, subjects performed the repetition task very well in this experiment, recalling 90% of non-target words correctly. Looking at the means for the
abstract stimuli first (See Table 4), there were semantically and phonologically related (SP) intrusions on 14% of possible trials, and purely semantically related (S) intrusions on 3% of possible trials. For the concrete stimuli, there were 4% SP intrusions and 1% S intrusions.

With subjects as the random factor, a 2 X 2 X 2 repeated measures ANOVA was performed with Type of Substitution (Induced/Spontaneous), Concreteness (Concrete/Abstract), and Type of Lure (SP/S) as the within-subjects factors. For the analysis with items as the random factor, a mixed ANOVA was performed with Concreteness (Concrete/Abstract) as the between-subjects factor and Type of Substitution (Induced/Spontaneous) and Type of Lure (SP/S) as the within-subjects factors.

The analyses revealed a main effect of Concreteness, $F_1(1, 35) = 46.77, \text{MSE} = 22.05, p < .001$, $F_2(1, 58) = 10.60, \text{MSE} = 81.06, p < .01$, with more intrusions on abstract than concrete sentences, as well as a main effect of lure, $F_1(1, 35) = 44.39, \text{MSE} = 25.86, p < .001$, $F_2(1, 58) = 14.74, \text{MSE} = 64.89, p < .001$, with more intrusions by the SP than S lures. The main effect of Type of Substitution was significant as well, $F_1(1, 35) = 43.36, \text{MSE} = 25.56, p < .001$, $F_2(1, 58) = 25.81, \text{MSE} = 35.78, p < .001$, with more intrusions when the lure was present than when it was not. The interaction of Concreteness and Lure was also significant, $F_1(1, 35) = 15.81, \text{MSE} = 26.14, p < .001$, $F_2(1, 58) = 5.31, \text{MSE} = 64.89, p < .05$, with the difference between SP and S lures much larger for abstract than concrete sentences. The interaction between Type of Substitution and Concreteness was also significant, $F_1(1, 35) = 21.39, \text{MSE} = 25.32, p < .001$, $F_2(1, 58) = 12.62, \text{MSE} = 35.78, p < .001$, with a much larger difference between induced and spontaneous substitutions for abstract than concrete sentences. Also, a
Table 4

Results of Experiment 4

<table>
<thead>
<tr>
<th>Stimulus Type</th>
<th>Intrusion</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Induced</td>
<td>SP</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td>14</td>
<td>3</td>
<td>8.5</td>
</tr>
<tr>
<td>Concrete</td>
<td>4</td>
<td>1</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>M</td>
<td>9</td>
<td>2</td>
<td>5.5</td>
</tr>
</tbody>
</table>

|               | Spontaneous | SP| S |
|               |             | M |   |
| Abstract      | 3          | 1 | 2.0|
| Concrete      | 1          | 1 | 1.0|
|               | 1          | 1 | 1.5|
| M             | 2          | 1 | 1.5|

significant Type of Substitution by Lure interaction was found, $F_{1}(1, 35) = 23.22$, $MSE = 24.53$, $p < .001$, $F_{1}(1, 58) = 17.60$, $MSE = 26.96$, $p < .001$, with the lure effect (SP > S) much greater for induced than spontaneous conditions. Finally, the three-way interaction between Type of Substitution, Concreteness, and Lure was significant, $F_{1}(1, 35) = 5.99$, $MSE = 22.03$, $p < .05$, $F_{1}(1, 58) = 4.08$, $MSE = 26.96$, $p < .05$, with the Lure x Type of Substitution interaction much more pronounced for abstract stimuli, due to the induced abstract SP intrusion rate of 14% which was much greater than all other conditions.

The spontaneous intrusion rates were quite low. For abstract stimuli, there were 3% SP intrusions and 1% S intrusions. For concrete stimuli, there were 1% intrusions for both. Planned contrasts were carried out to investigate the difference between SP and SponSP from the difference between S and SponS for both sentence types (with cloze controlled). For both stimulus types, this interaction contrast was significant, $t_{1}(35) =$
2.85, \( p < .01 \), \( t_2(29) = 3.17 \), \( p < .01 \) for concrete stimuli, \( t_1(35) = 4.10 \), \( p < .001 \), \( t_2(29) = 3.31 \), \( p < .01 \) for abstract stimuli. Next, the comparison of concrete semantic versus abstract semantic substitutions relative to spontaneous was conducted, \([Concrete (S - SponS)] - [Abstract (S - SponS)]\), and failed to yield the predicted result. In fact, the opposite finding was obtained, namely a higher rate of abstract than concrete semantic intrusions, \( t_1(35) = -2.53 \), \( p < .05 \), \( t_2(29) = -2.16 \), \( p < .05 \).

As a whole, the results supported the main hypothesis of a role for phonology, namely a higher intrusion rate for SP compared to S lures for both concrete and abstract stimuli. Such a result was obtained even when spontaneous intrusion rates were used as baseline, a result which failed to reach significance in Experiments 1 and 2. Further, recall that for the abstract stimuli, a significant length difference existed which was similar to that found for the stimuli in Experiment 1, namely a greater difference for the S lures. Specifically, the mean difference from the target words in number of letters (using absolute values) for the corresponding SP and S lures was -.93. Therefore, the effects of concreteness and lure type were investigated with the length difference for each lure from its target as a covariate. The ANCOVA (with items as the random factor) revealed a main effect of concreteness, \( F(1, 58) = 13.76 \), \( MSE = 105.88 \), \( p < .001 \). Also, a main effect of lure type was found, \( F(1, 57) = 10.64 \), \( MSE = 73.09 \), \( p < .01 \). Finally, with length as a covariate, the interaction between concreteness and lure type remained, \( F(1, 57) = 4.72 \), \( MSE = 73.09 \), \( p < .05 \).

The high semantic (S) intrusion rates found by Potter and Lombardi were not replicated in this experiment. In fact, when the spontaneous semantic intrusion rates were used as baseline, the rate of abstract semantic intrusions was 2% and the rate of concrete
semantic intrusions was 0%. Furthermore, the predicted advantage for concrete over abstract semantic lures was not obtained, but rather the opposite result was found. It was concluded that the abstractness of the Experiment 1 and 2 stimuli relative to Potter and Lombardi was not responsible for the low rate of semantic intrusions found in Experiments 1 and 2. Rather, the high rate of S intrusions found in Potter and Lombardi (1990), Experiment 2, is most likely due to the extraordinarily high cloze rates for their stimuli, which was 48% for their semantic lures and only 8% and 12% for the current concrete and abstract semantic lures, respectively. It must be emphasized, however, that even under conditions in which cloze rates were equalized for SP and S lures, a clear advantage of the former type was obtained in Experiment 4.

Further, it is noteworthy that the rate of SP intrusions found for abstract stimuli in Experiment 4 (14%) equaled that of Experiment 1 (14%). This supports the notion expressed earlier that if the semantic representation of the sentence is weak (as can be argued for abstract stimuli), there is a greater reliance on phonological information to perform sentence repetition, with the result that there is a high rate of SP intrusions relative to S intrusions. However, when the semantic representation of the sentence is stronger, the gap between SP and S intrusion rates is diminished due to less of a reliance on the phonological representation of the sentence.

In addition, the small but significant advantage for abstract semantic lures over concrete semantic lures was counter to prediction and deserves some speculation. The higher semantic substitution rate for abstract stimuli may be due to the higher cloze rate for the S lure than the target for abstract stimuli (as compared with concrete stimuli; see Table 3). However, the difference in cloze rates was small, and an Analysis of
Covariance performed with lure cloze rates as a covariate replicated the significant effects of concreteness, \( F(1, 58) = 10.41, \text{MSE} = 102.5, p < .01 \), type of lure, \( F(1, 57) = 19.22, \text{MSE} = 73.91, p < .001 \), and their interaction, \( F(1, 57) = 5.94, \text{MSE} = 73.91, p < .05 \), found earlier in the overall ANOVA. Furthermore, an analysis of the difference in semantic intrusion rates between abstract and concrete stimuli using (S lure) cloze rates as a covariate produced a marginally significant effect, \( F(1, 57) = 3.95, \text{MSE} = 28.1, p = .0516 \). A similar analysis performed for the SP intrusion rate difference between abstract and concrete stimuli elicited a highly significant result, \( F(1, 57) = 9.49, \text{MSE} = 143.7, p < .01 \).

At the outset of Experiment 4 it was hypothesized that the stronger semantic representation of concrete sentences (as opposed to abstract ones) would lead to more semantic intrusions, not less as was found. The greater S and SP intrusions in abstract than concrete sentences may suggest that a strong semantic representation acts to some extent as a barrier to errors of a semantic nature. One possible basis for such an argument, which will be elaborated upon in the General Discussion, is that a strong semantic representation leads to greater activation of correct lexical items (and a reduced error rate).

**Supplementary Analysis of Serial Position Effects**

Due to the many pieces of evidence from past research suggesting the possible existence of a verbatim trace (of some kind) lasting for relatively long periods of time, the possibility was entertained that SP substitutions might occur throughout all positions of
the 10 to 15 word sentences. As an example of this previous research, Anderson and Paulson (1977) found striking evidence that a verbatim representation of a sentence persists for a length of time far beyond the accepted limits of conventional STM theories. They viewed the verbatim representation as a representation of word order, but stopped short of implicating a phonological representation. In their experiment they presented subjects with pairs of sentences (study and test) and instructed them to respond as to whether the sentence logically followed from previous sentences. Anderson and Paulson were interested in any advantage provided by test sentences which were exact matches of previous study sentences versus sentences with altered voice (active/passive). Their results revealed a response time advantage of approximately 300 msec even when the test sentence had been delayed by presentation of fifteen intervening test and study pairs. In terms of time delay, there was a period of 450 seconds between the end of study and beginning of test. Therefore, Anderson and Paulson's results clearly provide support for a verbatim representation (of some kind) which lasts for 7.5 minutes.\textsuperscript{6,7}

Compatible results were also found by Garrod and Trabasso (1973). They presented subjects with four-sentence paragraphs followed by a question about the paragraph. They found that subjects were faster to judge the truth (or falsity) of the question if it was presented in the same voice as the pertinent sentence in the paragraph (a result found previously). The interesting result relevant to the present discussion was that such an advantage was found even for the first sentence of the paragraph suggesting that at least some verbatim information was retained for a fairly long time period (but see Perfetti & Goldman, 1974, for the influence of thematization of sentence elements on sentence recall).\textsuperscript{8}
Therefore, some exploratory analyses were carried out to investigate whether the influence of phonology would differ across serial position of the target. Specifically, the goal was to informally determine whether a higher SP (vs. S) intrusion rate would be found throughout all positions of the long sentences (for both concrete and abstract targets). It is important to point out, though, that targets were not counterbalanced or matched in any fashion across serial position, so these results should be viewed with caution. Further, locations were divided into groups because there were relatively few opportunities for lure substitutions at each serial position. Tables 5 and 6 present the rates of (induced and spontaneous) SP and S intrusions for five groups of target locations for abstract and concrete targets, respectively (using data with items as the random factor). Also indicated in the Tables are the number of observations per location group as well as the locations encompassed by each group (with 1 referring to the target word as the last word of the sentence; 2, the next to last word, and so on). Using the Bonferroni procedure to control for Type-I errors, the (induced) SP vs. S difference for each group was analyzed for both concrete and abstract targets. Using an alpha criterion of \(0.05/5 = 0.01\) for significance, no significant effects were found. An additional goal was to investigate if the difference between the rate of SP and S (induced) substitutions varied depending on the location of the target word in the sentence, and a non-significant Lure x Group interaction was found for abstract stimuli, \(F(4, 25) = .74, \text{MSE} = 146.4, p > .5\). However, the Lure effect did interact with Group for the concrete stimuli, \(F(4, 25) = 3.37, \text{MSE} = 15.31, p < .05\), as can be seen in Table 6 by the greater proportion of SP substitutions for positions near the beginning of the sentence. Similar analyses undertaken to investigate the Lure x Group interaction for the spontaneous intrusions
Table 5

Substitution Rates (%) by Condition as a Function of Location Group for ABSTRACT

**Targets**

<table>
<thead>
<tr>
<th>Group</th>
<th>SP</th>
<th>Condition S</th>
<th>SponSP</th>
<th>SponS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(N = 2) (Locations: 1-3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17.86</td>
<td>5.95</td>
<td>4.17</td>
<td>0.60</td>
</tr>
<tr>
<td>(N = 7) (Locations: 4-5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>22.22</td>
<td>2.78</td>
<td>5.56</td>
<td>2.08</td>
</tr>
<tr>
<td>(N = 6) (Locations: 6-7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8.33</td>
<td>1.85</td>
<td>1.85</td>
<td>0.93</td>
</tr>
<tr>
<td>(N = 9) (Locations: 8-10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13.89</td>
<td>4.17</td>
<td>2.08</td>
<td>0.00</td>
</tr>
<tr>
<td>(N = 6) (Locations: 11-13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(SponSP - SponS) yielded non-significant results, F(4, 25) = .4, MSE = 14.8, p > .8 for abstract stimuli, F(4, 25) = 2.14, MSE = 7.49, p > .1 for concrete stimuli. However, these analyses are all plagued by small (and/or unequal) sample sizes and no firm conclusions can be drawn. Nevertheless, eyeballing of the data reveals that for virtually all location groups for both concrete and abstract targets, there was a higher rate of intrusions for SP than S lures.
Table 6

Substitution Rates (%) by Condition as a Function of Location Group for CONCRETE Targets

<table>
<thead>
<tr>
<th>Group</th>
<th>SP</th>
<th>S</th>
<th>SponSP</th>
<th>SponS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (N = 9) (Locations: 1-3)</td>
<td>0.93</td>
<td>1.85</td>
<td>0.00</td>
<td>1.39</td>
</tr>
<tr>
<td>2 (N = 9) (Locations: 4-5)</td>
<td>1.85</td>
<td>0.00</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>3 (N = 6) (Locations: 6-7)</td>
<td>5.55</td>
<td>0.00</td>
<td>0.00</td>
<td>0.70</td>
</tr>
<tr>
<td>4 (N = 4) (Locations: 8-10)</td>
<td>10.42</td>
<td>0.00</td>
<td>6.25</td>
<td>1.04</td>
</tr>
<tr>
<td>5 (N = 2) (Locations: 11-13)</td>
<td>4.17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The data corresponding to the serial positions most associated with transient phonological storage were then inspected. Looking at data for target items in the last three positions of their respective sentences (locations 1-3), there were only small or non-existent effects for those positions for concrete and abstract targets, respectively. However, with so few data points contributing to the mean substitution rates for abstract
targets at late positions (Group 1, N = 2), more data would be needed to draw any meaningful conclusions. For concrete stimuli, it is unclear why induced SP substitution rates were greater for earlier sentence locations (in general); a better controlled study of serial position effects within this paradigm would provide some answers. One possibility is that, in accordance with arguments made regarding the higher rate of abstract semantic intrusions, the phonological-semantic record of final words is strongest relative to earlier sentence words or the recognition probe words presented before the sentence.

EXPERIMENT 5

Experiment 4 has provided strong evidence that a short-term phonological representation is retained and used during verbatim sentence repetition. In accordance with this hypothesis, it was predicted that it might be possible to demonstrate purely phonological lure intrusions in which the lure word bore no semantic relation to the target word. Although far fewer substitutions of the purely phonological kind were predicted, any such intrusions would provide powerful corroboration for the advocated view.

Therefore, in Experiment 5, only phonologically related lures which bore no semantic relation to their phonologically-related counterparts in the sentences were used (e.g. simulation/stimulation). The goal was to determine whether intrusions of P-lures would occur even though they would make the sentence nonsensical. To encourage such intrusions, the criterion for phonological relation was made much more stringent than in Experiments 1 and 2. In contrast to these experiments, in which a phonological relation was defined by having one or two initial phonemes in common with the target word, P-
lures in the current experiment were differentiated from their targets by at most only a few phonemes with most items being differentiated by one phoneme (e.g., simulation/stimulation). Further, targets and lures were matched in length (in number of syllables) and syllabic stress location as well (see Garrett, 1993). In addition, cloze rates for this experiment are likely to be zero, so they were not controlled.

Method

Participants. 30 Rice University students participated for course credit. Some had participated in either Experiment 1 or 2, but were not debriefed until they completed both experiments.

Materials. Twenty new sentences were created for Experiment 5 (Mean number of words per sentence = 11.9). The new lure words were related phonologically to the target words as described above (e.g. simulation/stimulation). That is, on average shared 87% of phonemes, and were also matched in number of syllables and stress pattern.

Design and Procedure. The design was similar to that of Experiment 1 with the exception that only phonologically related lures were used. There were ten sentences per condition (P or unrelated) counterbalanced into two sets with different random orders. The procedure was identical to that of Experiment 1. Further, to counter the potential criticism that a substitution might occur because subjects are unfamiliar with a target word (which were sometimes difficult; e.g., “corroboration”), a four-choice multiple choice synonym judgment task was administered to the subjects following the test session in which they were asked to circle the closest synonym provided for each target word. Any target word for which the response was incorrect was not used in the analysis of the data (i.e., any substitution to that target word was not counted).
Results and Discussion

Data from nine subjects was replaced due to a 40% or greater error rate on the button pressing task. The mean percentage correct on the probe task for the retained subjects was 77%. As shown in Table 1, subjects performed the recall task fairly well, recalling 81% of the words correctly (excluding targets). As predicted, purely phonological substitutions occurred on 7% of all possible trials. Additionally, spontaneous phonological substitutions occurred on 6% of all possible trials. However, 3% of the spontaneous P substitutions were lure words seen by other subjects in the other stimulus set, consistent with the "spontaneous" classification of Experiments 1 and 2, (termed "PES" substitutions: Phonological--Experimental Spontaneous) and 4% were intrusions of words not used as a lure anywhere in the experiment (termed "POS" substitutions: Phonological--Other Spontaneous). This distinction was made when the experimenter noted that subjects would often substitute phonologically related words which were not classified as lures. It was felt that such intrusions (of the POS type) were clearly relevant to the hypothesis and they were recorded. An intruded word was classified as an instance of a POS substitution when it contained at least three consecutive phonemes in common with the target (e.g., interjection -> injection); however, the vast majority of such substitutions contained nearly all phonemes in common with the target as well as its syllabic structure and stress (e.g., defection -> deflection; projection -> prediction).

The null hypothesis that there were equal rates of phonological substitutions for induced and spontaneous conditions was tested. A comparison of P vs. PES (i.e., intrusion was a lure on other trials) revealed more substitutions for the former condition,
t(29) = 2.84, p < .01, t(19) = 2.35, p < .05, suggesting that the presence of the lure word on the list increased the likelihood of a phonological substitution. The 6% aggregate substitution rate of the PES and POS types also clearly supports the hypothesis. The words that were substituted were not presented at any time to the subjects providing powerful evidence that a surface phonological representation of the sentence is retained and used during verbatim recall. However, the possibility (for this and all other experiments) that subjects simply misperceived the (target) words during presentation was acknowledged, and future research will be necessary to rule out such an explanation. Nevertheless, the lack of phonological substitutions in Experiment 1 can thus be accounted for by the more lax criterion for phonological similarity used there. Apparently, a very close phonological relation is necessary for semantically incoherent substitutions to occur, and the rate of such intrusions reached only 7%.

As in the preceding experiment, an analysis of serial position effects on the rate of P substitutions was carried out. Although the presence of intrusions for targets at all positions was not investigated formally, informal observation revealed that phonological intrusions occurred for targets presented both early and late in the long sentences. Specifically, target word locations were divided into four groups based on distance from the end of the sentence. Once again, the grouping was done because there were relatively few opportunities for lure substitutions at each serial position, and it is important to note that target words were not counterbalanced across positions. Table 7 displays the number of observations (target words) per group as well as the location of the target words in terms of number of words from the end of the sentence (e.g., 1 refers to the target being the last word of the sentence, 2 the next to last word, etc.). As shown in Table 7 (using
Table 7

**Mean Phonological Substitution Rates (%) as a Function of Location Group**

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.78</td>
</tr>
<tr>
<td>(N = 6)</td>
<td></td>
</tr>
<tr>
<td>(Locations: 1-6)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>9.33</td>
</tr>
<tr>
<td>(N = 5)</td>
<td></td>
</tr>
<tr>
<td>(Locations: 7-9)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.67</td>
</tr>
<tr>
<td>(N = 5)</td>
<td></td>
</tr>
<tr>
<td>(Location: 10)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.33</td>
</tr>
<tr>
<td>(N = 4)</td>
<td></td>
</tr>
<tr>
<td>(Locations: 11-12)</td>
<td></td>
</tr>
</tbody>
</table>

data with items as the random factor), phonological substitutions occurred for targets in all four groups of locations. Further, Tukey tests demonstrated no significant differences between any pair of means. Unfortunately, there were only two target words which were presented in the last three positions of their sentences, so only speculation about the serial positions most associated with transient phonological storage can be offered. Both targets appeared as the last word of their respective sentences, with one showing 20% P intrusions and the other 0%. Such a result, in conjunction with the other location group
means, tentatively provides support for the notion that a phonological memory
representation is maintained for all positions of long sentences. Overall, these results
suggest that a short-term phonological representation such as Baddeley’s (1986) is
inadequate to explain this result. This issue will be addressed in the General Discussion.
GENERAL DISCUSSION

The Role of a Phonological Representation in Verbatim Recall

The results of the experiments in the current study support the notion that a phonological representation is indeed used to perform sentence repetition: Experiments 1, 2, and 4 demonstrated that semantically and phonologically related (SP) lures substituted into the sentences at a higher rate than purely semantically related (S) lures, while Experiment 5 revealed that purely phonological substitutions could occur. Although cloze probabilities may weaken conclusions for Experiments 1 and 2, Experiments 4 and 5 provide strong support for the role of phonology.

However, according to Potter and Lombardi (1990), the main problem with advocating a surface representation used in verbatim recall was the nature of such a representation and how it could be used to recall long sentences:

"If the surface representation is, for example, an ordered string of word-sounds, it has a capacity of only about three to six words, judging by memory-span studies. Even if the surface representation is supplemented by a central executive system with some memory capacity (as in Baddeley’s multiple-components model of working memory, reviewed in Baddeley, 1986), there is no account of how the representation of the surface string is combined with word or phrasal chunks that are represented in a lexical, syntactic, or conceptual system or how these representations are mapped smoothly onto a single output during recall" (p. 634).

Potter and Lombardi (1990) undoubtedly believed in the existence of an “auditory-like
short-term representation” (p. 647) which is used for recall of unrelated word strings. However, they argued that the capacity of such a representation is only three to six words and can thus play only a supporting role in recall of long sentences. In addition, they argued that recall of a (random) word list is accomplished via retrieval of activated lexical entries by the acoustic trace (of the word list). Therefore, in their view, recall of sentences as well as random word lists is based on retrieval of activated lexical entries. The only difference is that a conceptual representation is used for lexical retrieval in sentence recall, while an ordered acoustic representation is used for lexical retrieval in word list recall. According to their account, conceptual representations in sentence recall activate all compatible lexical entries with the recently activated entries being selected due to greater activation levels relative to compatible but not recently activated entries.¹⁰ Therefore, as will be discussed in greater detail below, Potter and colleagues advocate a phonological short-term representation, albeit of a very small capacity.

Baddeley (1990) argues that memory span represents the number of items of whatever length that can be uttered in about two seconds. He further argues that the time it takes for the phonological STM representation to decay (from the phonological store in his model; i.e. its duration) is approximately two seconds, akin to the capacity of phonological STM postulated above. With regard to the results, even at a very fast reading speed of 2.5 words/second, it would take 4-6 seconds of storage capacity to completely store the stimulus sentences in the current experiments (10-15 words); a duration far above the limit of two seconds set by Baddeley (1990). (Although words are presented visually at 5 words/second in these experiments, phonological (i.e. inner speech) coding of the words would presumably lag behind visual processing, with a rate
closer to 2.5 words/second.) However, the results of the current experiments revealed the
presence of phonological substitutions for target words both early and late in the long
sentences. Recall the finding from Experiment 5 that phonological substitutions occurred
for all target location groups (Table 7), and the finding from Experiment 4 that higher
rates of SP than S intrusions occurred for virtually all target location groups for both
abstract (Table 5) and concrete (Table 6) targets. It was argued, as a result, that the
phonological storage capacity is increased and supported by other levels of STM, such as
the semantic and lexical levels (to be discussed in greater detail below).

In terms of the theoretical view that has been advocated, the evidence provided by
Anderson and Paulson (1977) and others strongly suggests that a “longer-term”
phonological trace is possible. (For evidence from standard memory research with recall
tasks, see Penney, 1979; Watkins & Todres, 1980; Balota, Cowan, & Engle, 1990;
Watkins & Watkins, 1980.) However, it would be acknowledged that although durable,
the trace would be substantially weakened after a long time period. Indeed, it would be
unlikely that a phonological representation of a perceived sentence lasts for the length of
time suggested by the results of Anderson and Paulson (1977), namely 7.5 minutes.
However, it is believed that the many different types of representation which fall under
the heading of “verbatim trace” (e.g. phonological, word-order, orthographic, etc.) as a
whole are indeed potentially that durable. In other words, if all such sources of
information are available to aid recall of a stimulus, then interactive support among them
will greatly increase the length of time during which recall of the stimulus will be
possible..

One possible explanation for the wide variance across experiments in the
literature of the durability of the verbatim representation could be the common contrast between recall versus recognition tasks. Although the current study is one of recall, it seems that most evidence of longer-term verbatim representations come from studies involving recognition (e.g., Long, 1994; Conway & Gathercole, 1987; Kintsch, McKoon, & Keenan, 1974; Bates, Kintsch, Fletcher, & Guiliani, 1980) while evidence of a less durable verbatim trace comes primarily from recall studies (e.g., Jarveila, 1971; Conway & Gathercole, 1987; also consider the vast literature on recency and modality effects in recall tasks). This phenomenon is interpreted in terms of activation levels. Presumably, a verbatim trace (outlined above) is retained in memory for (relatively) long periods of time. However, this trace decays in STM as does any short-term memory representation. In line with the current view, for recall of random word lists, surviving phonological information for early list items may not be sufficiently activated to unambiguously indicate a particular word. However, on a recognition task, there may be enough phonological information to affect responses, for example, causing longer times to reject a non-presented item that is phonologically similar to an early list item (e.g. Raser, 1972; Shulman, 1970). One could put forth an analogous argument for sentence stimuli, in that phonological information pertaining to sentence words tends to decay over time, yet is still available to be tapped by an appropriate task. Therefore, short verbatim memory spans found in recall tasks are not interpreted as being a true barometer for the existence or durability of a phonological trace, but rather serve as an indication of its strength.

Consistent with the discrepant findings found for recall versus recognition tasks, Conway and Gathercole (1987) provided evidence which could be interpreted as suggesting the existence of a longer-term phonological STM that can only be tapped with
a recognition task. In their study, subjects were presented with thirty list words and were instructed either to say aloud, mouth silently, or read silently each word; an incidental (i.e. unexpected) recognition test was later administered via a questionnaire. The results indicated, in several experiments, better recognition of spoken words than words which were read silently, and of greatest interest, this advantage was found over early and middle (as well as late) positions of the input list, suggesting that a phonological code was maintained even for early list words. Presumably, the code is not merely acoustic in nature for it would be unlikely to last over a list of thirty words. A later experiment also found similar effects when subjects simply listened passively to words instead of speaking them aloud. In a free recall experiment (Experiment 3), however, the advantage of spoken over read words was confined to the late list positions (the final third); a result more consistent with traditional short-term memory findings. Therefore, these findings as a whole support the notion that there is a phonological short-term representation which can persist over a list of as many as thirty words, yet is able to be tapped only by a recognition procedure (see also Bates, Kintsch, Fletcher, & Giuliani, 1980; Kintsch & Bates, 1977; Keenan, MacWhinney, & Mayhew, 1977).

The Multiple Capacity View of Short-Term Memory (STM)

Randi Martin and colleagues (e.g., Martin, 1993; Martin & Romani, 1994a, b; Martin, Shelton, & Yaffee, 1994) have argued that short-term memory should not be seen as a separate system that is drawn upon during language processing. Rather, they argued that verbal short-term memory should be seen as deriving from the procedures and
retentive capacities of language processing modules. In other words, Martin advocates the view that the phonological and semantic retention capacities involved in span tasks depend on language processing resources (Howard & Franklin, 1990; Barnard, 1985; Monsell, 1984, 1987). She argues for a multiple capacity view of short-term memory which contains, in addition to a capacity used to retain phonological information, capacities for retention of semantic and syntactic information (see also Monsell, 1984; Barnard, 1985; Wingfield, Tun, & Rosen, 1995). Indeed, although memory span has traditionally been thought to reflect solely phonological short-term memory, there is much evidence that suggests it is reflective of semantic and lexical information as well (e.g. Hulme, Maughan, & Brown, 1991; Martin & Romani, 1994a, b).

Due to the fact that Martin has focused primarily on comprehension processes, her multiple capacity view of STM will be elaborated in this context. She advocated the view that other levels of representation (such as syntactic and lexical-semantic) are linked (via spreading activation) to phonological information (and thus support verbatim recall). In fact, a similar view was expressed by Butterworth, Shallice, and Watson (1990) who argued that sentence recall was better than list recall due to the supportive influences of semantic and syntactic information (to the phonological code); however, unlike the view of Martin and others, they argued that the semantic and syntactic support provided by sentences originated from long-term (rather than short-term) memory.

Returning to the account of comprehension offered by Martin’s model, as each word of a sentence is heard, the phonological form is derived and the semantic and syntactic features of the words are accessed. On a word-by-word basis, the syntactic structure is developed, and propositions are then derived when information is available
for linking word meanings together. The model's relation to STM rests on the principle that each level of processing during sentence comprehension (and presumably production) possesses its own memory store.

Martin argued that her model could account for performance patterns produced by both normal subjects and brain-damaged patients. For memory of word lists, the first two levels (phonological and lexical-semantic/syntactic features) are used, while for sentence comprehension, the propositional level is most important. To perform sentence repetition, all levels are potentially involved in reconstruction of the input. Martin argued that, for normal subjects, the levels are differentially subject to decay or interference such that the phonological level is lost rapidly without the benefit of rehearsal, while the propositional level is most resistant to decay or interference processes. To account for neuropsychological performance of patients (such as EA, AB, and MW studied by Martin and colleagues; e.g. Martin, 1993; Martin & Romani, 1994a, b; Martin, Shelton, & Yaffee, 1994), it is postulated that brain damage may affect the rate of decay at one level and not others. For patient EA, who was thought to have a phonological STM impairment, the phonological level fades more rapidly than normal (leading to poor verbatim repetition). However, she is still able to rapidly access lexical, semantic, and syntactic information and construct the other levels of representation as the input is perceived resulting in good comprehension (Martin, 1993).

Results from the current set of experiments provided support for the interactivity notion. Using spontaneous substitution rates as baseline, Experiment 4 (the most highly controlled experiment) revealed, for abstract stimuli, a semantic (S) intrusion rate of 2% and an SP intrusion rate of 11%; recall that the rate of purely phonological substitutions
(using the PES spontaneous rate as baseline) from Experiment 5 was roughly 4%. The fact that the sum of the rates of pure S and pure P substitutions does not reach the level of SP intrusions suggests that semantic and phonological codes interact and support each other during sentence repetition. Apart from the current experiments, however, there is much other evidence which seems to support the notion of interactivity of semantic and phonological memory representations.

Turning now to other studies, Patterson, Graham, and Hodges (1994) have provided evidence which is consistent with the advocated position that the distinct STM codes support and strengthen each other. They administered a serial recall task to a patient PP who was diagnosed as having semantic dementia. PP was given lists of words three or four words long, which also varied on a familiarity dimension, and was instructed to immediately recall each sequence. It was found that sequences of words which were familiar (to PP) were recalled markedly better than sequences of unfamiliar words. Notably, PP made phonological errors when recalling the unfamiliar sequences which were characterized as migrations of phonological segments among words within the sequences (e.g. mint, rug → rint, mug). As a result, Patterson et al. argued that long-term semantic knowledge of words acts as a glue which binds the elements of phonological representations together during the constructional process of word production (or in this case, recall of word lists).

From the current perspective, the data from Patterson et al. can be taken as support for a multiple capacity view of STM, in which spreading activation from different short-term memory representations acts to support and strengthen other representations. If Patterson et al.’s patient PP had impaired semantic long-term memory for unfamiliar
words, then it is valid to conclude that her short-term semantic representations (i.e. the representation in semantic STM for the word sequence) for the same words would be impoverished. In fact, a reasonable mechanism by which the long-term semantic information can support phonological representations is transfer of the relevant semantic knowledge to semantic STM so that it can more directly assist the phonological construction process (which is presumably executed in phonological STM). According to the advocated view of STM, possession of a weak semantic STM representation will put more responsibility on the shoulders of the other memory representations (e.g. phonological) resulting in errors, similar to the errors produced for repetition of nonword sequences. The errors found by Patterson et al. for patient PP were phonological in nature, involving (for example) migration of phonological segments throughout the word sequence, and it is argued that such a result would be expected if too much strain were placed on the phonological STM representation due to insufficient support from the semantic STM representation.

Butterworth, Campbell, and Howard (1986) presented a patient RE with the task from Wingfield and Butterworth (1984) in which tape-recorded text is presented to the subject; the subject stops the tape at self-determined intervals and recalls the immediately preceding text. Patient RE had a serial recall span of 4 digits and made errors recalling single sentences verbatim. Notably, she made many errors to sentence words in the terminal positions (no recency effect) suggesting a poor phonological STM. However, when presented with the Wingfield and Butterworth task, she recalled segments as well as controls (even up to 15 words!). In addition, patient RE was asked to repeat sentences from the Token Test (DeRenzi & Vignolo, 1962). It is important to point out that these
sentences are “lexically repetitive and semantically arbitrary” (Butterworth et al., 1986, p. 204); for example, “Touch the large white circle and the small green square.” Patient RE did poorly on this task despite the fact that she did as well as controls on the Wingfield and Butterworth (1984) task. Clearly, this is striking evidence that semantic (and pragmatic, syntactic, etc.) information strongly supports other (presumably phonological) codes during verbatim repetition (see also Bates et al., 1980; Bramwell, 1897, reprinted in 1984).

The Mechanism of Verbatim Sentence Repetition: Two Frameworks

Although there is far from conclusive evidence as to the dynamics of the mechanism responsible for immediate verbatim sentence repetition, the available evidence suggests a few possible accounts. Two frameworks will be considered for the current discussion, namely that provided by Bock and Levelt (1994), and the multiple capacity view of STM offered by R. Martin and colleagues (discussed in detail in the previous section; with emphasis on a similar model offered by N. Martin and Saffran (1992)).

Potter and Lombardi have claimed that the process of sentence repetition uses the same production mechanisms as those used in normal speech production. Before considering such a claim, however, a brief review of these mechanisms is warranted, and Bock and Levelt (1994) will be used as a guide. According to their modular view, speech production proceeds as follows: The speaker has a message that she wishes to convey and this (conceptual) information is used by the functional level processes of lexical selection and function assignment; the message selects lexical items (“lemmas”) for production; then function assignment occurs, which involves assigning syntactic relations
or grammatical functions (e.g. subject-nominative, object-dative); then positional processing occurs (which entails the processes of constituent assembly and inflection) which is responsible for fixing the order of elements in an utterance; finally, phonological encoding is executed. As described by Bock and Levelt, the process of constituent assembly involves “the creation of a control hierarchy [i.e. a framework] for phrasal constituents that manages the order of word production and captures dependencies among syntactic functions” (p.947). The other component of positional processing, inflection, involves (at least for English) the generation of elements which correspond to information about tense and number (among other things). Bock and Levelt also claim that (for normal speech production) each level (message, functional level, positional level, phonological encoding) is only influenced by information at the previous level. So, for example, functional processes would be affected by the message level but not the phonological level. They also argue that language production is incremental so that higher levels (e.g. functional) need not complete their processing before lower level processing (e.g. positional) begins.

Potter and Lombardi (1990) advocate the use of these language production mechanisms during verbatim recall of sentences. They argue that short-term sentence recall, similar to long-term sentence recall, begins with a conceptual representation of the sentence (i.e. the message); subsequent functional, positional, and phonological processing occurs which is no different from normal language production. However, Potter and Lombardi (1990) argued that, unlike long-term recall, short-term recall has available a set of recently activated lexical entries corresponding to the words in the sentence to be recalled. As a result, the message level would preferentially select those
activated lexical entries during functional processing. Alternatively, as in the current set of experiments, a lexical selection error could occur (see Garrett, 1980). Notably, this account is consistent with the notion advocated by Bock and Levelt (1994), namely that the process is unidirectional, with phonological processing having no influence upon earlier processes. Unfortunately, this account, though parsimonious, has been shown by the current findings to be incomplete.

In the current set of studies, substantial evidence has been provided that a phonological representation is used during sentence recall, and further, that phonological information is used to select lexical items during functional processing (recall the advantage of SP over S intrusions found in all relevant experiments). Therefore, as argued by Potter and Lombardi (1990), sentence recall involves more than simply reading off a phonological record of the sentence. However, a (substantial) modification of Potter and Lombardi's (1990) view is proposed, namely that phonological, as well as semantic, information is used to select lexical items for recall.

To account for the role of phonology in sentence recall, the account offered by Potter and Lombardi (1990) must be modified to incorporate bidirectional information processing in their model, or alternatively, that separate memory stores (perhaps those used to store the products of comprehension processes) send information to the production system and thereby influences its processes; i.e. that phonological information stored elsewhere as a result of comprehension processes is used to select lexical items for production (during sentence recall). A bidirectional (interactive activation) account (of lexical selection) would be consistent with the approach of Dell (1986), who has argued that in normal production, phonological similarity increases the probability of a semantic
substitution (Dell & Reich, 1981). Stemberger (1985) has also argued that, although infrequent, word substitutions involving both strong semantic and strong phonological similarity are still more likely than chance (e.g. "lobster" for "oyster"; Garrett, 1993).

In the case of purely phonological substitutions (Experiment 5), it can be argued that the information maintained in the postulated memory store activates lexical entries (during functional processing) sufficiently to override the selection most consistent with the message level output. Further, one intriguing footnote to this account is the finding that subjects were occasionally found to change the meaning of the sentence during recall to accord with the intruded word. Consider the example: "The disagreement led to an altercation in the school parking lot" --&gt; "The disagreement led to an alteration in the rules of school parking." This evidence seems to suggest that the message level can be altered subsequent to a lure intrusion (at the functional level). On the surface, this evidence seems to suggest bidirectional processing between the functional and message levels, yet more evidence would be required to assert such a claim especially considering the possibility that subjects may have misperceived the stimuli during presentation.

Perhaps a more parsimonious account of the data would be possible through a model which incorporates different levels of representation in a bidirectional spreading activation network. N. Martin and Saffran (1992) offered a model based on Dell's (1986) interactive activation network model of speech production which focuses on the interactive processes involved in lexical selection. It is unique in the sense that it can account for both input and output processes. Furthermore, it consists solely of semantic, lexical, and phonological levels and purposefully neglects syntactic and multi-word utterance processes. Of course, with regard to sentence-level tasks (such as sentence
repetition), selection of lexical items would be followed by insertion (or linkage) of the items into a syntactic frame for production. It is noteworthy that in the N. Martin and Saffran adaptation of Dell’s model, they point out that along the lines of R. Katz and Goodglass (1990), deficits in one level of representation can be revealed as errors of another type. For instance, as Katz and Goodglass argued in their account of semantic errors in repetition (commonly found in deep dyslexia), rapid loss of phonological information which normally serves to stabilize semantic representations (via activation at the lexical level) would lead to inadequate activation of semantic features (at the semantic level of the network), thus resulting in semantic errors in repetition. According to Dell’s model, nodes decay over time and feedback serves to stabilize primed lexical nodes.

A thorough and detailed account of the spreading activation processes of the model will not be discussed here (see N. Martin & Saffran, 1992); rather, the ability of its general framework to account for the results of the present study will be evaluated (see also Martin & Lesch, 1996, for a discussion of how the N. Martin & Saffran model relates to that of R. Martin and colleagues). Using the N. Martin and Saffran (1992) model, the process of sentence repetition in the Potter and Lombardi (1990) paradigm can be elaborated as follows: The speech production process begins with the conceptual layer (which consists of semantic nodes); this can be seen as analogous to the conceptual representation of the sentence in the Potter and Lombardi paradigm. From this point, activation spreads to the lexical level where nodes consistent with the conceptual representation will be activated. Presumably, activation of the lexical level has already occurred from presentation of the sentence words themselves (as well as the list words
containing possible lures). If lures had not been presented, the conceptual representation would activate lexical nodes consistent with it; since the words actually presented in the sentence have already primed their corresponding lexical nodes, they have a higher activation level (than other semantically compatible lexical nodes) and are thus selected resulting in flawless sentence repetition. In contrast, if semantically related lures are presented in the list, they will activate their corresponding lexical nodes as well making correct selection of the target lexical node more difficult (potentially causing intrusions).

Although the model provides an excellent framework for elaborating Potter and Lombardi's theoretical view, it was found in the current set of studies that phonological information is involved in the process of sentence repetition as well. Fortunately, N. Martin and Saffran's network model is able to deal with this finding with equal elegance. In fact, in their study of patient EA, R. Martin, Shelton, and Yaffee (1994) argued that the model could account for her tendency to paraphrase in sentence repetition by postulating that activation at the phonological level decreased abnormally rapidly relative to the decrease in activation at the semantic level.¹⁶

According to the current interpretation, a phonological representation of the sentence is available to guide lexical access and this claim can be realized in the network model simply as activation at the phonological level. When the to-be-repeated sentence is perceived, it activates nodes at the phonological level (for each word). For the target word in particular, its corresponding phonological activation will send activation to the target lexical node (as well as phonologically related nodes). Concurrently, the conceptual representation (of the target word) will send activation to the target lexical
node as well as semantically related lexical nodes. As would be predicted by this account, if lexical nodes are consistent with both the semantic (conceptual) and phonological representations of the target word, their activation level will be greater than other words which are just semantically related, only phonologically related, or completely unrelated. Further, since the target word and lure word are already activated (by being presented in the sentence or list), they end up with higher activation levels than other semantically and/or phonologically related words and are thus more likely to be selected (preventing a high error rate in normals). In accordance with this line of reasoning, it was found that semantically and phonologically related (SP) lures were more likely to substitute for the target word than lures only related semantically. Similarly, the purely phonological substitutions found in Experiment 5 can be explained by phonological activation (by the phonological lure word as well as the target word) of phonologically related lexical nodes, with one such related node being the already primed lure word.

Also, N. Martin and Saffran’s model provides a foundation for the account that was provided for the results of Experiment 4. First, the greater (SP - S) intrusion rate difference for abstract than concrete stimuli alluded to earlier can be explained by the model as follows. If the target words are abstract, semantic feedback will be less efficient in supporting the priming of the lexical node, and thus feedback from the phonological nodes will have more influence. On the other hand, concrete target words will have more efficient or effective feedback from the semantic level to the lexical level, resulting in a diminished influence of phonology.
Also, it was likewise argued that semantic intrusions could be more likely to occur for abstract than concrete sentences due to the stronger semantic representation of concrete sentences. According to the model, a strong semantic representation would send great amounts of activation to correct lexical items (i.e. the words of the sentence to be recalled), and thus fewer semantic intrusions would be expected.

Conclusions

The experiments reported here provide strong support for the hypothesis that a phonologically-based memory representation is used in the process of sentence repetition. It was argued that this representation works in an interactive fashion with semantic (and possibly syntactic) memory representations to allow accurate repetition of long sentences (the “Combination” view). Furthermore, evidence has been provided which suggests a phonological representation of greater capacity and durability than previously thought, due to the supportive influence of other memory codes. It is argued that models of STM must account for the interactive mutual support of memory representations demonstrated by sentence repetition studies, and further research is necessary to evaluate the claim that phonological short-term memory is more long-term and prominent than previously thought.
Footnotes

1 Cole & Jakimik (1980) review evidence that during auditory sentence processing, semantic constraints provided by one word are used during recognition of an immediately following word.


3 There were really only four important sets; the multiple of 5 only affected the unrelated condition. The semantic judgment task was created so that each target would be presented (within the sentence) paired with a lure or unrelated word presented within its own copy of the same sentence as well. However, there was no particular unrelated word that best represented the unrelated condition; therefore, all five were used on separate versions of the task, creating the five versions of the four sets (corresponding to the SP, S, P, and U lures respectively).

4 The actual duration used in the Potter and Lombardi (1990) experiment was 517 msec, but 520 msec was used due to technical limitations.

5 A rating for the word “reply” used by Potter and Lombardi (1990) was unavailable, so this mean is based on 19 stimulus targets.

6 It is acknowledged, as Anderson and Paulson found, that there is a substantial weakening of the verbatim trace over the first 30 seconds (of delay). However, after this initial dropoff, the trace is remarkably durable and resistant to further decay.

7 However, the criticism could be raised that the meaning of the sentences is slightly altered with voice conversion. Therefore, the trace need not necessarily be phonological, but rather meaning-based.
8 An implication of their findings is that the results of Anderson & Paulson, 1977, and Garrod & Trabasso, 1973, may have an alternative explanation. It is possible that people retain thematic information such as the referent of the (grammatical) subject rather than a verbatim record (at least for longer delay intervals).

9 The total proportion of spontaneous phonological substitutions is 6% because the “PES” and “POS” substitution rates were based on different numbers of “total possible trials”: While PES substitutions could be made for only Unrelated condition trials, POS substitutions could be made for both Phonological as well as Unrelated condition trials.

10 In the case of a nonsense word list, the acoustic trace cannot retrieve any activated lexical entries and must rely solely on itself for recall (resulting in a shorter nonsense word than word span).

11 Which would also account for the larger span found for words than nonwords and larger span for sentences than word lists.

12 A much simplified version of the theory proposed by Bock & Levelt (1994) is being outlined.

13 A lemma refers to a word as a syntactic entity, containing information (if applicable) such as word class, gender and subcategorization frame properties.

14 Bock and Levelt use traditional grammatical relations terminology (subject, direct object, etc.) to refer to where the elements that are assigned certain grammatical functions actually appear in the sentence (a result of positional processing). So, in English, the element that is assigned the nominative function appears in subject position.
This memory store would be distinct from the store used during phonological encoding, which is responsible for maintaining the wordshape frame as it is constructed.

They also acknowledged the possibility that EA had an additional deficit to the lexical level.
References


