RICE UNIVERSITY

Semantic Short-Term Memory and Resolution of Interference: Patient, ERP and fMRI Data

by

Andrew Cris Hamilton

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Randi C. Martin, Elma Schneider
Professor, Chair
Psychology

Tony Ro, Associate Professor
Psychology

E. Darcy Burgund, Assistant Professor
Psychology

Geoff Potts, Assistant Professor
Psychology

Eric Margolis, Associate Professor
Philosophy

HOUSTON, TEXAS

APRIL 2005
INFORMATION TO USERS

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 bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send 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.

UMI Microform 3170793
Copyright 2005 by ProQuest Information and Learning Company. All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346
Abstract

Experiment 1 presents data from a patient with a semantic short-term memory deficit, patient ML, that indicate profound susceptibility to interference. For example, although this patient cannot reliably recall three items in a serial recall task, he paradoxically shows exaggerated effects of proactive interference in short-term memory. However, this patient’s difficulty with interference appears to be limited to tasks involving verbal stimuli– other data show that patient ML performs normally on two nonverbal tasks that require resolution of interference.

Experiment 2 attempted to identify ERP components related to proactive interference in one of the tasks administered in Experiment 1. This task, the recent negatives task, is a convenient measure of susceptibility to proactive interference. Moreover, Experiment 2 added an additional manipulation motivated by a unique effect discovered during the testing of patient ML (Hamilton, 2004), whereby the patient performed much better on the recent negatives task when repetition was minimized and the number of stimuli presented within the task were expanded. Two ERP components, a frontally distributed N400 effect and parietally distributed late positive component (LPC), were found to respond to the manipulation of recency and repetition.

Experiment 3 and Experiment 4 examined a language comprehension paradigm known to differentiate between patients with semantic and phonological short-term memory deficits. This task requires detection of semantic anomalies in phrases in which multiple adjectives appear before or after a noun – multiple adjectives appearing before a noun are believed to place greater demands on semantic short-term memory relative to when adjectives appear after a noun. Thus, patients with semantic short-term memory
deficits are especially poor at detecting anomalies in the before condition. Experiment 3 uses the parietally distributed N400 as an indirect measure of short-term memory demands to corroborate behavioral and patient data. Relative to the "after" condition, the "before" condition did elicit smaller N400s, consistent with the idea that integration of adjectives in the before condition differs from integration in the after condition.

Experiment 4 uses a modification of the sentence anomaly task employed in Experiment 3 to identify brain areas engaged in short-term maintenance of semantic representations. It was hypothesized that the before condition would produce greater activation in the left inferior frontal gyrus, a region that has been related to short-term maintenance of semantic representations. Results from Experiment 4 are discussed in terms of the organization of maintenance and control processes important in semantic short-term memory.

Finally, data from Experiments 1-4 are discussed in terms of their implications for theories of semantic short-term memory deficits and the associations of semantic short-term memory deficits with particular deficits of language comprehension and production.
ACKNOWLEDGMENTS

I would like to thank Dr. Randi C. Martin for her generous supervision, support and giving me the opportunity to study at Rice University. Thanks are due to Dr. Phil Burton, who patiently supervised the analysis of fMRI data. I am also indebted to Dr. E. Darcy Burgund for her assistance with fMRI analysis. Thanks are also due to Dr. Geoff Potts and Laura Martin, who opened their laboratory and equipment to me and provided much direction in the analysis of ERP data. Michelle Lipszyc coordinated the programming of Experiment 3 and collection of pilot data. Dr. Chin-Lung Yang programmed Experiment 3. Cari Zimmerman contributed to collection of the pilot data for Experiment 4. I would like to thank the members of my dissertation committee, Drs. E. Darcy Burgund, Eric Margolis, Randi C. Martin and Tony Ro, for their comments on this work and accommodating my schedule.

Finally, I would like to thank my father, H.G. Hamilton, and wife, Paula Lewis, for their indefatigable personal and financial support. This dissertation is dedicated to my mother, Barbara A. Hamilton (1962-1990).
Table of Contents

ABSTRACT...........................................................................................................ii
Acknowledgements..............................................................................................iv
List of Tables.......................................................................................................vi
List of Figures.......................................................................................................vii
Introduction........................................................................................................1
Experiment 1: Dissociations Among Tasks of Inhibition...........................................5
Experiment 2: Recent Negatives Task in ERP.......................................................7
  Method..............................................................................................................13
  Results.............................................................................................................15
  Discussion.......................................................................................................23
Experiment 3: Before/After Sentence Anomaly Task in ERP.................................31
  Method..............................................................................................................38
  Results.............................................................................................................42
  Discussion.......................................................................................................44
Experiment 4: fMRI of Before/After Anomaly Paradigm.......................................47
  Method..............................................................................................................54
  Results.............................................................................................................57
  Discussion.......................................................................................................61
General Discussion..............................................................................................70
References..........................................................................................................79
Appendix A.........................................................................................................126
List of Tables

Table 1  Behavioral Data – ERP Recent
         Negatives Task – Undergraduate
         Subjects

Table 2  Stimuli for Experiment 3 – Before
         Condition

Table 3  Stimuli for Experiment 3 – After
         Condition

Table 4  Experiment 3 – Sample Sentences
Table 5  Filler trials in Experiment 3
Table 6  12 ROIs identified by task vs.
         baseline contrast.
Table 7  Stereotaxic coordinates of ROIs
         demonstrating significant
         Before/After X Time interactions in
         Experiment 4 compared to several
         neuroimaging studies of interference
         resolution, short-term maintenance
         of semantic representations and
         short-term memory load.
**Lists of Figures**

<table>
<thead>
<tr>
<th>Figure 1</th>
<th>Patient ML accuracy in the Open and Closed versions of the Recent Negatives Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2</td>
<td>Anterior Posterior Channels used for Frontal N400 (FN400) analyses</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Experiment 2: FN400 component – Recent Negative v. Nonrecent Negatives in Open and Closed Conditions.</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Experiment 2 - FN400 – Open vs. Closed X Recent vs. Nonrecent.</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Experiment 2 – FN400 effect in Closed condition with positive trials.</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Experiment 2 – FN400 effect in Open condition with positive trials.</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Negative and Positive Trials in Closed Condition</td>
</tr>
<tr>
<td>Figure 8</td>
<td>FN400 for Open condition: Interaction of Trial Type X Laterality</td>
</tr>
</tbody>
</table>
Figure 9
Experiment 2 – Late Positive
Component (550ms – 650ms) –
Open and Closed conditions.

Figure 10
Experiment 2 – Posterior Superior
Channels used in analyses of Late
Positive Component.

Figure 11
Experiment 2: Late Positive
Component – 550ms – 650ms: Set
Size X Recency Interaction

Figure 12
Experiment 2: Late Positive
Component for Recent vs. Nonrecent
negatives in Open condition

Figure 13
Schematic of short-term memory
representations necessary in the
Before condition in Experiment 3

Figure 15
Sentence Anomaly Task (a.) reaction
time data for control Ss (b.) accuracy
data for semantic STM patient AB
(c.) accuracy data for semantic STM
patient ML (d.) accuracy data for
phonological STM patient EA
Figure 16  
Before vs. After X Sensibility interaction with visual presentation of sentences

Figure 17  
Experiment 3: N400 effect at parietal channels.

Figure 18  
Experiment 3: Main effect of Sensibility

Figure 19  
Experiment 3: After/ Before X Laterality interaction.

Figure 20  
Twelve ROIs defined by $t > 11.32$

Figure 21  
ROIs 1, 2, and 4

Figure 22  
ROI 3

Figure 23  
ROI 6 – Before/ After X Sensible/ Anomalous X Time interaction

Figure 24  
ROI 10

Figure 25  
ROI 17

Figure 26  
ROI 18

Figure 27  
ROI 25

Figure 28  
ROI 28

Figure 29  
Two ROIs in left middle temporal gyrus

Figure 30  
(a.) ROI 28 (BA 40) in Experiment 4
(b.) parietal region identified by
Becker et al. (1999) as anatomical locus of phonological store.
Semantic Short-Term Memory and Resolution of Interference:

Patient, ERP and fMRI Data

Distinct patterns of patient performance on a number of tasks support a dissociation between semantic and phonological capacities in short-term memory (STM). For example, patients with semantic short-term memory deficits typically demonstrate no advantage of words relative to nonwords (no lexicality effect), while patients with phonological short-term memory deficits fail to demonstrate normal phonological similarity effects. Patients with semantic short-term memory deficits perform better on a rhyme probe task (requiring phonological maintenance) relative to category probe task (which requires maintenance of semantic information), while phonological short-term memory patients perform better on the category probe than rhyme probe task (Martin & Romani, 1994; Martin, Shelton, & Yaffee, 1994). It is believed that patients with semantic short-term memory deficits are unable to accurately maintain lexical-semantic representations in short-term memory and must therefore rely on phonological representations for any short-term maintenance that they manage. On the other hand, patients with phonological short-term memory deficits are unable to maintain phonological information and depend upon lexical-semantic representations during recall. Similar patterns of dissociations have been reported by N. Martin and Saffran (1997).

Of relevance to the present studies, Martin and Lesch (1996) reported a particularly interesting observation in patients with semantic short-term memory deficits. When presented with serial recall tasks, these patients made numerous intrusions of previously presented list items. Thus, although these patients have great difficulty recalling a list of even three words, they nevertheless experience spontaneous intrusions
of words from previous lists. This effect is surprising, given that if patients’ short-term memory deficits are attributed to reduced maintenance capacity or an abnormally rapid rate of decay, one might expect fewer intrusions from previous lists. Interestingly, this paradoxical effect of intrusions in patients with greatly limited short-term memory has only been observed in cases of semantic short-term memory deficits.

In Experiment 1 (included as an appendix), these intrusions are conceptualized as an apparent failure to inhibit processing of previously presented items, which produces exaggerated effects of proactive interference. Data from Event Related Potentials (ERP) are then presented to further examine the temporal dynamics and neural basis of proactive interference. Another ERP study uses the N400 ERP component as an indirect measure of the consequences of increased semantic short-term memory load during sentence comprehension. A fourth study attempts to elucidate the neural substrate of semantic short-term memory using the superior spatial resolution of functional magnetic resonance imaging (fMRI). Thus, these experiments may potentially answer several questions regarding the relationship between semantic short-term memory and interference resolution.

Anatomically, semantic short-term memory deficits are associated with lesions of the inferior frontal areas of the left hemisphere. This observation is intriguing given recent neuroimaging data implicating inferior frontal areas in functions such as semantic selection (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997) and resolution of proactive interference (Jonides, Smith, Marshuetz, Koepppe, & Reuter-Lorenz, 1998). Other interference effects associated with semantic short-term memory deficits, possibly attributable to failures of inhibition, have also been reported – for example, Freedman,
Martin, and Biegler (2004) reported that patients with semantic short-term memory deficits demonstrated greatly exaggerated interference effects when producing a conjoined noun phrase to describe two semantically related items (compared to two unrelated items). These observations suggest a possible role for impaired inhibitory function in semantic short-term memory deficits. Such an inhibitory deficit may also have important implications for semantic short-term memory patients' speech production and comprehension deficits.

Given the prominent role of inhibition and resistance to interference in many contemporary models of working memory (Chiappe, Hasher, & Siegel, 2000; Hasher, Quig, & May, 1997; Hasher, Stoltzfus, Zacks, & Rypma, 1991; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2000, 2002, 2003; Kane, Hasher, Stoltzfus, Zacks, & Connelly, 1994; Lustig, May, & Hasher, 2001; May, Hasher, & Kane, 1999; Stoltzfus, Hasher, Zacks, Ulivi, & Goldstein, 1993) and the aforementioned patient performance suggesting extreme vulnerability to interference, the initial goal of this research is to examine the performance of a semantic short-term memory patient on tasks reported to be associated with inhibition. These data answer several questions. First, would a patient with a semantic short-term memory deficit show deficits on all tasks, both verbal and nonverbal, requiring resolution of interference? Furthermore, would such a patient show deficits on tasks that make no obvious demands on short-term memory? Secondly, given the unusual number of intrusions during short-term recall, would semantic short-term memory patients show similar effects on other tasks requiring resolution of proactive interference? These questions were addressed in Experiment 1.
Other studies explored neural processes involved in inhibition and interference resolution in semantic short-term memory. Experiment 2 attempted to identify an ERP component that indexes proactive interference in a probe recognition task tapping short-term memory (the recent negatives task). Finding such a component would be useful in determining the temporal dynamics of interference in the recent negatives task. For example, at what point in a task is interference resolved or inhibition necessary? Moreover, further elucidation of such components might be applicable to ERP studies of patients with deficits of interference resolution.

Experiment 3 also used electrophysiological methodology to examine the N400 as an indirect measure of semantic short-term memory using a task adapted from patient studies (Martin and Romani, 1994). This sentence anomaly task is believed to place unique demands on semantic short-term memory.

In addition to two studies employing ERP methodology, an fMRI experiment was conducted. Although several studies (Barde & Thompson-Schill, 2002; Collette et al., 2001; Crosson et al., 1999; Martin, Wu, Freedman, Jackson, & Lesch, 2003; McDermott, Petersen, Watson, & Ojemann, 2003) have examined the neural substrate of semantic and phonological processing used in short-term memory, the results are quite disparate, with a number of different tasks employed. The current experiment uses a task known to discriminate between patients with phonological and semantic short-term memory deficits. Data from this study can be compared with previous neuroimaging studies of proactive interference in short-term memory (D'Esposito, Postle, Jonides, & Smith, 1999; Jonides et al., 2000; Jonides et al., 1998; Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003; Postle, Berger, Goldstein, Curtis, & D'Esposito, 2001) and other
interference paradigms. Thus, data from Experiment 4 will be compared with studies relating resolution of interference to activity in the left inferior frontal gyrus.

EXPERIMENT 1

Complete details of Experiment 1 can be found in Hamilton and Martin (2005), which is included as Appendix A. Relevant experiments are summarized here. Patient ML was tested on two verbal tasks of inhibition, a Stroop task and a recent negatives task, and two nonverbal tasks, a nonverbal spatial analogue of the Stroop task and the antisaccade task.

In the Stroop task, color words (RED, BLUE, YELLOW, GREEN, PURPLE) appear in different colored fonts (ie., red, blue, yellow, green, purple). The subjects’ task is to name the color that the stimulus appears in – for example, if the word RED was presented in a blue font, the subject would respond “BLUE”. The prepotent tendency to read the word conflicts with the instruction to name the color and is the source of interference. The interference effect is characterized by subjects taking longer to name the color on incongruent trials (the word RED appearing in blue) than on neutral trials (a row of asterisks, *****, appearing in blue). The recent negatives task is another task of interference and inhibition which takes advantage of an interference effect in short term memory. This effect is seen when previously presented material interferes with the ability to quickly judge whether a currently presented word appeared in a list. The antisaccade task is sensitive to one’s ability to inhibit reflexive eye movements to sudden onsets of stimuli presented in the periphery of the visual field. Instead of making eye movements to the stimulus, subjects are asked to make an eye movement in the opposite
direction, in order to detect a target on the opposite side of the display. Finally, the nonverbal Stroop requires the resolution of conflict when the local direction and spatial position of a stimulus conflict. For example, when presented a right-pointing arrow on the left half of a display, subjects have to indicate which direction the arrow is pointing. This conflict between the direction the arrow is pointing and its spatial position is the source of an interference effect. For a more thorough description of each of these tasks, see Appendix A.

The Stroop and antisaccade tasks were chosen because they loaded on an inhibition factor in a latent variable analysis of executive function (Miyake et al., 2000). The recent negatives task (Monsell, 1978) was chosen as a way to examine proactive interference in short-term memory. The Stroop task is perhaps the quintessential measure of inhibition and interference. The recent negatives task has been used extensively in the neuroimaging literature (D'Esposito et al., 1999; Jonides et al., 2000; Jonides et al., 1998; Nelson et al., 2003; Postle et al., 2001) and is a convenient means of inducing proactive interference in short-term memory. The nonverbal Stroop task was developed by the authors to be analogous to the conventional Stroop task, in that it required resolution of nonverbal conflict. The antisaccade task is a nonverbal task requiring suppression of reflexive saccades to the sudden onset of a cue. This task, although having no obvious short-term or working memory requirements, has often been reported to be correlated with working memory ability (Kane et al., 2001; Mitchell, Macrae & Gilchrist, 2002; Roberts, 1994).

Patient ML's interference effects on the Stroop task and recent negatives task were well outside the range of tested controls matched for age and education. Thus, ML
was impaired on both verbal tasks but performed normally on the nonverbal tasks, suggesting a distinction between inhibition in verbal and nonverbal domains. ML’s data also represent a dissociation between Stroop and antisaccade performance, two tasks that load on a single factor in factor analytic studies (Miyake et al., 2000). These data suggest that ML’s difficulty with inhibition and interference resolution does not extend to all tasks – ML has difficulty only with verbal tasks.

Thus, the results of Experiment 1 are at odds with many of the individual differences studies examining these tasks with healthy subjects. Many of these studies have proposed that a single inhibitory mechanism, localized to a single brain region, the dorsolateral prefrontal cortex, is involved in performance of all of these tasks (see Kane & Engle, 2002). Hamilton and Martin (2005) proposed that instead of a common region in the brain, the correlations among these tasks may instead result from the activity of a common neurotransmitter, such as dopamine. In this scheme, separate brain regions may be involved in various inhibition tasks, but all of these areas are “fueled” by the same neurotransmitter. See Hamilton and Martin (2005) in Appendix A for a complete discussion of these data and their implications for organization of inhibitory mechanisms in the brain.

EXPERIMENT 2

Although the recent negatives task (Monsell, 1978) has been used extensively in the functional neuroimaging literature (D'Esposito et al., 1999; Jonides et al., 2000; Jonides et al., 1998; Nelson et al., 2003; Postle et al., 2001), attempts to use electrophysiological methodologies with the task have not yet appeared in the literature. Electrophysiological methods (e.g. ERP) have the advantage of providing superior
temporal resolution, a dimension that may prove important in the recent negatives task. Although previous fMRI work suggests that the resolution of interference occurs during the presentation of the probe (D'Esposito et al., 1999), ERP is better suited to examining such temporal questions.

The recent negatives task is a simple probe recognition paradigm in which subjects are presented a list of items followed by a probe. Subjects indicate whether or not the probe appeared in the present list. However, a further manipulation alters the recency of the probe as described below. There are four conditions.

**Recent negative**- A list is presented, followed by a probe that did not appear in the list (a negative probe). However, the probe did appear in the list presented immediately before the present trial.

For example:

<table>
<thead>
<tr>
<th>List</th>
<th>Probe</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>K L M P</td>
<td>B</td>
<td>Response: NO</td>
</tr>
<tr>
<td>T V R X</td>
<td>P</td>
<td>Response: NO</td>
</tr>
</tbody>
</table>

In this example, the probe “P” did not appear in the present list, although it did appear in the list presented immediately before the current trial. The match between the probe and any persisting representation of the previous list is the source of the interference effect. The effect is behaviorally indicated by longer reaction times and poorer accuracy in this condition relative to the Non-Recent condition.

**Non-Recent Negative**- In this condition, a probe is not present in the current list but did appear previously. However, in the non-recent condition, the letter did not appear in either of the two previous trials.
For example:

<table>
<thead>
<tr>
<th>List</th>
<th>Probe</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>K V R X</td>
<td>T</td>
<td>“NO”</td>
</tr>
<tr>
<td>G L D P</td>
<td>J</td>
<td>“NO”</td>
</tr>
<tr>
<td>W M S Z</td>
<td>F</td>
<td>“NO”</td>
</tr>
<tr>
<td>B D F C</td>
<td>X</td>
<td>“NO”</td>
</tr>
</tbody>
</table>

**Recent Positive** - In addition to the negative trials, positive trials are also presented. In the recent positive trials, the probe letter appeared in the current list and was presented in the list presented immediately before.

For example:

<table>
<thead>
<tr>
<th>List</th>
<th>Probe</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>F L D J</td>
<td>W</td>
<td>“NO”</td>
</tr>
<tr>
<td>B N X D</td>
<td>D</td>
<td>“Yes”</td>
</tr>
</tbody>
</table>

**Nonrecent Positive** - In this condition the letter probe did appear in the current list (necessitating a positive “yes” response), but did not appear in the two previous lists. For example:

<table>
<thead>
<tr>
<th>List</th>
<th>Probe</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>K V R X</td>
<td>T</td>
<td>“NO”</td>
</tr>
<tr>
<td>G L D P</td>
<td>J</td>
<td>“NO”</td>
</tr>
<tr>
<td>W M S Z</td>
<td>F</td>
<td>“NO”</td>
</tr>
<tr>
<td>B X F C</td>
<td>X</td>
<td>“Yes”</td>
</tr>
</tbody>
</table>
Previous work (Hamilton, 2004) has demonstrated a novel effect in patient ML's performance on the recent negatives task. Namely, his performance differed according to the number of stimuli that were presented in a given block of trials - when only 16 trials were repeatedly presented (as was the case with the original Monsell paradigm), ML performed poorly on both recent and nonrecent negative trials (see Figure 1a). However, when the number of trials was expanded, ML was much less accurate on recent negatives, while his performance improved on nonrecent negatives (see Figure 1b). Moreover, when only 16 stimuli were used, patient ML showed substantial facilitation for recent positives relative to nonrecent positives as measured by accuracy. This manipulation has been conceptualized as manipulating “inhibition demands” required in the recent negatives task. When only 16 items are repeated throughout the task, there is increased interference among items (presumably due to increased activation and familiarity of all items), thereby increasing demands for inhibition of previous items and making the discrimination between old and new items generally more difficult. When a greater number of items are presented, there are fewer repetitions of stimuli, thus producing less interference among items.

This effect, first observed in patient ML, is explored in Experiment 2 with healthy undergraduate subjects. It was hypothesized that the manipulation of stimulus-set size would be useful in identifying which ERP components were related to resolution of interference in the recent negatives task. For Experiment 2, blocks in which 16 stimuli are repeatedly presented will be referred to as “closed” blocks, while those with a greater number of unique stimuli will be referred to as “open” blocks. Ideally, components
related to resolution of interference and inhibition in the recent negatives task should also respond to the manipulation of stimulus-set size ("open" v. "closed" conditions in the task). Again, Experiment 2 will potentially provide useful data regarding the temporal dynamics of proactive interference in a short-term memory paradigm. Candidate components could also be used to further examine the deficits in patients showing exaggerated interference effects in the recent negatives task.

No published studies have specifically examined the recent negatives task in the ERP environment; thus, no specific ERP component has been previously identified as sensitive to the recency manipulation. However, previous EEG work has identified several components that have been related to response inhibition in the go/ no-go task. In the go/ no-go task, subjects are to respond to a target stimulus (go trials) and to withhold response to a non-target stimulus (no-go trials). A negative component with a latency of 150- 400 ms after presentation of the stimulus is named the N2. Given that larger N2 amplitudes are observed following a no-go response relative to a go response, the N2 has been related to inhibition (Jodo & Kayama, 1992). However, it is important to note that the go/no-go paradigms have typically involved nonverbal stimuli. As demonstrated in Experiment 1, there is compelling evidence that the mechanisms involved in inhibition in a nonverbal domain are very different from the mechanisms engaged for the inhibition of irrelevant verbal information. Thus, while the N2 is of potential interest here, differences between the recent and nonrecent negative trials may not be apparent in this component.

Other components of interest were the frontal N400, which has been related to familiarity processes in recognition memory tasks (Curran, 1999), and a P600 or late
positive component (LPC) which has often been related to recollection or response confidence (Finnegigan, Humphreys, Dennis & Geffen, 2002). Thus, differences between the recent negative vs. nonrecent negative trials may be apparent in the N400 and late positive components.

Behaviorally, one might expect that blocks with “closed” stimulus sets (small stimulus-set size) would generate more interference throughout all trials. As discussed previously, in the same experiment presented here, patient ML’s accuracy was very poor on both the recent and nonrecent negatives in the closed condition. However, his accuracy on the nonrecent negatives improved greatly in the open condition. Given that both of these blocks were presented during the same experimental condition, it seem unlikely that a systematic response bias (i.e., a bias to say yes to all negative trials) can explain ML’s poor performance on recent and nonrecent negative trials in the closed condition. One explanation is that patient ML experienced considerable proactive interference on all of the negative trials, such that sufficient interference existed to induce a “yes” response, even on the nonrecent negatives trials.

However, it is important to emphasize that behaviorally, interference effects are calculated by subtracting the reaction times (RT) of nonrecent negatives from recent negatives. As suggested by patient ML’s accuracy data, reaction times on the nonrecent negatives in the closed condition might also be influenced by small stimulus-set size. In fact, pilot data suggested that the effects of the closed vs. open conditions are seen primarily in the RT for the nonrecent negatives, which were longer in the closed condition. Because there is greater interference on the non-recent negatives in the closed than the open condition, a smaller interference effect for recent vs. non-recent negatives
may be observed. That is, in the closed blocks, nonrecent negatives may be subject to interference as well, resulting in longer RTs in this condition and thus a smaller interference effect. Therefore, the contrast between nonrecent negatives in the closed and open blocks may prove important in detecting any effect of stimulus set size.

Method

Subjects. Twenty-six subjects were recruited from the Rice University community. All subjects participated in four blocks each of two conditions.

Materials. There were two types of trial blocks – one presented blocks of trials repeatedly sampled from the same list of 16 words (“closed” condition), while the other condition sampled items from a much larger list of words (“open” condition). Trials were presented in a fixed pseudo-randomized order. The closed lists consisted of 16 one-syllable words of comparable imageability and frequency ratings as determined by Francis-Kucera ratings (Francis and Kucera, 1982). For the open list, over 500 one-syllable words were used. There were 4 blocks within each condition, for a total of 8 blocks. Each block was comprised of 39 trials. Additionally, there was a practice block of 10 trials, which used 16 consonant letters instead of words.

In the closed conditions, there were 4 types of trials: recent negative, non-recent negative, recent positive, and non-recent positive. Each trial consisted of three words and a probe word. Subjects judged whether the probe word was in the immediately preceding list. Half of the trials were negative trials in which the probe did not match an item in the present list. The other half of the trials were positive trials. For recent negative trials, the probe word was present in the previous list. For non-recent negative trials, the probe did not appear in the previous two lists, but may have occurred prior to that in the closed
condition. In order to minimize repetition of stimuli, there was no recent vs. non-recent distinction for positive trials in the open condition. In other words, all positive trials were nonrecent.

Each type of trial was presented an equal number of times over the eight blocks. Orders were counterbalanced across subjects, such that half received AABBAABB and half received BBAABBAA.

**EEG data acquisition and signal processing.** EEG data were acquired with a 128 channel Electrical Geodesics system (Electrical Geodesics, Inc., Eugene). EEG data were acquired continuously and referenced to the vertex with .1 - 100 Hz analog filtering and digitized at 250 Hz. The EGI Geodesic Sensor Net is a dense sensor array in a lightweight elastic thread structure containing plastic pedestals. Each pedestal contains a silver/silver chloride electrode housed in a synthetic sponge. The sponges are soaked in a saline solution to render them conductive. Application of all 128 channels takes approximately 15 minutes.

**Stimulus presentation and behavioral response collection.** Stimulus presentation and behavioral response collection was controlled by E-Prime 1.0 (PST, Pittsburgh). Visual stimuli were presented on an Apple 15" flat-panel active matrix Studio display to reduce 60 - 75 Hz monitor refresh electrical noise associated with CRT displays. Manual responses were collected with a 4-key microswitch keypad (Electrical Geodesics, Inc., Eugene, OR).

Subjects were seated in an adjustable chair and placed their chin in a chinrest. The chinrest was placed such that subject's eyes were 50 cm from the center of the screen. Subjects were instructed to remain as still as possible, with their eyes on the
fixation stimulus throughout the block. Subjects were encouraged to refrain from blinking as much as possible while the stimuli were being presented. Breaks were provided every 3 - 4 minutes so that subjects could rest their eyes.

Subjects were told that they would be viewing a set of 3 words followed by a brief delay and then a probe word. Subjects were asked to make a “yes” or “no” response to the probe by pressing a corresponding key on a response box using their left hand. Subjects were instructed to respond as quickly as possible without sacrificing accuracy.

Each trial began with an initial fixation point for 1500 ms. Three words were then presented serially for 1000 ms each with no inter-stimulus interval. All words were lowercase and white on a black background. Following the third list item, a 3000 ms retention interval was accompanied by five red asterisks that appeared in the center of the screen. Finally, a lowercase probe word was presented. The probe remained on the screen until the subject made a key press response.

Subjects first received 10 practice trials using consonant letters. After the practice block, subjects were allowed to initiate the experiment with a key press when they felt comfortable. Each block of 39 trials lasted about 5-7 minutes depending on the subject. Between each block, subjects were given a rest break.

Results

Behavioral Data

Behavioral data appear in Table 1. Eight subjects were removed due to unusable ERP data.

Reaction Time (RT). RTs on error trials were eliminated from the analysis. Overall, reaction times were slower in the closed than open set condition; however, the
size of the interference effect was similar in the two set size conditions. A 2 (Stimulus-Set Size - open v. closed) X 2 (Recency- recent negative v. nonrecent negative) repeated measures ANOVA, which included only the negative trials, confirmed these impressions. There was a main effect of Recency, $F(1, 17) = 49.907, p < .0001$, and a main effect of Stimulus-Set Size, $F(1,17) = 5.26, p = .035$, but no Stimulus-Set Size X Recency interaction.

For the closed condition, a 2 (Trial Type - negative vs. positive) X 2 (Recency – recent vs nonrecent) repeated measures ANOVA, which included positive trials, revealed a main effects of trial type $F(1, 17) = 15.82, p = .001$ and Recency $F(1, 17) = 15.905, p = .001$. The Trial Type X Recency interaction was also significant, $F(1, 17) = 12.179, p = .003$. On the negative trials, subjects took 81 ms longer to reject recent negatives (845ms) than nonrecent negatives (764ms), $t(17) = 4.24, p = .001$. The 9ms difference between recent positives (725ms) and nonrecent positives (734ms) was not statistically significant $t(17)=-.802, p = .434$. Thus the facilitation effect on positive trials observed for patient ML (Hamilton, 2004) was not apparent in undergraduate RT data.

On the open trials, a one-way ANOVA including positive trials showed a significant main effect of condition, $F(2, 34) = 21.705, p < .0001$. Subjects took 88 ms longer to correctly reject recent negatives (808ms) than nonrecent negatives (720ms), $t(17) = 5.937, p < .0001$. On the open trials, the mean response time on positive trials was 715ms.

There was no significant difference between the positive trials in the open condition compared to the nonrecent positives in the closed condition.
Accuracy. As shown in Table 1, accuracy was high in all conditions. A 2 (Stimulus-Set Size) X 2 (Recency) repeated measures ANOVA on error rates, using only negative trials, revealed a main effect of Stimulus-Set Size, $F(1,17) = 6.080, p = .025$, but no effect of recency, $F(1,17) = 2.937, p = .105$, and no Stimulus-Set Size X Recency interaction, $F(1,17) = 3.353, p = .085$.

For the closed condition a 2 (Trial Type – Negative vs. Positive) X 2 (Recency – Recent vs. Nonrecent) repeated measures ANOVA showed a significant interaction of Trial Type X Recency, $F(1, 17) = 9.751, p = .006$. This interaction reflects lower accuracy in the recent negative compared to the nonrecent negatives $t(1, 17) = 1.939, p = .069$ and higher accuracy in the recent positives compared to the nonrecent positives, $t(1, 17) = 2.793, p = .012$. Thus, when considering accuracy, the interference effect in the closed condition was marginally significant, while the facilitation effect in the positive trials did reach statistical significance.

For the open condition a one-way repeated measures ANOVA revealed a significant main effect of Trial Type, $F(2, 34) = 6.023, p = .006$. This effect is explained by lower accuracy in the positive trials when compared to each of the negative conditions (Recent Negatives vs. Positives, $t(17) = 2.468, p = .025$; NonRecent Negatives vs. Positives, $t(17) = 3.575, p = .002$.)

ERP Analysis

Data from 8 subjects were excluded from the ERP analysis. These data were removed due to an insufficient number of usable trials per cell. Unusable trials were attributable to subject movement, equipment failures and unacceptable impedances. Thus, 18 subjects were included in the ERP analysis.
EEG data were segmented off-line into 1000 ms epochs spanning 200 ms pre-stimulus onset to 800 ms post-stimulus onset. Data were digitally screened for artifact (eye blinks or movements, subject movement, or transient electronic artifact) and contaminated trials were eliminated. Remaining data were sorted by condition and averaged to create the ERPs. Averaged ERP data were digitally filtered at 20 Hz lowpass to remove residual high-frequency noise, baseline corrected over the 200 ms pre-stimulus period, and re-referenced into an average reference frame to remove topographic bias due to choice of reference site. Bad channels were then marked and replaced using interpolation from surrounding channels’ measurements surrounding the bad channel. The subject-averaged ERPs were averaged together to produce the the grand average waveforms. Statistical analyses were performed on the subject-averaged ERPs with the subject averages serving as the observations. The waveform plots were constructed from the grand averaged data. Statistics were extracted for temporal components that, from visual inspection and previous literature, were considered to be of theoretical and statistical significance.

While one of the advantages of high-density ERP nets is the greater spatial resolution relative to traditional EEG recordings, Experiment 2 used a more conservative regional averaging technique as recommended Dien and Santuzzi (2005). Although the regional approach sacrifices the fine resolution provided by high-density electrode nets, it addresses potential objections to selection of channels guided by mere visual inspection. Channels were averaged by region, based on the following dimensions – left/right, anterior/posterior, and superior/inferior. Such a strategy yields eight separate regions.
The extracted statistics for each component were then analyzed. Averaged amplitudes for all relevant channel regions during the selected temporal epoch were submitted to a 2 (Stimulus-Set Size) X 2 (Recency) X 2 (Laterality) repeated measures ANOVA using the Super ANOVA program. Given that the open and closed conditions had a different number of conditions (four conditions in Closed – recent negative, nonrecent negative, recent positive, nonrecent positive; three conditions in Open – recent negative, nonrecent negative, positive), it was not possible to calculate a single omnibus ANOVA to include all of the conditions in Experiment 2.

Contrary to findings from go/no-go paradigm, no difference in the N2 was evident between the recent and non-recent negatives. However, a frontal N400 effect and a later positive component over posterior channels did appear sensitive to these manipulations. As mentioned earlier, these components have been identified in previous studies of recognition memory.

**Frontal N400 component.** Analyses of the Recent Negatives and Nonrecent Negatives are reported first. Positive trials, which appear in Figures 5 and 6, were included in separate ANOVAs for open and closed conditions, as reported below. Mean amplitudes for left and right anterior superior regions (which included frontal areas where the N400 was apparent, see Figure 2) were computed over a large temporal window spanning 300 ms – 500 ms, which encompassed the entire negative component, with a peak at 350 – 375 ms (see Figure 3). These values were submitted to a 2 (Recency) X 2 (Stimulus-Set Size) X 2 (Laterality) repeated measures ANOVA. The results showed a significant Set Size X Recency interaction, $F(1, 17) = 5.323, p = .0339$ (Figure 4). In the Closed condition, there was a main effect of recency, $F(1, 17) = 7.208, p = .0157$, in
which the nonrecent negatives showed a larger negative deflection during the time
window. However, there was no such effect in the Open condition, $F(1, 17) = .184, p = .6732$. There were no other significant main effects or interactions.

Figure 5 shows the recent and nonrecent negatives in the closed condition relative
to the positive trials. It is important to examine positive trials for the FN400 to examine
issues of familiarity. If the FN400 is sensitive to familiarity, one would hypothesize that
positive trials would generally have a greater positive deflection than negative trials, and,
further, one might hypothesize that the deflection would be most positive for Recent
positive trials which should be more familiar than Nonrecent positive trials. Figure 6
shows the same for the Open condition. Visual inspection of the waveforms in the closed
condition shows that the negative trials show a greater negativity during this time
window. However, the recent negatives show a smaller N400 effect, such that they
qualitatively look more similar to positive trials. A 2 (Trial Type – Negative vs. Positive)
X 2 (Recency ) X 2 (Laterality) repeated measures ANOVA was calculated for the closed
condition. There was a main effect of Trial Type (negative vs. positive), $F(1, 17) = 13.527, p = .0019$, whereby the positive trials had a greater positive amplitude when
compared to the negative trials. There was also a main effect of laterality, $F(1, 17) = 5.006, p = .0389$, whereby amplitudes for right hemisphere channels were greater (more
positive) than the left hemisphere. There was also a Trial Type X Recency interaction, $F(1, 17) = 5.818, p = .0275$ (see Figure 7). This interaction is explained by a greater
negativity in the Nonrecent trials compared to the Recent trials in the negative condition,
$t(17) = 2.68, p = .016$, but no such difference in the positive condition, $t(17) = .234, p = .817$. 
By comparison, in the Open condition, a 3 (Trial Type – Recent Negative, Non-recent Negative, Positive) X 2 (Laterality – Left, Right) repeated measures ANOVA revealed a main effect of Trial Type, $F(2, 34) = 19.201, p = .0001$ and an interaction Trial Type X Laterality, $F(2, 34) = 4.716, p = .0156$ (Figure 8). A significant main effect of Trial Type was apparent in both the left hemisphere, $F(2, 34) = 6.831, p = .0032$, and the right hemisphere, $F(2, 34) = 30.039, p = .0001$. In the left hemisphere, there was no significant difference between the recent negative and nonrecent negative trials, $t(17) = .738, p = .47$. Both negative trials had a smaller amplitude than the positive trials (recent negatives vs. positives, $t(17) = 3.708, p = .002$; nonrecent negatives vs. positives, $t(17) = 3.227, p = .005$). Likewise, in the right hemisphere, there was no significant difference between the recent negative and nonrecent negative trials, $t(17) = .11, p = .91$, and both the recent negative and nonrecent negative trials had a smaller amplitude than the positive trials (recent negatives vs. positives, $t(17) = 6.70, p < .0001$; nonrecent negatives vs. positives, $t(17) = 7.28, p < .0001$). However, there was a greater positive amplitude for the positive trials on the right when compared to positive trials on the left, $t(17) = 3.594, p = .002$. Thus, the the FN400 was more positive for positive trials over the right hemisphere.

**Late Positive Component.** Another component also appeared to reflect differences between Recent and Non-Recent Negatives (Figure 9). However, this component was localized over parietal regions. Thus, a temporal window of 500ms – 800 ms over the posterior superior regions (Figure 10) was used for this analysis.

A 2 (Recency) X 2 (Set Size) X 2 (Laterality) repeated measures ANOVA was calculated for only the negative trials. There were no main effects of Laterality, Recency
or Set-Size. The two way Recency X Stimulus-Set Size interaction was significant, $F(1, 17) = 11.764, p = .0032$. (Figure 11). The Recency X Stimulus-Set Size interaction reflected a greater positive deflection for NonRecent Negatives relative to Recent Negatives in the Open condition, $F(1, 17) = 11.036, p = .004$, but no difference between the two in the Closed condition, $F(1, 17) = 2.225, p = .1541$.

A one-way repeated measures ANOVA comparing the three trial types was conducted separately for the Open condition. Figure 12 shows both negative and positive trials in the Open condition. This analysis revealed a significant main effect of Condition, $F(2, 34) = 5.619, p = .0078$. Follow-up contrasts revealed that recent negatives had smaller amplitudes than nonrecent negatives, $t(17) = 3.32, p = .004$ and positive trials, $t(17) = 2.5, p = .02$. However the nonrecent negatives did not differ from positive trials, $t(17) = .617, p = .544$.

A 2 (Trial Type – Negative vs. Positive) X 2 (Recency) X 2 (Laterality) repeated measures ANOVA was calculated for the closed condition and revealed Trial Type X Laterality interaction, $F(1, 17) = 4.539, p = .0480$ and an interaction of Trial Type X Recency, $F(1, 17) = 7.409, p = .0145$. The Trial Type X Laterality interaction is explained by a greater positivity for the positive trials than the negative trials only in the left hemisphere, $t(17) = 2.158, p = .046$, but not in the right hemisphere, $t(17) = .065, p = .949$. The Trial Type X Recency interaction is explained by a greater positive amplitude for the Nonrecent trials compared to the Recent trials but only for positive trials, $t(17) = 2.55, p = .021$. However there was no effect of Recency in the negative trials, $t(17) = 1.49, p = .154$. 
Discussion

Experiment 2 was conducted to further examine proactive interference in the Recent Negatives Task using the superior temporal resolution of ERP. In addition to the manipulation of probe recency, a novel manipulation of stimulus-set size was included. This manipulation was motivated by patient ML’s performance in pilot experiments that indicated that using open stimuli sets improved ML’s performance on the nonrecent negative trials. It was hypothesized that using a closed set of stimuli created a great deal of proactive interference even on the nonrecent trials, given that all stimuli had been presented repeatedly. Consequently, the difference in interference between the recent and non-recent negatives was smaller than in the open set condition where the nonrecent negatives had not been presented previously. Thus it was predicted that an ERP component sensitive to interference or inhibition might show a smaller difference in the closed than open condition. As discussed in the introduction, previous data suggested that an N2 effect might reflect inhibition – however, no such effect was found that was sensitive to the recent/non-recent negative manipulation in either the open or closed sets.

Two components were identified that were sensitive to the recent vs. nonrecent negatives manipulation; however, a different interaction between the probe recency and stimulus-set size conditions appeared in each component. For the FN400 (300ms – 500ms window), the Recent vs. Non-Recent differences were seen in the Closed condition. For the Late Positive Component (LPC) (500-800ms window), the Recent v. NonRecent effects were seen in the Open condition. The FN400 effects were observed
over frontal areas, whereas the late positive component differences were observed over parietal channels.

There is surprisingly little published ERP data using probe recognition paradigms tapping short-term memory. Most of the work looking at recognition memory has examined long-term memory paradigms. Data from such paradigms will be reviewed first, followed by a discussion of one study that used a short-term memory recognition memory paradigm.

Much of the research examining old/new effects in the ERP data have emphasized dual-process models of recognition memory, which are comprised of a “familiarity” process (assumed to be an automatic process) and a “recollection” process (which is a more strategic). Recollection is often thought to include conscious retrieval of the “encoding episode”. ERP studies have related an N400 component to familiarity processes, while a later component, the Late Positive Component (LPC), has been related to more conscious recollection in recognition memory (for a review, see Mecklinger, 2000). However, as detailed below, others have related the LPC to memory trace strength, discriminability or decision accuracy/judgment confidence.

Finnigan, Humphreys, Dennis, and Geffen (2002) recorded ERPs while subjects made old/new recognition judgments in a study-test paradigm for presented words (subjects merely indicate whether or not they have seen a presented word in a previous study list). Among the old words, some had been presented either once (‘weak’) or three times (‘strong’) during the study list. Not surprisingly, subjects were more likely to respond ‘old’ for strong words relative to weak words. Finnigan et al. reported an N400 component that was more negative for new words and “weak” words compared to
"strong" words that had been repeated three times. However, unlike the present data, this N400 was distributed over parietal channels. In addition, Finnigan et al. found a late positive complex (LPC) that appeared sensitive to decision accuracy (which the authors relate to decision confidence). The authors found that the LPC was more positive for correct than incorrect responses. However, the topography of this component also differed slightly from the LPC found in the present data, in that the Finnigan et al. LPC was lateralized to the left, whereas there was no significant lateralization of the LPC in Experiment 2.

If one conceptualizes the recent negative probes in Experiment 2 as being "stronger" (given a more recent presentation of a probe in a previous trial) it is not surprising that recent negatives elicit a less negative (i.e. smaller) N400 component compared to nonrecent negatives (which might be considered "new" in the context of the recent negatives task). Similarly, given that RTs are faster and accuracy higher for nonrecent negatives relative to recent negatives, the observation that nonrecent negatives have a more positive LPC is consistent with Finnigan et al.'s suggestion that the LPC indexes decision confidence or decision accuracy. Given less interference on nonrecent negatives, subjects are more confident in their decisions as indicated by shorter RTs, higher accuracy and more positive LPCs. However, the LPC component seems an unlikely candidate to mediate resolution of interference given its late onset. Given that average RTs are ~700-800ms, it would seem that any component mediating inhibition or resolution of interference would necessarily be earlier. Therefore, the N400 might be a more likely component reflecting proactive interference or the consequences of some inhibitory process in the recent negatives task.
With regard to the frontal distribution of the N400 in the present study, Curran (1999) has reported a similar distribution for a component that he termed the FN400 (or Frontal N400). This component peaked at a latency of 388 ms, similar to the latency observed in the Experiment 2 data. The FN400 also demonstrated an ‘old/new’ effect by which new items were more negative than old items. This is consistent with data in Experiment 2 if one conceptualizes Recent Negative trials as being ‘old’ and Non Recent negatives being ‘new’ by comparison. Similar old/new effects have been observed in studies by Rugg (1987), Rugg, Brovedani, and Doyle (1992), and Van Petten, Kutas, Kluender, Mitchiner, and McIsaac (1991). It is important to note that the frontal distribution of the N400 reported in Experiment 2 suggests that this component differs from the parietally distributed N400 commonly reported in sentence comprehension paradigms (as reported in Experiment 3). Whereas the N400 commonly reported in the language comprehension literature is thought to reflect semantic integration processes, the FN400 has been attributed to an “amodal familiarity process” related to a “matching-process” (Curran, 1999; Curran & Dien, 2003). Although the frontal distribution of the N400 in the present data suggests that it is separable from the parietally distributed “semantic N400”, some researchers have continued to assume that the N400 reported in the old/new literature also reflects ease of semantic integration, with repeated words or stimuli more easily integrated than new words.

Studies by Crites, Delgado, Devine and Lozano (2000) examined a probe recognition task that used pictures as stimuli. Each list consisted of four pictures in each memory list followed by a probe. Thus, this task is a short-term memory task. Crites et al. related their findings to the old/new literature described above – specifically, they
found an N400 component like that reported in the continuous recognition and repetition priming experiments. In their experiment, the N400 responded to the "recency" (i.e., serial position) of the list item matching a positive probe in that the N400 was smaller the closer the list item to the probe.

The "familiarity" explanation of the FN400 nicely accommodates data obtained in the Closed condition in Experiment 1. However, one obvious question raised by Experiment 2 is why different effects are observed in the Open and Closed condition. One speculative explanation is that the Closed condition, which promotes more interference by virtue of greater repetition of items, causes subjects to rely more heavily on familiarity processes, which are indexed by the FN400. On the other hand, the Open condition, which only repeats items as necessary to form the recent negative condition, creates less interference and allows subjects use of more explicit recollection, which is reflected by the LPC.

However, one could easily make the opposite prediction regarding the familiarity in the closed condition. That is, the high familiarity of all items in the closed condition (and, subsequently, greater proactive interference in this condition) prevents familiarity from being a useful source of information from which subjects can make a judgment. Instead, the Closed condition should require more strategic recollection and result in recent negative vs. nonrecent negative differences in the LPC, which presumably is an index of recollection. The present data from the closed condition support half of this hypothesis - that is, the FN400 appears to appropriately index the relative familiarity among each of the conditions in the Closed condition. Ranking the four conditions in the closed condition in terms of familiarity, one would expect that recent positives >
nonrecent positives > recent negatives > nonrecent negatives. Visual inspection of the peak amplitudes of the N400 (in terms of amplitude) reveal recent positives > nonrecent positives > recent negatives > nonrecent negatives (see Figure 5). However, this correspondence between the FN400 and familiarity is not apparent in the Open condition. Although positive trials show a smaller N400 relative to negative trials (consistent with a familiarity hypothesis), there appear to be no differences between the recent negative and nonrecent negative trials in the Open condition (see Figure 6).

However, the data from the Closed condition are at odds with the second part of the hypothesis – specifically that the Closed condition should rely more heavily on explicit recollection given the high familiarity among all trials. No significant differences were found between recent negatives and nonrecent negatives in the LPC in the Closed condition (see Figure 9). However, significant recent negative vs. nonrecent negative differences were found in the Open condition. Specifically, the nonrecent negative trials had a larger LPC than the nonrecent negatives. One explanation for this effect is that the nonrecent negatives are subject to less interference affording stronger “recollection”.

This pattern is in the Open condition is also consistent with a judgment accuracy or confidence explanation of the LPC. The nonrecent negative trials have shorter reaction times, greater accuracy and a larger LPC compared to the recent negatives that have longer reaction times, lower accuracy and a smaller LPC. In other words, subjects are generally more confident in their response to nonrecent negative trials, resulting in a larger LPC, and less confident in their response to nonrecent negative trials, resulting in a
smaller LPC. Obviously, further experiments are needed to fully elucidate the functional significance of these ERP components in the Open and Closed conditions of this task.

Patient data may provide further insight into the functional significance of the FN400 and LPC reported in Experiment 2. Patients showing exaggerated interference effects would be useful in identifying which components are related to interference and the resolution of proactive interference. For patients demonstrating exaggerated interference effects in the recent negatives task, it would be interesting to determine the relationship among the N400, LPC and the patients’ behavioral data. For example, patient ML, who at times appears more likely to respond with a positive response to recent negative trials (suggesting that proactive interference is so overwhelming for the recent negatives that ML mistakenly assumes these trials to be positive), it would be interesting to determine if the N400 for incorrect recent negatives was distinguishable from N400s to positive trials. Unfortunately, attempts to analyze ERP data from patient ML have been difficult, owing to the difficulty of analyzing ERP data from a single subject. Several administrations of the experiment would likely be necessary to collect an adequate number of correct trials for ERP analysis, which would undoubtedly introduce problems with the patient’s strategic attempts to modify performance.

Although data from Experiment 2 do not identify a single ERP component that can be unequivocally related to resolution of interference in the recent negatives task, the data do identify potential components that may be examined in future research. These data represent an informative first step in identifying ERP components sensitive to proactive interference in a short-term memory paradigm. The FN400 may prove
especially useful for examining interference processes in short-term memory. For example, in a modification of the recent negatives task, Hamilton (2004) manipulated the phonological and semantic relatedness of probes to previously presented items in the same and previous list and found significant interference for both semantic and phonological relatedness. Given that the behavioral interference effects were of similar magnitude (as measured by reaction times), it is difficult to determine whether these interference effects result from similar mechanisms. Future ERP work may possibly address this question by determining whether the phonological and semantic manipulations have similar effects on the FN400 and LPC components reported in Experiment 2.

EXPERIMENTS 3 AND 4

Experiments 3 and 4 employed variations of a sentence anomaly task to engage semantic short-term memory. This task requires detection of semantic anomalies in phrases in which multiple adjectives appear before or after a noun or multiple nouns appear before or after a verb. Patient data (Martin & Romani, 1994; Martin & He, 2004) indicate that multiple adjectives appearing before a noun place greater demands on semantic short-term memory relative to adjectives appearing after a noun. Similarly, several nouns appearing in the subject position before a verb caused a greater semantic short-term memory load than several nouns appearing in the object position after a verb. Experiment 3 examined this task in the ERP environment. The N400 component was examined as an indirect measure of integration demands in an auditorily presented sentence anomaly paradigm. Experiment 4 uses functional magnetic resonance imaging (fMRI) to examine a modified version of this task using visual presentation. Data from
Experiment 4 are especially useful in corroborating patient data which suggest that the left inferior frontal gyrus is particularly important for semantic short-term memory and, consequently, performance on this task. Data from Experiment 4 also provide a first step in addressing questions raised by data in Experiment 1. Namely, is it possible that a common mechanism is responsible for patient ML’s “verbal inhibition deficit” and his difficulty in maintaining unintegrated semantic representations during sentence comprehension? That is, could failures of inhibition or interference resolution be a causal factor in ML’s difficulty in this comprehension task? While it is possible that the association of susceptibility to interference and difficulty with certain language processing tasks may be explained as a mere consequence of damage to a large region of the left inferior frontal gyrus (encompassing separate regions critical for interference resolution and language processing, respectively), a more intriguing hypothesis is that interference resolution processes are also engaged in some language comprehension tasks. This question will be addressed by comparing brain regions activated in Experiment 4 with previously identified areas thought to be crucial to interference resolution in a number of different paradigms.

Experiment 3 – Before/After Sentence Anomaly Task in ERP

Experiment 3 employs a sentence anomaly task thought to place particular demands on semantic short-term memory. Martin and Romani (1994) used this task to demonstrate that patients with semantic short-term memory deficits have particular difficulty in accurately maintaining semantic representations necessary to detect semantic anomalies present in a sentence. Martin and Romani manipulated short-term memory load by altering the number of adjectives that were to be maintained before eventual
integration with a noun. Specifically, patients received sentences with adjective/noun phrases in which the adjectives either preceded or followed the corresponding noun (i.e., “She saw the green, shining, bright sun, which pleased her.”) or “The sun was shining, bright and green, which pleased her.”). In a set of analogous sentences, nouns either preceded or followed a corresponding verb (e.g. The muffins, cookies, and cakes baked in the oven. vs. They baked muffins, cookies and cakes in the oven). For additional examples of sentences used in Experiment 3, see Table 2 and Table 3.

Martin and Romani (1994) presented a model delineating the levels of short-term memory necessary for sentence comprehension. A schematic for each of the sentence constructions (adjectives before and adjectives after) borrowed from Martin and Romani (1994) appear in Figure 13 and Figure 14. For each type of sentence, representations develop over time as each word is heard. First, the phonological form is derived followed by semantic representations. On a word-by-word basis, the syntactic structure is formulated. Importantly, message-level propositions are derived as soon as representations are available for linking the semantic representations for individual words together. For sentence comprehension, the proposition level is critical. Martin and Romani (1994) argued that these levels are differentially subject to decay or interference - it is thought the phonological and semantic levels are lost (or interfered with) rapidly, while the proposition level is the most persistent. Moreover, data from brain-damaged patients indicate these phonological and semantic levels are dissociable.

Patients with phonological short-term memory deficits rapidly lose access to phonological representations, making verbatim sentence repetition difficult. Instead, these patients have a tendency to paraphrase sentences. However, because these same
patients do adequately maintain semantic representations, they possess very good sentence comprehension. In contrast, patients with semantic short-term memory deficits fail to maintain semantic representations, which presumably results in poor comprehension of sentences that demand retention of word meanings for some duration before a proposition can be formulated.

This hypothesis was confirmed by Martin and Romani (1994) and Martin and He (2004). These studies reported that patients with semantic short-term memory deficits had particular difficulty detecting sentence anomalies when multiple adjectives had to be maintained before subsequent integration with the noun they modify or multiple nouns had to be maintained until the verb was processed to determine their role with respect to the verb. By contrast, semantic short-term memory patients had less difficulty detecting anomalies when adjectives appeared after a noun or nouns after the verb. Multiple adjectives appearing after a noun allow for immediate integration of the adjectives and nouns into a higher-order proposition. Multiple nouns appearing after a verb can be immediately assigned their role with respect to the verb as they are processed. Thus, because of the delayed integration in the “before” condition and immediate integration in the “after condition, Martin and Romani reasoned that the before condition placed a greater demand on semantic short-term memory than did the after condition.

In addition to the Before vs. After manipulation, Martin and Romani (1994) and Martin and He (2004) manipulated the number of adjectives that appeared before or after the noun (or, in the noun–verb phrases, the number of nouns appearing before a verb). Specifically, one, two or three adjectives (or nouns) appeared before or after the noun (or verb). Patient data from Martin and Romani (1994) and Martin and He (2004)
demonstrating the different patterns of performance on this task appear in Figure 15. Figure 15a presents reaction time data from healthy control subjects which demonstrates a significant Before/After X Distance interaction, whereby reaction times increase as a function of the number adjectives to be maintained, but only in the before condition (reaction times for anomaly decisions were measured from the onset of the anomalous word and from the onset of the corresponding word in the sensible sentences). Figure 15b and 15c present accuracy data for two patients with semantic short-term memory deficits. Figure 15d presents data from patient EA, a patient with a phonological short-term memory deficit. Apparent in patient AB and ML’s data is the exaggerated effect of distance in the before condition for the semantic short-term memory deficits. While these patients have good accuracy for one adjective (near 90% for both patients), their accuracy declines dramatically when two or three adjectives must be integrated in the before condition. This effect is not apparent in the after condition. The exaggerated effect of distance in the before condition is not apparent in patient EA, who has a phonological short-term memory deficit.

It is also important to note that when only one adjective is required to be integrated in either the before or after condition, reaction times among healthy control subjects is equivalent. Thus, it is difficult to attribute any differences in the before vs. after condition to differences in the degree of anomaly inherent in the stimuli for the before and after conditions.

Experiment 3 used this paradigm in the ERP environment to provide converging evidence that presenting multiple prenominal adjectives places unique demands on short-term memory. The N400 component, an ERP component thought to index processes
important in semantic integration, is used as an indirect measure of short-term memory load. A brief review of the N400 literature follows.

A number of ERP components have been associated with language processes. One of the first language-related components to be identified was the N400, first reported by Kutas and Hillyard (1980). Subjects in this study were asked to read sentences silently as they were presented serially, one word at a time. Twenty-five percent of the sentences ended with semantically incongruous words. Relative to sensible words, the incongruous words elicited a large negative deflection in the waveform peaking around 400ms. Moreover, words that were more incongruous elicited larger N400s than less incongruous words. Thus, many researchers have concluded that the N400 appears to be sensitive to the difficulty of semantic integration (assuming that more incongruous words are more difficult to integrate than less incongruous or sensible words). Moreover, since it was first reported in 1980, it is now known that most content words elicit an N400 (Kutas & Van Petten, 1994). The amplitude of this component is determined by some aspect of the processes necessary for integration of words into a higher-order semantic interpretation (or proposition) (van Berkum, Hagoort & Brown, 1999). However, the semantic integration hypothesis is complicated by the finding that pronounceable nonwords also elicit an N400. If the N400 indexed access to only the meaning of a word, one might predict an absence of N400s for nonwords (Kutas & Van Petten, 1994). Also, at least one theory has proposed that instead of indexing semantic integration, the N400 reflects the inhibition of semantic representations irrelevant to context (Debruille, 1998).

It is important to again note that Experiment 3 uses the N400 as an indirect measure of semantic short-term memory processes. In other words, Experiment 3 does
not assume that the N400 is an index of short-term maintenance of semantic
representations per se. Instead, it is assumed that the increased load associated with
maintaining three unintegrated adjectives presented before a noun will make semantic
anomalies less salient, resulting in smaller N400s in the before condition relative to the
after condition. Moreover, it is important to note that this hypothesis is somewhat
inconsistent with a conceptualization of the N400 as reflecting difficulty of semantic
integration. Many have assumed that the N400 is responsive to difficulty in integrating a
word with previous context (Brown & Hagoort, 1999; Kutas & Van Petten, 1994). For
example, a highly anomalous word is assumed to be more difficult to integrate than a less
anomalous word, thus the former elicits a larger N400 effect. From this perspective, one
might predict that maintenance of multiple prenominal adjectives or multiple nouns in the
subject noun phrase would result in larger N400s during semantic integration relative to
sentences in which adjectives followed the noun or the nouns followed the verb. The
present experiment makes the opposite prediction – namely, that N400s will be smaller in
the before condition, based on the assumption that the before condition will result in less
salient semantic anomalies. While this prediction may at first seem at odds with much of
the N400 literature, it is important to note that previous studies have not examined the
role of short-term memory load in modulating the N400.

In the materials for Experiment 3, the semantic congruity or incongruity was
matched in the before and after conditions. Thus, integration difficulty in terms of
semantic congruity did not differ between conditions. Instead, the two conditions were
hypothesized to differ in integration difficulty because of the need to maintain or retrieve
degraded semantic representations in the “before” condition but not in the “after” condition.

A prior study by Yang, Martin, and Potts (2002) employed the same “Before/After” paradigm with visually presented stimuli. (In fact, their stimuli are the same stimuli recorded for auditory presentation in Experiment 3). In the Yang et al. study, each word was presented for 200ms with a 400 ms inter-word interval. Yang et al. reasoned as outlined above – that is, that semantic anomalies in the ‘Before’ condition should elicit a smaller N400 effect relative to anomalies in the ‘After’ condition. Yang et al. found an interaction such that the difference in the N400 between anomalous and sensible sentences was greater in the after condition than in the before condition (see Figure 16).

Experiment 3 replicates the Yang et al. study, but uses auditorily presented stimuli instead of visually presented sentences. The long inter-word intervals in the Yang et al. study are typical of ERP experiments on sentence comprehension as researchers aim at having the ERP for one word diminish substantially before the next is presented. However, presenting stimuli at 600 ms per word results in unnaturally slow reading time. During normal reading, the average word fixation is about 250 ms per word, with the time varying depending on word length, frequency, and grammatical class. Auditory presentation provides for more natural processing of sentences. However, using spoken auditory stimuli provides less experimental control over the presentation of stimuli as the words run together and thus the ERP for one word will blend into the next. However, some previous ERP studies have successfully used auditory presentation in studying
sentence comprehension (see Friederici, 2002, for a review). Experiment 3 was an attempt to replicate the results of the Yang et al. visual study using auditory presentation.

For Experiment 3, the typical N400 effect elicited by anomalous words was expected to be smaller in the before condition compared to the after condition. This result is predicted on the grounds that the before condition places greater demands on short-term memory (specifically semantic short-term memory), thus making anomalous words less salient.

Method

Subjects. Twenty-five undergraduate students at Rice University (16 females; age range = 18 – 21; mean age= 19.5) participated for course credit. Subjects were right-handed and native speakers of English with normal or corrected-to-normal vision. Participation in the experiment required no current psychiatric diagnosis or history of neurological illness or injury.

EEG data acquisition and signal processing. The electroencephalogram (EEG) was recorded using 128 channel Electrical Geodesics system (Electrical Geodesics Inc, Eugene). EEG data were acquired continuously referenced to the vertex with .1 - 100 Hz analog filtering and digitized at 250 Hz. The EGI Geodesic Sensor Net is a lightweight elastic thread structure containing plastic pedestals. Each pedestal contains a silver/silver chloride electrode housed in a synthetic sponge. The sponges are soaked in a saline solution to render them conductive. Application of all 128 channels takes approximately 15 minutes.

An Apple Macintosh 266 MHz PPC G3 computer with Mac OS 9.04 controlled the data acquisition and stimulus generation using EGIS software (Electrical Geodesics
Inc., Eugene). Visual stimuli were presented on an Apple 15" flat-panel active matrix Studio display to reduce 60 - 75 Hz monitor refresh electrical noise associated with CRT displays.

In all experiments, subjects were seated in an adjustable chair with their chin in a chinrest. The chinrest was placed so that subject's eyes were 50 cm. from the center of the flat-panel screen. The chair was adjusted for comfort. Subjects were instructed to remain as still as possible, with their eyes on the fixation mark, throughout the block. Subjects were requested to refrain from blinking as much as possible while the stimuli were presented. Breaks were provided every 3 - 4 minutes so that subjects could rest their eyes.

**Materials.** The EEG was recorded as the subjects listened to each sentence. Each sentence was presented auditorily through two speakers positioned in front of the subject. All stimuli were recorded by a male voice. Each sentence was preceded by a fixation mark that lasted for 800 ms. A total of 60 experimental sentences were constructed for the experiment. Half of the experimental sentences were Adjective-Noun/Noun-Adjective constructions. The other half were Noun-Verb/Verb-Noun constructions. Thus, for all of the experimental sentences, there were either three adjectives preceding or following a noun or three nouns preceding or following a verb. Four versions of each experimental sentence were constructed from the combination of two factors: Before vs. After (eg. delayed vs. immediate integration) and Plausibility (sensible vs. anomalous). Each anomalous sentence was identical to its sensible counterpart, with the exception of the substitution of an anomalous word. Because of the need to have a large number of stimuli in each condition, the distance manipulation used in Martin and Romani (1994) and Martin and He (2004) was not included in Experiment 3.
In addition, the critical words appeared at either the 7\textsuperscript{th} or the 10\textsuperscript{th} positions in the sentences. That is, the position of critical words determining whether a sentence was sensible or anomalous was the same for the before and after conditions. This is important given that the N400 has been reported to vary as a function of the position of a word within a sentence (see Kutas and Van Petten, 1994). Specifically, N400s are typically larger for words appearing earlier in a sentence. The typical explanation for such effects is that earlier words are bereft of context and thus are more difficult to integrate, resulting in larger N400 effects. Examples of representative sentences in Experiment 3 are presented in Table 4.

A set of 300 filler sentences was also included to obfuscate the manipulations of interest in the experiment. Approximately 60 filler sentences were anomalous. The anomalous filler sentences introduced both semantic violations and grammatical violations. In addition, the position in which the anomalous words appeared within the filler sentences was varied and counterbalanced. Specifically, anomalous words appeared at early, middle or late positions within the filler sentences. For examples of filler sentences, see Table 5.

A version of each sentence appeared once in each of the four conditions (Sensible vs. Anomalous, Before vs. After). Each block was comprised of 60 experimental sentences and 75 filler sentences (15 anomalous and 60 sensible). Four blocks of 135 sentences were presented in a single session. Thus, the sentences were counterbalanced such that they only appeared twice during each session, but always appeared in different conditions. This minimized the possibility that subjects could keep track of which versions of the sentences they had encountered. Each subject participated in two experimental
sessions separated by no less than one week. The order of block presentation was counterbalanced across subjects, such that each subject received all four versions of the experimental sentences. The inter-trial interval (ITI) was 1000 ms.

Subjects were not required to make a key press to indicate a “sensible” or “anomalous” judgment. Instead, a “comprehension challenge” was presented every 5 to 10 trials. The “comprehension challenge” asked subjects to embellish the sentence they had just heard. This allowed the experimenter to make sure that the subjects were adequately engaged in the task. If the experimenter considered the subject’s response to be inadequate, the subject was encouraged to attend to sentences more closely. The comprehension challenge was chosen over a sensible/ anomalous response in order to minimize any strategic processes that might complicate interpretation of the ERP data. However, because of this design, no behavioral data are available for Experiment 3.

Data Processing. EEG data were segmented off-line into 1400 ms epochs spanning 400 ms pre-stimulus to 1000 ms post-stimulus. Data were digitally screened for artifact (eye blinks or movements, subject movement, or transient electronic artifact) and contaminated trials were eliminated. Remaining data were sorted by condition and averaged to create the ERPs. Averaged ERP data were digitally filtered at 8 Hz lowpass to remove residual high-frequency noise. An 8 Hz filter was necessary to remove higher frequency noise presumably attributable to auditory presentation of words. Data were then baseline corrected over the 200 ms pre-stimulus period, and re-referenced into an average reference frame to remove topographic bias due to choice of reference site. The subject-averaged ERPs were averaged together to produce the grand average waveforms. Statistical analyses were performed on the subject-averaged ERPs with the subject
averages serving as observations. The waveform plots were generated from grand-averaged data averaged across channels of interest.

As recommended by Dien and Santuzzi (2005), channels were averaged by region, based on the following dimensions – left/right, anterior/posterior, and superior/inferior. Such a strategy yields eight different regions – anterior superior, anterior inferior, posterior superior, and posterior inferior on the left and right hemisphere. Although the regional approach sacrifices the fine resolution provided by high-density electrode nets, it addresses potential objections to the large degrees of freedom that multiple electrodes present. Given the typical parietal distribution of the N400 in language tasks, Experiment 3 examined channels over the posterior superior channel groupings.

A large temporal window spanning 200ms to 600ms was used for the analysis of the N400. Previous research has indicated that auditory presentation of sentences elicits N400s with an earlier onset relative to visually presented sentences (Brown & Hagoort, 1999). Given that the speech signal is continuous, it has been proposed that early ERP components are “smeared” for auditorily presented sentences, giving auditorily evoked N400s a different morphology when compared to visually presented stimuli.

Results

Grand-averaged data were submitted to a 2 (Before v. After) X 2 (Sensible v. Anomalous) X 2 Laterality (Left posterior superior channels v. Right posterior superior channels) repeated measures ANOVA. Given the typical parietal distribution of N400 effects in language paradigms, the left and right posterior superior channels were chosen. Channels along the mid-line were excluded from the analysis to allow tests of laterality
effects. Analyses were conducted using both latency and amplitude as dependent measures.

**Latency Analysis**

Visual inspection of waveforms suggested that sensible and anomalous trials might differ in latency of the peak N400 component (see Figure 17a). Thus a latency analysis was conducted using a window that included both components. Latency analyses use the latency of the peak amplitude in each condition as the dependent variable (as opposed to average amplitude over a specified time window). A large temporal window of 200ms - 600ms was chosen to capture most of the N400 component, beginning shortly after the onset of the negative deflections.

Grand-averaged data were submitted to a 2 (Before v. After) X 2 (Sensible v. Anomalous) X Laterality (Left posterior superior channels v. Right posterior superior channels) repeated measures ANOVA.

With latency as a dependent measure, there was a trend toward a main effect of Before/After \( F(1, 24) = 3.704, p = .0662 \). There was no main effect of Sensible/Anomalous, \( F(1, 24) = .874, p = .3590 \). There was a main effect of Laterality, \( F(1, 24) = 5.872, p = .0233 \). The After/Before X Sensibility interaction was not significant, \( F(1,24) = 2.430, p = .1321 \).

**Amplitude Analysis**

A 2 (Before v. After) X 2 (Sensible v. Anomalous) X 2 (Laterality -Left v. Right) repeated measures ANOVA was performed on a 200ms – 600ms window (see Figure 17). ANOVA revealed a main effect of Sensibility \( F(1, 24) = 7.78, p = .0102 \) (Figure 18). There was also a Before/After X Laterality interaction \( F(1, 24) = 4.761, p = .0391 \).
(Figure 19). The Before/After X Sensibility interaction $F(1,22) = .218, p = .6445$ was not significant.

Given the significant Before/After X Laterality interaction, further analysis examined the left and right regions separately. For the left hemisphere, there was a main effect of Before/After, $F(1,24) = 5.372, p = .0293$. There was also a main effect of Sensible/Anomalous, $F(1,24) = 6.197, p = .0201$. The Before/After X Sensibility interaction was not significant, $F(1,24) = .011, p = .9181$. In the right hemisphere, there was a trend toward a main effect of Sensibility, $F(1,24) = 3.359, p = .0793$, but no main effect of Before/After, $F(1,24) = .875, p = .3588$, nor any interaction between Before/After and Sensibility, $F(1,24) = .674, p = .4196$. Thus, the Before/After X Laterality interaction is explained by a less negative N400 in the Before than the After condition over the left hemisphere but no difference over the right hemisphere.

Discussion

Experiment 3 provides converging electrophysiological evidence that presenting multiple adjectives before a noun (or multiple nouns before a verb) places greater demands on integration processes during sentence comprehension. This effect was evidenced by a smaller N400 component in the before condition relative to the after conditions.

These data provide a new finding regarding the relationship between short-term memory demands, semantic integration and the N400 in auditory sentence comprehension. Brown and Hagoort (1999) noted that the N400 is likely related to a meaning integration process that integrates activated meanings into a message-level representation. A number of effects are thought to support the hypothesis that integration
difficulty determines the magnitude of the N400 (see Kutas and Van Petten, 1994 and Brown and Hagoort, 1999 for reviews). First, when lists of words are presented, N400s are smaller when successive words are semantically related relative to semantically unrelated. It is assumed that subjects always attempt to extract meaning from words and that semantically related words afford an easier extraction of meaning. Secondly, N400s for words appearing early in a sentence are typically larger than N400s for words appearing later in a sentence. Presumably, later words benefit from the preceding context, thus making integration less difficult and eliciting smaller N400s (Kutas and Van Petten, 1994). However, as noted above, the “semantic integration hypothesis” does not accommodate all of the N400 data – pronounceable nonwords also elicit N400s. Moreover, Rugg (1984) observed larger N400s for nonrhyming words that appear in a task that required rhyme judgments.

The apparent abundance of evidence supporting the “semantic integration hypothesis” prompted Brown and Hagoort (1999) to conclude, “the easier the integration process is, the smaller the amplitude of the N400” (p. 223). The present data are clearly at odds with such a position, as behavioral data strongly indicate that integration is more difficult in the before condition compared to the after condition. Specifically, reaction times are significantly longer and accuracy poorer in the before condition relative to the after condition. Nevertheless, the greater integration difficulty in the before condition was associated with smaller N400s in Experiment 3.

However, by making a subtle distinction in the semantic integration hypothesis, it is possible to reconcile the observation of smaller N400s with the increasing integration difficulty. In fact, the smaller N400 in the before condition is very consistent with
previous conceptualizations of the N400 if one assumes that the before condition is an impediment to semantic integration and context formation. Given that the before condition prevents integration of prenominal adjectives (or integration of nouns with verbs), a weaker context is formed. Given a weak formation of context in the before condition (i.e., higher level propositions used in comprehension of the sentence are not formed), this condition results in smaller N400s. Thus, N400s are smaller in both sensible and anomalous sentences in the before condition, given that integration and context has been only weakly established. The weak context in the before condition results in smaller N400s in both the sensible and anomalous trials. A similar reasoning has been used to explain N400 anomalies in schizophrenic patients. Some studies have reported smaller N400 effects in sentence anomaly tasks among schizophrenics (although others have paradoxically reported larger N400s). A similar failure to adequately form and maintain context during language comprehension has been used to explain abnormal N400s in a number of tasks using schizophrenic patients (see Kumar and Debruille, 2004, for a brief review).

Although the present data do provide further support for differences between the before and after conditions, they do differ in some regards from the data obtained for the same paradigm using visual presentation of sentences. The absence of a Before/After X Sensibility interaction in the auditorily presented experiment is the most conspicuous inconsistency. While most of the N400 research has been conducted with visual presentation, the N400 is not modality dependent. However, the presentation modality has been reported to alter the latency of the N400 (Brown and Hagoort, 1999). Moreover, earlier components of the ERP are often not observed when sentences are
presented auditorily. Brown and Hagoort (1999) have proposed that this is attributable to the "continuous physical stimulation of the speech signal which gives rise to a series of temporally overlapping stimulus components" (p. 222). Brown and Hagoort suggest that this, combined with the "refractory period" of auditorily-evoked ERPs, produces a "smearing" of earlier temporal components. Similarly, others have noted that the temporal differences of the auditorily evoked N400 are related to the fact that most words are identified before their acoustic offset has been completed and that a number of lexical candidates are available given incomplete acoustic information at any given moment. These factors conspire to "smear" the auditorily evoked N400 when averaged over different words and across multiple subjects. This increased variability may explain the absence of the Before/After X Sensibility interaction in the present data.

Finally, the left-lateralized N400 reported in Experiment 3 is also of interest, given that the N400 is typically thought to be lateralized to the right. In a review of the N400 literature, Kutas and Van Petten (1994) have reported that the auditory N400 is typically more symmetrically distributed or lateralized to the left hemisphere. Thus, the present data are consistent with previous studies of the auditory N400.

EXPERIMENT 4: fMRI OF BEFORE/ AFTER ANOMALY PARADIGM

Experiment 4 uses a variation of the Before/After sentence anomaly paradigm described above in Experiment 3. The fMRI environment necessitated several modifications to the paradigm. Whereas ERP allows a great amount of flexibility in the length and timing of sentence presentation (given that an epoch of interest can be conveniently removed off-line), fMRI demands much more stringent control of task parameters. Therefore, the paradigm was modified such that entire sentences were not
presented. Instead, only the four critical words from the adjective-noun or noun-verb phrase were presented. For example, instead of presenting the sentence “She saw the green, shining, bright sun, which pleased her.” Experiment 4 presented only “green, shining, bright sun”. Subjects were asked to make a “sensible” v. “anomalous” judgment (by making a key press) based on whether the four words “go together”. Experiment 4 used adjective-noun phrases as well as noun-verb phrases in “Before” and “After” conditions.

A number of neuroimaging studies have examined the short-term maintenance of semantic and phonological representations. These studies are briefly reviewed below. However, Experiment 4 used a paradigm not employed in previous studies. This task was chosen for its demonstrated ability to engage semantic short-term memory. Performance on this paradigm discriminates between semantic short-term memory patients and phonological short-term memory patients (Martin & Romani, 1994, as described previously in Experiment 3). In addition, because this task is thought to minimize phonological contributions to short-term memory it maybe especially useful in engaging semantic short-term memory capacities.

The neural dissociation of semantic and phonological short-term memory, first reported in patients by Martin and colleagues (Martin & He, 2004; Martin & Romani, 1994; Martin et al., 1994) has been tested in several studies using functional neuroimaging. These studies have produced somewhat inconsistent results, with some studies showing dissociable neural substrates (Crosson et al., 1999; Martin et al., 2003; McDermott et al., 2003; Poldrack et al., 1999; Shivde & Thompson-Schill, 2003), while others have shown no such dissociation (Barde & Thompson-Schill, 2002). However,
these experiments have used very different experimental paradigms. Experiment 4 uses a paradigm demonstrated to engage semantic short-term memory in behavioral studies both with healthy subjects and patients with neurological damage. A brief review of the existing neuroimaging studies relevant to dissociable phonological and semantic short-term memory capacities follows. Although some of the studies did not explicitly test short-term memory, one could assume that most of the tasks make at least some demand of short-term memory. For example, McDermott, Petersen, Watson, and Ojemann (2003) used a task in which subjects were “asked to attend to the relations” among a lists of semantically- or phonologically-related words. Presumably, short-term memory is required to attend to relations among a serial list of visually presented words.

Poldrack, Wagner, Prull, Desmond, Glover and Gabrieli (1999) used fMRI to demonstrate distinct areas of prefrontal activation associated with phonological and semantic “processing”. In the semantic task, subjects were required to make a “concrete v. abstract” judgment on visually presented words. The phonological task required subjects to count the number of syllables in a word or nonword. Thus, it could be argued that this study did not actually tap short-term memory. However, this study has been influential in subsequent fMRI studies of semantic and phonological short-term memory and is thus included here. Poldrack et al. (1999) reported that the anterior region of the ventrolateral prefrontal cortex (BA 47/45) was associated with semantic processing, while a more posterior region of the prefrontal cortex (BA 44) was associated with phonological processing. Note that such results are not entirely consistent with patient data reported by Martin and colleagues. Patient data suggest that phonological short-term memory is related to the function of the left inferior parietal lobe, while semantic
short-term memory is related to left inferior frontal areas. However, it is important to note that any short-term memory task might engage phonological rehearsal, which may very well be related to activity of left frontal regions involved in motor planning and articulation.

Collette, Majerus, Dabe, Degueldre, Delfiore, Luxen and Salmon (2001), used positron emission tomography to examine the differences in neural activation between short-term memory for words compared to nonwords. Nonwords were used because, unlike words, they have no semantic features associated with them. Consequently, nonwords are thought to rely more heavily on phonological short-term memory. In this study, short-term memory for words (presumably tapping semantic short-term memory more than non-words) was associated with activation in the middle temporal gyrus (BA 21) and the tempo-parietal junction (BA 39). Again, these data are at variance with data from patient studies, which have reported that areas in the inferior parietal lobe, such as BA 39, are more likely to support phonological short-term memory.

Crosson et al. (1999) also used fMRI to examine separable semantic and phonological components operating in short-term memory. Although Crosson et al. (1999) report many areas of differing activation, increased activity in the inferior frontal gyrus, anterior to Broca's area (near BA 47), was reported during tasks requiring semantic short-term memory. For tasks requiring phonological short-term memory, more posterior areas were activated including areas around the inferior temporal-occipital junction (BA 37). Phonological short-term memory tasks also activated prefrontal areas near BA 44 and 46, but did not overlap with activity observed in the semantic short-term memory tasks.
Martin, Wu, Jackson, Freedman, and Lesch (2002) also reported a fMRI study examining phonological and semantic short-term memory tasks. In this study, large inferior-frontal, mid-frontal and left parietal activations were implicated in short-term memory load (i.e., when the task required retention of more items) in both semantic and phonological short-term memory tasks. The phonological short-term memory task (in this case, a rhyme probe task) was associated with activation in left inferior parietal lobe (BA 40) and more posterior areas of the frontal lobe. The semantic short-term memory task also elicited frontal activation, but this activation was more anterior to the activation related to phonological short-term memory (although this anterior activation was only a statistical trend).

In contrast, Barde and Thompson-Schill (2002) found no differences between phonological and semantic short-term memory tasks. These authors argue that short-term memory is organized by “process” (in other words, the type of operation being performed on given information) with no distinctions determined by “type” of material to be remembered (e.g. verbal or non-verbal information or semantic v. phonological information). These authors argue that their data strongly implicate inferior frontal areas in short-term memory, regardless of the type of information that must maintained.

As noted above, McDermott, Petersen, Watson, and Ojemann (2003) have examined differences between “attention to semantic and phonological relations” among lists of words using fMRI. In this study, subjects were simply asked to think about the relations among words in a list of 16 items. The words either rhymed (beep, weep) or were semantically related (bed, rest). Although this task is not a short-term memory task per se, it is assumed that the task does require some degree of support from short-term
memory. Performance of the task on the semantically related lists was associated with activation in the left inferior gyrus in BA 47 and BA 44/45, as well as activation in the left superior/middle temporal cortex in BA 22/21. Performance on the phonological lists was correlated with activation in BA 6/44, posterior to the regions activated in semantic lists, as well as activation in BA 40 and precuneus (BA 7). These data are more consistent with an anterior/posterior distinction between semantic and phonological retention in short-term memory as proposed by Martin and colleagues. However, it is again important to note that the phonological task activated posterior regions of the inferior frontal lobe.

In a study using methodology similar to that of McDermott et al. (2003), Shivde and Thompson-Schill (2003) found activation in anterior regions of the left inferior frontal gyrus and the left middle temporal gyrus for semantic lists and the left superior parietal lobe for phonological information. The Shivde and Thompson-Schill task merely required subjects to indicate whether or not two words separated by a 10 second delay either shared a vowel sound or were semantically related. Again, this task does not place a particularly demanding load on short-term memory, although it is presumed that the task does engage short-term memory mechanisms necessary for maintenance across the 10-second delay interval. Of further interest, Shivde and Thompson-Schill suggest that their task may be a relatively pure task of semantic maintenance, given that there were no activations in brain areas typically associated with phonological rehearsal (ie. BA 44 and BA 6). The authors suggest that their task’s demands (maintaining only one item across a delay period) may have been well below the typical short-term memory span, and thus did not engage rehearsal processes necessary for other short-term memory tasks. Also of
note, the Shivde and Thompson-Schill (2004) study seems to be inconsistent with the group's earlier attempt to dissociate phonological and semantic maintenance in short-term memory (Thompson-Schill and Barde, 2002). However, the tasks were very different, with the Barde and Thompson-Schill (2002) task requiring the ordering of concrete nouns according to size. Thus, the latter experiment likely engaged executive processes, thereby obfuscating activations related purely to semantic maintenance. The Shivde and Thompson-Schill (2004) paradigm is presumably a purer means of engaging such maintenance processes.

The fMRI literature motivates several predictions regarding brain areas to be activated by the before/after anomaly task. First, it is predicted that areas of the left inferior frontal gyrus will show greater activation in the before condition relative to the after condition. More specifically, given that this task is thought to place particular demands on semantic short-term memory, it is predicted that the before condition will elicit greater activation in areas reported in the previously published studies of short-term semantic maintenance. Specifically, the before condition should engage more anterior areas of the left inferior frontal lobe.

Of further interest are areas of the left inferior parietal areas, which are implicated in phonological short-term memory. Moreover, given a model of semantic processing in which control processes in the left frontal areas act upon semantic representations localized to the temporal lobe (see Martin, 2003), another prediction is that the temporal areas will show no differences between the before and after condition. This assumes that the same amount of semantic processing occurs in both the Before and After conditions,
but that the frontal control processes are more necessary for maintenance of semantic representations in the Before condition.

While there is an abundance of research examining semantic anomaly tasks using ERP, fewer studies have examined anomaly tasks in fMRI. Friederici, Ruschemeyer, Hahne and Fiebach (2003) examined both semantic and syntactic anomalies using fMRI. They reported that semantic anomalies generated activation bilaterally in the insula and superior temporal areas. These temporal activations are consistent with intracranial recordings that have localized the neural generators of the N400 to temporal areas (McCarthy, Nobre, Bentin, and Spencer, 1995). Consistent with these studies, Kiehl, Laurens and Liddle (2002) found bilateral anterior temporal activation for semantically incongruent words. This activation also extended into the left inferior frontal lobe. Thus, it was hypothesized that any sensible vs. anomalous differences would be observed in temporal and bilateral inferior frontal areas. However, it is important to note that the previous studies were designed to identify brain areas responsive to semantic violations and not concerned with semantic short-term memory.

Method

Subjects. Thirteen subjects (10 female) were recruited from the Rice University community. Ages ranged from 18-34 (mean = 25 years). Subjects reported no history of psychiatric illness, brain injury or brain trauma. All subjects were right-handed. Scanning was conducted at Baylor College of Medicine’s Human Neuroimaging Laboratory. One subject was dropped from analysis due to excessive head motion in the scanner.
Design. A within-subjects 2 (Before v. After) X 2 (Sensibility) design was used for Experiment 4. For adjective-noun blocks, there were 20 trials in each of the 4 conditions (anomalous-before, anomalous-after, sensible-before, sensible-after). In the noun-verb blocks, there were 16 trials per condition (anomalous-before, anomalous-after, sensible-before, sensible-after). For the noun-verb stimuli, it was necessary to use verbs that could accommodate the same nouns in the Before condition (cakes, pies, cookies, baked) and in the After condition (baked, cakes, cookies, pies).

An event-related design with varying inter-trial intervals (ITIs) of 6s, 8s, 10s, and 12s was used. Each subject participated in a total of 4 runs, where 1 run consisted of a before block and an after block. There were 2 runs consisting of adjective-noun stimuli and 2 runs consisting of noun-verb stimuli. The presentation order was counterbalanced across subjects.

Materials. Stimuli consisted of 80 adjective-noun stimuli and 64 noun-verb stimuli. In half of the trials, the anomaly always appeared in the same serial position, farthest away from the noun (i.e., rusty, old, red, swimsuit). In the other half of the trials, the anomaly appeared closest to the noun (i.e. adorable, cute, fluffy, turtle), or in the middle of the other two adjectives (i.e., adorable, fluffy, cute, turtle). The “serial position” of the anomaly was manipulated to prevent subjects from merely attending to the first and last words to detect an anomaly. Given that there is no a priori reason to anticipate behavioral or hemodynamic effects resulting from differences among trial types, all trials were included in the behavioral and imaging analyses. The before and after conditions were collapsed over the adjective/noun and noun/verb phrases to increase statistical power.
Procedure. First, a screen appeared informing subjects which of the four conditions they would receive (Adjective/Noun- Before, Adjective/Noun-After, Noun/Verb- Before, Noun/Verb-After). Stimuli were presented using rapid serial visual presentation; words appeared serially on a screen for 500 ms each. Subjects then had to decide whether the adjectives "made sense" with the noun (sensible) or not (anomalous) by pressing a button box with their left hand. Subjects followed the same procedure for noun-verb blocks.

Data Acquisition. MRI scanning was performed on a 3T, head-only, Siemens Allegra MRI scanner (software version Syngo MR 2002B, Erlangen, Germany). An echo planar imaging (epi) sequence was used with an echo time of 40ms, a repetition time of 2000ms and a 90° flip angle. Twenty-six 4 mm axial slices were collected per volume, covering the entire brain for most subjects. The field of view was 220mm and the acquisition matrix was 64 X 64, resulting in a 3.44mm in-plane resolution.

In the adjective-noun blocks, there were 304 volumes per run; in the noun-verb blocks 244 volumes were collected per run. One high resolution structural scan consisting of 194 1 mm slices was acquired at the beginning of each scanning session. The order of presentation of the adjective-noun and noun-verb blocks was counterbalanced across subjects. All stimuli were presented with E-Prime. Subjects viewed the stimuli using a mirror mounted on the head coil which allowed a view of a projection screen which was located behind the scanner.

Data were viewed and analyzed using the AFNI software package (Cox, 1996). Two-dimensional motion correction was achieved by aligning images on a slice-by-slice basis for each input 3d time dataset that resulted from each block, relative to a pre-
specified base image in that dataset. A 4 mm Gaussian blur was used to smooth the data, using AFNI’s 3dvolreg function.

The data were then normalized in order to calculate the percent signal change and a deconvolution analysis, using AFNI’s 3dDeconvolve program, was used to estimate the hemodynamic response for each condition at each voxel. The intensity values derived from this analysis were used as the dependent variables in the ANOVAs and contrasts. The ANOVA was computed with AFNI’s 3dANOVA program.

The comparison of task vs. baseline, thresholded at $t = 11.33$, $p = 1.22$, was used to identify twelve peak maxima that reliably showed greater activations in the experimental task compared to baseline. A sphere with a 5mm radius (from the peak maxima) was used to create regions of interest (ROI). Within these regions, all non-zero voxels were averaged at each of seven time points for each subject. These resulting values were used as the dependent measure in a 2 (before vs. after) X 2 (sensible vs. anomalous) X 7 (time points) repeated measures ANOVA. Talairach coordinates for all twelve ROIs appear in Table 6. Figure 20 presents all 12 ROIs, numbered in order of magnitude of the respective $t$ value associated with their maxima.

Results

Behavioral Data. Behavioral data were acquired from 10 subjects while they performed the task in the scanner. Data from two other subjects was not available due to experimenter error. Subjects took 100ms longer to respond to the before condition (869 ms) compared to the after condition (769 ms), $t(9) = 3.159$, $p = .01$. Eight of ten subjects showed longer reaction times for the before than the after condition in the expected direction.
Before/After Contrasts. As described above, ROIs were created with a 5mm radius surrounding maxima that exceeded $t = 11.32$. Of the twelve maxima identified, 10 were located in the left frontal areas. Statistics for ROIs yielding significant main effects of Before vs. After, main effects of Anomalous vs. Sensible or significant interactions with time are reported below. ROIs are ranked in order of magnitude of $t$ for the maxima in the task vs. baseline contrast (see Table 6). All interactions with Time reported below are for the quadratic trend as this trend would most closely mimic the pattern of the hemodynamic response.

Two ROIs demonstrated a significant main effect of Before/After and an interaction of Before/After X Time. An additional ROI demonstrated significant main effects of Before vs. After, but no interaction of Before/After X Time. One ROI showed a significant main effect of Anomalous vs. Sensible and two ROIs yielded a Anomalous vs. Sensible X Before vs. After X Time interaction.

ROI 1. The first ROI was located in the left inferior and middle frontal gyri (LIFG) in BA 9 (coordinates, $x = -46$, $y = 9$, $z = 28$). This ROI was associated with the largest $t$ value in the analysis ($t = 14.76$). This ROI demonstrated a significantly greater activation in the Before condition relative to the After condition, $F (1,11) = 5.037$, $p = .046$ and an interaction of Before/After X Time, $F (1,11) = 10.191$, $p = .009$ (Figure 21a).

ROI 2. The second ROI showing statistically significant effects was located more anteriorly (coordinates, $x = -41$, $y = 39$, $z = 9$) in the LIFG compared to ROI 1. This region was near BA 10 and BA 46. This region demonstrated significantly greater
activation in the before than the after condition $F(1, 11) = 9.938, p = .009$ (Figure 21b). In this ROI, there was also a significant interaction of Before/After X Time, $F(1,11) = 5.040, p = .046$. In addition, ROI 2 also showed a main effect of Anomalous vs. Sensible, $F(1, 11) = 7.785, p = .018$. However, there was no interaction of Anomalous/Sensible X Time and this effect is not included in a figure.

**ROI 3.** ROI 3 was located near Broadman’s Area 44 (coordinates, $x = -52, y = 10, z = 11$) and demonstrated a significant Anomalous/Sensible X Before/After X Time interaction, $F(1, 11) = 5.703, p = .036$ (Figure 22).

**ROI 6.** ROI 6 was located in the left medial frontal gyrus near BA 32 and the cingulate gyrus (coordinates, $x = -1, y = 10, z = 44$). ANOVAs revealed a significant Anomalous/Sensible X Before/After X Time interaction in this region, $F(1, 11) = 4.992, p = .047$ (Figure 23).

**ROI 10.** Finally, another ROI showing a significant main effect of Before/After and was located in the left middle frontal gyrus, near BA 8 (coordinates, $x = -46, y = 11, z = 37$). This region showed a significant main effect of Before vs. After, $F(1, 11) = 7.32, p = .02$, but no interaction of Before/After X Time (Figure 24).

Notably, other areas in the left frontal areas that have been implicated in phonological rehearsal did not show significant Before/After effects. A region in the left insula (ROI 4, Figure 21c) ($x = -42, y = 18, z = 13$) showed no effect of Before/After, $F(1, 11) = .000, p = .994$ and no interaction of before/after X time, $F(1, 11) = .028, p = .871$. ROI 5 in the left superior frontal gyrus ($x = -4, y = 10, z = 55$) showed no effects of before/after, $F(1, 11) = 3.063, p = .108$, or interaction of before/after X time, $F(1, 11) = 1.161, p = .304$. Another region, located in Brodmann’s Area 44 (ROI 3) also
failed to show a main effect of Before/After, $F(1, 11) = .539$, $p = .478$, or Before/After X Time interaction $F(1, 11) = 1.957$, $p = .189$. This suggests that the Before/After manipulation is effective in engaging semantic short-term memory processes that are not attributable to differences in phonological rehearsal.

**Additional ROI analyses.** Given that left parietal and left middle temporal areas are also of relevance to the present study, further ROI analyses were conducted with the Task vs. Baseline contrasted at $t = 9.92$, $p = 1.0^{18}$. Given that this yielded 21 additional ROIs (for a total of 33), only ROIs located in the left or right inferior frontal areas, left or right parietal areas or left middle temporal areas were analyzed with ANOVAs. Only those regions yielding significant main effects or interactions with the time factor are reported here. All interactions with Time are for the quadratic trend.

**ROI 17.** This ROI was located near the right insula (coordinates, $x = 38$, $y = -17$, $z = 7$) and yielded a marginally significant interaction of Before/After X Time, $F(1,11) = 4.665$, $p = .054$ (Figure 25).

**ROI 18.** This ROI was located in right inferior frontal gyrus (coordinates, $x = 50$, $y = 13$, $z = 15$), near Broadman’s Area 44. This area showed a significant interaction of Anomalous/Sensible X Time, $F(1,11) = 13.455$, $p = .004$, with greater activation in the anomalous relative to the sensible condition (Figure 26).

**ROI 25.** A ROI in the right inferior parietal lobe, near Broadman’s Area 40 (coordinates, $x = 41$, $y = -45$, $z = 40$) showed a main effect of Before/After, $F(1,11) = 6.529$, $p = .027$, whereby activation was greater in the before condition compared to after (Figure 27). In addition, there was a significant Before/After X Time interaction, $F(1,11) = 6.158$, $p = .030$. 
ROI 28. A region in the left inferior parietal lobe (x = -45, y = -49, z = 43), near Broadman's 40 also showed a marginally significant effect of Before/After, $F(1, 11) = 4.483$, $p = .058$ and a trend toward a Before/After X Time interaction, $F(1, 11) = 3.876$, $p = .075$. The response function for this ROI appears in Figure 28. This ROI is included here given the relevance of parietal areas in models of short-term memory.

ROI Analysis of Left Middle Temporal Areas

Given the importance of middle temporal regions to models of semantic processing (see above), the threshold for identifying ROIs was lowered to $t = 6.89$, $p = .11$, which identified over 90 additional ROIs. For the purposes of the present study, only two ROIs located in the middle temporal lobe were analyzed, in order to address a priori hypotheses regarding the role of middle temporal areas in semantic processing.

Two ROIs were identified, both in the middle temporal gyrus near Broadman's Area 21. Neither of the ROIs showed main effect of Before/After or interaction of Before/After X Time (main effect, Before/After, ROI 43, $F(1, 11) = 1.283$, $p = .282$; ROI 47, $F(1, 11) = .673$, $p = .429$). The response functions for two middle temporal ROIs appear in Figure 29.

Discussion: Experiment 4

Experiment 4 examined a task reported to discriminate between semantic and phonological short-term memory deficits. However, Experiment 4 used a modified version of the sentence anomaly task used in Experiment 3. The behavioral data collected during Experiment 4 are notable, given that they demonstrate that merely
presenting the critical phrases elicits the same types of "integration costs" when they are removed from the context of a complete sentence.

Given the previous data of Martin and Romani (1994) and Martin and He (2004), it was hypothesized that the Before/After manipulation would produce differences in hemodynamic response in areas of the left inferior frontal areas. Of secondary interest was to determine whether activations in Experiment 4 would overlap with areas related to interference resolution, specifically, the recent negatives task (Jonides et al., 1998). Finally, activity in the temporal lobe was also of interest, given the role of the temporal lobe in semantic processing.

The first hypothesis was confirmed, in that the Before/After manipulation produced statistically significant differences in hemodynamic response in two ROIs in the left inferior frontal gyrus. Each of the ROIs showing the before vs. after effect were compared to a number of relevant studies in the neuroimaging literature. Table 7 presents the Talairach or MNI coordinates (in RAI mm) for the two relevant ROIs in Experiment 4 as well as the coordinates reported in several relevant studies of semantic short-term memory and other studies examining various interference paradigms.

The first ROI \((x = -46, y = 9, z = 28)\) is near an area reported by Martin, Wu, Freedman, Jackson, and Lesch (2003) as being involved in short-term memory load (using a contrast of 4 items vs 1 item). Moreover, the maximum peak in this ROI is near an area reported by Jonides et al. (1997) in a neuroimaging study of the n-back task. It is also near an area reported to be responsive to selection demands involved in semantic retrieval (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). The observation that the greatest region of activation in Experiment 4 is near an area reportedly involved in
load effects in short-term memory provides additional evidence that the maintenance of unintegrated adjectives in the Before/After paradigm does, in fact, place a greater load on short-term memory processes.

The second ROI showing before vs. after effects (coordinates, $x = -41, y = 39, z = 9$) is located near an area reported by McDermott, Petersen, Watson, and Ojemann (2003) in a study of short-term maintenance of semantic and phonological representations. McDermott et al. reported an area in the LIFG (coordinates, $x = -43, y = 39, z = 0$) that indicated greater activation during maintenance of semantic information when compared to maintenance of phonological information. This is an intriguing observation, given that patients with semantic short-term memory deficits have a particularly difficult time in detecting anomalies when three semantic representations must be maintained in an unIntegrated form. The second ROI reported in Experiment 4 may represent an area crucial to the short-term maintenance of semantic representations. One may speculate that the first ROI represents an area devoted to general short-term memory processes, while the second ROI is important in the maintenance of semantic information. This idea will be further pursued in the General Discussion.

Moreover, the second ROI is also near an area reported by Poldrack, Wagner, Prull, Desmond, Glover and Gabrieli (1999) as demonstrating greater activation in a “semantic processing” task (abstract vs. concrete judgments) when compared to a “phonological processing” task (syllable counting). Poldrack and Wagner (2004) have proposed what they term the “semantic-phonological hypothesis” of left inferior frontal organization. According to this hypothesis, “semantic processing” is associated with activity of the anterior regions of the left inferior frontal areas, while “phonological
processing” is localized to more posterior areas of left inferior frontal regions. While the term “processing” is rather vague and somewhat inaccurate (brain damaged patients with damage to left inferior frontal areas do not have deficits of semantic knowledge and perform well on semantic tasks such as category judgments, see Martin and He, 2004, for an example), Experiment 4 provides some evidence for at least one aspect of this “semantic-phonological hypothesis”. Specifically, the anterior regions of the left inferior frontal areas are indeed important for maintenance and/or controlled retrieval of semantic knowledge (see Badre & Wagner, 2002 for a review of the controlled vs. automatic retrieval distinctions in the anterior and posterior divisions in LIFG).

Thus, these data support the previous hypotheses proposed by Martin and Romani (1994) that maintaining three unintegrated prenominal adjectives places greater demands on short-term memory compared to when adjectives appear after a noun. Moreover, greater activation in LIFG areas associated with maintenance of semantic representations supports the hypothesis that the before condition places greater demands on semantic short-term memory (Martin & He, 2004).

In addition, Experiment 4 provides data relevant to a current debate concerning the organization of semantic processing in the brain. A considerable body of neuropsychological research indicates that long-term representations of semantic knowledge are localized to the temporal lobes (Mummery et al. 2000). This and other evidence has prompted some researchers to propose a semantic processing network in which frontal areas are responsible for maintenance and control of semantic representations, while the representations being acted upon are located in temporal lobes.
During semantic short-term maintenance, temporal areas are presumably active due to the top-down signals from inferior frontal areas (Shivde and Thompson-Schill, 2004).

Additional support for this hypothesis is derived from the observation that temporal areas are not sensitive to tasks placing varying demands on control processes during semantic processing tasks (see Thompson-Schill et al., 1997 for an example). However, Noppeney, Phillips and Price (2004) have recently challenged this model with a study that reported temporal activity that varies with what they term "semantic executive processes". Experiment 4 provides data more consistent with a model of semantic processing represented by a frontal-temporal network, in which the temporal lobes subserve long-term semantic knowledge representations that are acted upon by inferior frontal control processes.

It is important to note that areas traditionally implicated in phonological rehearsal are located in BA 44 and the insula. For example, Awh et al. (1996) reported that two areas, in BA 44 and BA 6 (in the supplemental motor area), are related to phonological rehearsal. In addition, Dronkers (1996) identified the left insula as being crucial in articulatory planning. In the present data, these areas related to phonological rehearsal and motor planning show no differences between the before and after conditions. Thus, it appears that the differences between the before and after conditions observed in the left inferior frontal gyrus are not due to mere differences in phonological rehearsal.

One problem inherent in attempts to dissociate semantic and phonological short-term memory is the difficulty in finding a task that fully dissociates the two processes. That is, phonological and semantic processing are obligatorily engaged during comprehension of words. One way that researchers have attempted to circumvent this
problem is by using non-words – words that have no semantic meaning attached. However, the use of nonwords in a phonological task does not address the presence of phonological processing in the semantic tasks.

Given the difficulty of effectively dissociating semantic and phonological processing, Experiment 4 suggests the Before/After anomaly paradigm presented in Experiment 4 may be an especially useful paradigm in elucidating the neural substrate of semantic short-term memory. The rapid visual presentation may minimize (although not eliminate) phonological rehearsal and the Before v. After conditions provide a convenient contrast by which semantic short-term memory load may be examined. Future studies could also include the distance manipulation (presenting either one, two or three words for integration) employed by Martin and Romani (1994) to explore effects of load in the task.

In addition to frontal regions discussed above, Experiment 4 also revealed a right parietal area (x = 41, y = -45, z = 40) that demonstrated significant before vs. after effects. In addition, a nearly homologous region in the left parietal lobe (x = -45, y = -49, z = 43) showed a significant trend toward a Before/After X Time interaction. This right parietal area is near BA 40 and is reported to be involved in verbal working memory by Jonides et al. (1998). However, in a review of the neuroimaging literature examining verbal working memory, Becker, MacAndrew and Fiez (1999) identified two areas within the parietal lobe that had been previously related to the “phonological store” – an area in more superior parietal areas (mean coordinates: x = -33, y = -48, z = 39) and another more inferior area of the parietal lobe (mean coordinates: x = -52, y = -27, z = 22). Becker et al. (1999) argued that the latter, more inferior area, seemed to better meet
the criteria for the phonological store. Importantly, the inferior area is more consistent
with the neuropsychological studies of phonological short-term memory deficits. These
studies have reported that the inferior supramarginal and angular gyri are related to
phonological short-term memory (Vallar & Papagno, 1995). The authors suggest that the
more superior area may be related to visual attention and note that tasks of visual
attention typically elicit bilateral activation of the superior parietal lobe, but that these
activations are typically larger in the right hemisphere compared to the left. The parietal
ROIs in Experiment 4 appear superior to the areas thought to be related to phonological
maintenance and storage. In addition, the areas showing significant effects of Before vs.
After, fall outside an ROI reported by Martin et al. as being important in phonological
short-term memory. Therefore, it seems possible that the parietal activations in
Experiment 4 represent greater attentional load in the Before condition relative to the
After condition and not necessarily engagement of phonological short-term memory
processes. Figure 30a shows the relative position of the parietal ROI in experiment 4
compared to the areas typically related to phonological short-term memory in Figure 30b
(as identified by Becker et al. 1999 and Martin et al., 2003).

Other evidence supports the assertion that the parietal activations in Experiment 4
may be related to attention rather than phonological short-term memory. LaBar,
Gitelman, Parrish and Mesulam (1999) noted that many neuroimaging studies have
confounded working memory and spatial attention, making it difficult to assess the
functional relevance of different parietal regions. Specifically, many working memory
tasks use a probe recognition task in which list items are presented simultaneously in a
spatial configuration and are then followed by a probe. To address this shortcoming, the
authors designed verbal working memory and spatial attention tasks that minimized overlapping cognitive features. The working memory task was a 2-back task that presented a series of letters foveally. Thus, it required no shifts of spatial attention. The spatial attention task was a variation of the Posner paradigm, in which attentional shifts are cued by a centrally presented arrow. Interestingly, the authors found that an area very similar to the left parietal ROI in Experiment 4 (ROI 28) was involved in both the verbal working memory task and the spatial attention task. The authors proposed that the overlap was related to the shifting of attentional focus over space, time or cognitive domains. Of particular note, the task used in the LaBar et al. (1999) experiment used a working memory task that did not require shifts of spatial attention. Instead, all stimuli in the working memory task were presented foveally, much like the rapid serial visual presentation in Experiment 4.

The aim of Experiment 4 was to identify areas of the LIFG engaged by a task that taps semantic short-term memory. However, Experiment 4 also provides an opportunity to examine hemodynamic response to semantic violations in the task. Three ROIs demonstrated significant interactions of Sensible/Anomalous X Time or Sensible/Anomalous X Before/After X Time. The first was ROI 3 near left BA 44, which showed an interaction of Sensible/Anomalous X Before/After X Time. Another, ROI 18, was located in right BA 44. ROI 6, in the left medial frontal gyrus, near BA 32, showed a significant Sensible/Anomalous X Before/After X Time interaction in which the anomalous trials showed greater activation than sensible trials, but only in the after condition. In addition, ROI 2, implicated previously in short-term semantic maintenance, showed a main effect of Sensible/Anomalous, but no interaction with time.
Kiehl, Laurens and Liddle (2002) examined a sentence anomaly task in which subjects read sentences that ended in semantically congruent or incongruent words. In this study, the authors report bilateral activation of the inferior frontal gyrus. These areas of activation (x = -48, y = 32, z = 4) are anterior to the inferior frontal areas showing Sensible/Anomalous X Time effects in Experiment 4. However, the area responding to semantic violations that was reported by Kiehl et al. (2002) is near ROI 2 which showed significant Before/After X Time differences and a main effect of Anomalous Sensible in Experiment 4.

Friederici, Ruschemeyer, Hahne and Fiebach (2003) also examined a sentence anomaly task in which semantically incongruous sentences were contrasted with semantically congruent sentences. However, this study used auditory presentation of sentences. This study also found bilateral activations in the inferior frontal gyri, in what the authors describe as the insula. Again, these reported areas differ from those in Experiment 4. The areas responding to semantic violations in Friederici et al. (2003) are posterior and inferior to those responding to the semantic violations in Experiment 4. Importantly, both the Kiehl, Laurens and Liddle (2002) and the Friederici et al (2003) experiments also reported significant temporal activation in response to semantic violations. Thus, these data correspond to previous research that has used intra-cranial recordings to identify areas involved in the generation of the N400. Studies using intra-cranial recordings have typically reported superior temporal areas as the source of the N400 to semantic violations (McCarthy, Nobre, Bentin & Spencer, 1995). However, Experiment 4 found no superior temporal activations showing Sensible vs. Anomalous differences.
Finally, previous studies examining semantic violations during sentence comprehension have not reported activations in anterior cingulate. However, Experiment 4 revealed an ROI in anterior cingulate that showed a significant Sensible/Anomalous X Before/After X Time interaction. This interaction is explained by greater activation in the anomalous condition, but only in the After condition. The anterior cingulate (ACC) has been studied extensively in the cognitive neuroscience literature (see Botvinick, Cohen and Carter, 2004, for a recent review). The canonical view of ACC function is that it plays a role in conflict monitoring and triggers strategic control when conflict is detected. For example, the ACC is commonly reported to respond to conflict in the Stroop task, flanker tasks, local-global tasks and the go/no-go paradigm (Botvinick, Cohen & Carter, 2004). Moreover, the ACC is reliably activated by the commission of errors. Inasmuch as semantic anomalies represent conflict, it would seem unsurprising that this task would activate ACC. However, as noted above, other neuroimaging studies examining semantic violations have not reported ACC involvement. Regardless, the Before/After X Sensible/Anomalous interaction is intriguing, given that it reinforces the hypothesis put forth in Experiment 3. Specifically, that semantic anomalies are less salient in the Before condition compared the After condition. Consistent with this hypothesis, the fMRI data reveal greater activation in the ACC for the Anomalous condition, but only in the After condition.

GENERAL DISCUSSION

Broadly, the present experiments have addressed questions related to semantic short-term memory as first proposed by Martin, Shelton and Yaffee (1994). Experiment 1 addressed an apparent paradox by which patients with semantic short-term memory
deficits show intrusions of previously presented list items during short-term recall (Martin and Lesch, 1996). Furthermore, patients with semantic short-term memory deficits also show unusual interference effects when naming semantically related pairs of pictures (Martin, Freedman and Biegler, 2004). These two findings suggested that semantic short-term memory patients might have difficulty with suppressing or inhibiting some types of irrelevant information. Experiment 1 presented data from a number of tasks commonly assumed to tap inhibition. Experiment 1 yielded several interesting findings. First, patient ML showed dramatically exaggerated effects of interference on the Stroop task, a task that makes minimal demands of short-term memory. Secondly, consistent with the previously reported intrusions during serial recall, ML showed greatly exaggerated susceptibility to proactive interference in the recent negatives task. In contrast, patient ML performed well within the range of healthy age and education matched controls on a non-verbal analogue of the Stroop task as well as the antisaccade task. ML’s normal performance on the antisaccade task is especially interesting given that antisaccade performance has been reported to correlate with working memory ability (Roberts et al. 1994, Kane et al. 1999). Experiment 1 proposes that the pattern of performance of patient ML suggests the dissociability of inhibition for verbal compared to nonverbal tasks.

Experiment 2 more closely examines the recent negatives task in the ERP environment. Although a considerable number of studies have examined the recent negatives task using neuroimaging, none have yet used the superior temporal resolution of ERP to investigate the task. Experiment 2 incorporated a novel effect identified in patient ML’s performance on the recent negatives task. Specifically, patient ML’s
performance was much better when repetition of items was minimized and items were repeated only as necessary to form the recent negatives trials. Experiment 2 compared the standard recent negatives task, which repeats a set of 16 items throughout the task (the "closed" condition), with a version that minimized repetition (the "open" condition). This manipulation was designed to help identify components that were uniquely related to proactive interference. Data from Experiment 2 are somewhat equivocal. Although two components were found to differ between recent and nonrecent negatives, the interactions of recency (recent negatives v. nonrecent negative) X repetition (open v. closed sets) differed for each component. Specifically, a N400 effect, distributed over frontal areas, was larger for nonrecent negatives compared to recent negatives, but only in the closed condition. By contrast, a temporally later effect, a Late Positive Component (LPC), was observed to differ between recent and nonrecent negatives, but only for the open condition. This component was localized over parietal channels. These components are related to the literature, which has related the FN400 to a "familiarity process" (Curran, 1999) and the LPC to a "recollection process" or "response confidence". However, it is unclear why recollection processes as indexed by the LPC were not apparent in the Closed condition, given that one may intuitively predict that Closed condition would rely more heavily on recollection processes (given that familiarity and interference is high in all conditions of Closed sets).

Experiment 3 used a paradigm reported by Martin and Romani (1994) and Martin and He (2004) to be particularly difficult for patients with semantic short-term memory deficits. This sentence anomaly task has two types of sentence constructions. The "before" condition places multiple adjectives before a noun (in other trials, multiple
nouns appear before a verb), while in the after condition, multiple adjectives follow the noun. Behavioral data from healthy subjects indicate that before condition is more difficult (as indicated by longer reaction times and poorer accuracy) than the after condition, presumably due to the greater difficulty in integrating semantic representations in to a message-level proposition. This task is particularly difficult for patients with semantic short-term memory deficits. Experiment 3 uses the parietally distributed N400 (which differs from the frontal N400 in Experiment 2) as an indirect measure of the consequences of integration. It was hypothesized that N400s would be smaller in the before condition relative to the after condition, primarily because semantic anomalies would be less salient. Ideally, this effect would be seen in the interaction of Before/After X Sensible/Anomalous – semantic anomalies would elicit smaller N400s in the before condition compared to the after condition. Instead, Experiment 3 revealed an interaction of Before/After X Laterality (a significant difference was seen in the before vs. after contrast over left hemisphere channels, but not the right hemisphere). This can be reconciled with the previous literature examining the N400, if one assumes that the before condition undermines the formation of context, which leads to smaller N400 effects.

Experiment 4 uses a variation of the before/after sentence anomaly paradigm to examine short-term memory demands in the before vs. after condition. Given that fMRI necessitates precise stimulus timing, Experiment 4 used rapid serial visual presentation to present only the critical phrases from the sentence anomaly task. Behaviorally, reaction times were longer in the before condition relative to the after condition, replicating the general effect that is seen when complete sentences are presented. The imaging data
showed that two regions in the left inferior frontal areas showed statistically significant differences in hemodynamic responses between the before and after conditions. The first of these ROIs was very similar to an area reported by Martin et al. (2004) which showed differences between short-term maintenance of one word and four words. Moreover, this area is similar to areas reported by Jonides et al. (1997) which were active during an n-back task. Given that this area of the left inferior frontal gyrus is sensitive to load and shows significant before vs. after differences in hemodynamic response, these data provide further evidence that the maintenance of unintegrated semantic representations (ie. words) places greater demands on short-term memory.

A second ROI located more anteriorly to the first also showed before vs. after differences. This ROI was near areas reported to show greater activation in maintenance of semantic representations relative to phonological maintenance (McDermott et al. 2003). This provides further evidence that the before condition in this paradigm places greater demands specifically on semantic short-term memory processes.

*Is inhibition the causal mechanism explaining associations of semantic short-term memory deficits and comprehension and production deficits?*

Experiment 1 concluded with speculation that the inhibition deficit reported for patient ML might be causally related to his deficits on some tasks of language production and comprehension. Although these suggestions are obviously speculative, similar inhibition accounts have been invoked to explain language processing difficulties in some types of aphasia. For example, Robinson, Blair, and Cipolotti (1998), attributed a case of dynamic aphasia to the inability to resolve interference “under conditions when there are many competing verbal response options and a greater amount of mutual inhibition..”.
Even more intriguing, the Robinson, Blair and Cipolotti (1998) patient had damage in left BA 45 – the same areas implicated in the neuroimaging studies of the recent negatives task.

Associations of inhibition deficits and certain types of short-term memory and language processing deficits have raised the interesting possibility that all of these deficits share a common mechanism. For example, one might speculate that accurate maintenance of unintegrated adjectives in the Before/ After sentence anomaly paradigm might place particular demands on inhibitory processes that are involved in activating relevant semantic representations and inhibiting related, but irrelevant, representations (for a similar argument see Thompson-Schill, Bedny and Goldberg, in press). For example, with weaker context in the Before condition, prenominal adjectives would activate more semantic knowledge (given that the adjectives are not constrained by context) which would have to be eventually inhibited in order to form a proposition during semantic integration. In fact, some have termed this collection of putative control processes necessary for language processing as “semantic executive processes”, which include retrieval, selection and control of depth of semantic analysis (Noppeney, Phillips and Price, 2004). Without such a mechanism to keep relevant semantic representations active and unfettered by competing information, maintenance might fail. Thus, one might expect to see areas of the left inferior frontal gyrus, specifically BA 45, to be active during demanding language comprehension tasks, reflecting interference resolution processes ascribed to this area. Data from Experiment 4 begin to address this question.

Stated simply, are differences between the Before and After conditions apparent in BA 45? The answer from Experiment 4 is an equivocal “no”. Differences between the
Before and After conditions were found in an area thought to be implicated in effects of load (maintenance of one item vs. maintenance of four items) supporting the idea that the Before condition places greater demands on short-term memory. Another region showing Before/ After differences was found in an area reported to show greater activation in semantic maintenance relative to phonological maintenance, supporting the idea that the Before condition places particular demands on semantic short-term memory. However, both of these ROIs were centimeters (typically in the y dimension, see Table 7) from areas reported to be involved in resolution of interference in the recent negatives task.

However, it would be premature to abandon the idea that the Before condition of the anomaly task makes demands of control processes involved in interference resolution. In fact, the first ROI reported in Experiment 4 is within millimeters of a region reported by Thompson-Schill et al. (1997) to be related to one such control process, which the authors termed “selection”. Thompson-Schill et al. manipulated selection demands in three semantic retrieval tasks and attributed activity of a region centered around MNI coordinates, $x = -49$, $y = 8$, $z = 30$ to be involved in selection of semantic information among potential competing responses (but see Badre & Wagner, 2002, for a different account).

Furthermore, the ROI 1 reported in Experiment 4 is also near an area reported to show greater activation in incongruent trials relative to congruent trials in the Stroop task (Leung, Skudlarski, Gatenby, Peterson and Gore, 2000). Given that the Stroop task is the widely considered to be the prototypical paradigm for the study of interference resolution and inhibition, the proximity of ROI 1 in Experiment 4 to areas showing the
"Stroop effect" suggests the maintenance of semantic representations in the Before condition might indeed rely on some areas of the brain that are also involved in interference resolution.

In sum, the present experiments address several important issues regarding the relationship between semantic short-term memory and interference resolution. First, given the data presented in Experiment 1, it does seem that patients with semantic short-term memory might be especially susceptible to interference in the verbal domain. However, whether difficulties with resolution of interference are causally related to difficulties in maintaining semantic representations in short-term memory is an open question. One possibility is semantic processing, including semantic retrieval necessary for many of the semantic short-term memory tasks, depends more heavily on resolution of interference from innumerable semantic competitors. Thus, without a properly functioning inhibitory mechanism capable of resolving the interference in short-term memory, the system becomes functionally overwhelmed with competing semantic alternatives. Importantly, this explanation is very different from most conceptualizations of short-term memory deficits, which have most often assumed a rapid decay of representations in short-term memory. By contrast, the present data suggests that semantic short-term memory deficits might result from an over-activation of representations in short-term memory which serves to undermine short-term memory performance. Data demonstrating that a patient with a severe short-term memory deficit shows exaggerated effects of proactive interference on a short-term memory task strongly suggests that a simple decay account is not sufficient to explain at least some cases of such deficits.
Of course, another, albeit less interesting, possibility is that the deficits of inhibition and deficits of semantic short-term memory are mere associations resulting from two separable mechanisms residing in spatially proximal regions of the left inferior frontal gyrus. For example, patient ML, reported in Experiment 1, has a lesion that encompasses most of the left inferior frontal gyrus. His behavioral deficits, which include both exaggerated interference on verbal tasks and a severe short-term memory deficit, may be attributable to having damage to separable inhibitory and semantic short-term memory mechanisms which are localized to different regions of the left inferior frontal gyrus. However, data from Experiment 4 suggest that similar control processes may be recruited for comprehension tasks that place demands on semantic short-term memory and tasks that require resolution of interference in the Stroop task and semantic selection tasks. Thus, the hypothesis that the same control processes may play a fundamental role in both semantic short-term memory and language comprehension and production is a question worthy of further pursuit.
References


Table 1. Behavioral Data – ERP Recent Negatives Task – Undergraduate Subjects

<table>
<thead>
<tr>
<th></th>
<th>Closed Blocks</th>
<th></th>
<th></th>
<th>Open Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recent Negative</td>
<td>Nonrecent Negative</td>
<td>Recent Positive</td>
<td>Recently Negative</td>
</tr>
<tr>
<td>RT</td>
<td>845</td>
<td>764</td>
<td>725</td>
<td>733</td>
</tr>
<tr>
<td>Accuracy</td>
<td>96.30</td>
<td>98.43</td>
<td>97.58</td>
<td>95.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Recent Negative</th>
<th>Nonrecent Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>808</td>
<td>720</td>
<td>715</td>
</tr>
<tr>
<td>Accuracy</td>
<td>98.43</td>
<td>98.72</td>
<td>96.94</td>
</tr>
</tbody>
</table>
### Table 2: Stimuli for Experiment 3 – Before condition

**Before/Delayed Integration**
*(greater semantic STM demands)*

<table>
<thead>
<tr>
<th>Adjective Noun Phrases</th>
<th>Anomalous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensible</td>
<td>Anomalous</td>
</tr>
<tr>
<td><em>A fluffy, small, surprised kitten</em> came out of the room.</td>
<td><em>A fluffy, small, surprised shriek</em> came out of the room.</td>
</tr>
<tr>
<td>Noun Verb Phrases</td>
<td>Anomalous</td>
</tr>
<tr>
<td>Sensible</td>
<td>Anomalous</td>
</tr>
<tr>
<td><em>Jeeps, men and women</em> were <em>crowding</em> the streets.</td>
<td><em>Jeeps, men and women</em> were <em>walking</em> the streets.</td>
</tr>
</tbody>
</table>
Table 3: Stimuli for Experiment 3 – After condition

<table>
<thead>
<tr>
<th>Adjective Noun Phrases</th>
<th>Anomalous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensible</td>
<td></td>
</tr>
<tr>
<td>The kitten was <em>fluffy, small and surprised</em> as it came out of the room.</td>
<td>The shriek was <em>fluffy, small and surprised</em> as it came out of the room.</td>
</tr>
<tr>
<td>Noun Verb Phrases</td>
<td></td>
</tr>
<tr>
<td>Sensible</td>
<td></td>
</tr>
<tr>
<td>The streets were <em>crowded by Jeeps, men and women.</em></td>
<td>The streets were <em>walked by Jeeps, men and women.</em></td>
</tr>
</tbody>
</table>
Table 4: Experiment 3- Sample sentences. Word in italics indicates the critical word determining sensible v. anomalous status.

Critical word was the 7th word within the sentence.

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th>Conditions</th>
<th>Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adj-N/N-Adj</td>
<td>Before-sensible</td>
<td>She saw the green, shining, bright <em>emerald</em>, which pleased her.</td>
</tr>
<tr>
<td></td>
<td>Before-anomalous</td>
<td>She saw the green, shining, bright <em>sun</em>, which pleased her.</td>
</tr>
<tr>
<td></td>
<td>After-sensible</td>
<td>The emerald was shining, bright and <em>green</em>, which pleases her.</td>
</tr>
<tr>
<td></td>
<td>After-anomalous</td>
<td>The sun was shining, bright and <em>green</em>, which pleased her.</td>
</tr>
<tr>
<td>Np-V/V-Np</td>
<td>Before-sensible</td>
<td>The fruit, sandwiches and chips were <em>eaten</em> at lunch today.</td>
</tr>
<tr>
<td></td>
<td>Before-anomalous</td>
<td>The plates, sandwiches and chips were <em>eaten</em> at lunch today.</td>
</tr>
<tr>
<td></td>
<td>After-sensible</td>
<td>They ate the chips, sandwiches and <em>fruit</em> at lunch today.</td>
</tr>
<tr>
<td></td>
<td>After-anomalous</td>
<td>They ate the chips, sandwiches and <em>plates</em> at lunch today.</td>
</tr>
</tbody>
</table>

Critical word was the 10th word within the sentence.

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th>Conditions</th>
<th>Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adj-N/N-Adj</td>
<td>Before-sensible</td>
<td>In the forest, she saw the fierce, large, brown <em>bear</em> so she hid.</td>
</tr>
<tr>
<td></td>
<td>Before-anomalous</td>
<td>In the forest, she saw the fierce, large, brown <em>rock</em> so she hid.</td>
</tr>
<tr>
<td></td>
<td>After-sensible</td>
<td>In the forest, the bear was large, brown and <em>fierce</em> so she hid.</td>
</tr>
<tr>
<td></td>
<td>After-anomalous</td>
<td>In the forest, the rock was large, brown and <em>fierce</em> so she hid.</td>
</tr>
<tr>
<td>Np-V/V-Np</td>
<td>Before-sensible</td>
<td>Late this morning, the pants, shirts and washcloths were <em>folded</em>, as instructed.</td>
</tr>
<tr>
<td></td>
<td>Before-anomalous</td>
<td>Late this morning, the soaps, shirts and washcloths were <em>folded</em>, as instructed.</td>
</tr>
<tr>
<td></td>
<td>After-sensible</td>
<td>Late this morning, maids folded the washcloths, shirts and <em>pants</em>, as instructed.</td>
</tr>
<tr>
<td></td>
<td>After-anomalous</td>
<td>Late this morning, maids folded the washcloths, shirts and <em>soaps</em>, as instructed.</td>
</tr>
</tbody>
</table>
Table 5: Filler trials in Experiment 3

<table>
<thead>
<tr>
<th>Filler Trials</th>
<th>Semantic anomaly</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>early position</strong></td>
<td>The desk <strong>drunk</strong> by the hungry, lonely, homeless man was smelly and not very tasty.</td>
</tr>
<tr>
<td><strong>middle position</strong></td>
<td>The graceful woman began to dial the <strong>dance</strong> in front of the crowd that had gathered.</td>
</tr>
<tr>
<td><strong>late position</strong></td>
<td>Earlier that morning, the foolish young girl was burnt by the <strong>book</strong>.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Grammatical anomaly</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>early position</strong></td>
<td>He <strong>training</strong> and effort have been rewarded with a victory.</td>
</tr>
<tr>
<td><strong>middle position</strong></td>
<td>Janie had worked for newspapers before she <strong>write</strong> her own novel last year.</td>
</tr>
<tr>
<td><strong>late position</strong></td>
<td>Just before the dinner, the pleasant, excited hostess came in and asked everyone to <strong>sitting</strong> down.</td>
</tr>
</tbody>
</table>
Table 6: 12 ROIs identified by task vs. baseline contrast. ROIs were determined by selecting voxels significant at $t=11.33$, $p = 1.22$. ROIs appear in order of magnitude of $t$ value for maxima used to identify original ROI. However, when several voxels exceeding threshold were within 10mm, a center of mass was calculated among the voxels and the ROI was calculated around this center. The $t$ value that appears in the table represents the $t$ value for this geometric center. Coordinates are in standardized Talairach space (RAI mm).

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>z</th>
<th>$t$ value of center of mass of ROI</th>
<th>Anatomical Structure</th>
<th>BA</th>
<th>Significant Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>-46</td>
<td>9</td>
<td>28</td>
<td>14.762</td>
<td>LIFG/LMFG</td>
<td>9</td>
<td>B/A, B/A X Time</td>
</tr>
<tr>
<td>-41</td>
<td>39</td>
<td>9</td>
<td>12.227</td>
<td>LIFG/LMFG</td>
<td>46</td>
<td>B/A, B/A X Time</td>
</tr>
<tr>
<td>-52</td>
<td>10</td>
<td>11</td>
<td>11.351</td>
<td>LIFG</td>
<td>44</td>
<td>A/S X B/A X Time</td>
</tr>
<tr>
<td>-42</td>
<td>18</td>
<td>13</td>
<td>10.068</td>
<td>LIFG/Insula</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>10</td>
<td>55</td>
<td>12.932</td>
<td>LSFG</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>10</td>
<td>44</td>
<td>10.83</td>
<td>L medial FG</td>
<td>32</td>
<td>A/S X B/A X Time</td>
</tr>
<tr>
<td>-14</td>
<td>-6</td>
<td>19</td>
<td>10.108</td>
<td>L Caudate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-48</td>
<td>22</td>
<td>29</td>
<td>10.736</td>
<td>LMFG</td>
<td>9, 46</td>
<td></td>
</tr>
<tr>
<td>-43</td>
<td>-3</td>
<td>29</td>
<td>12.206</td>
<td>L Precentral G</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>-46</td>
<td>11</td>
<td>37</td>
<td>10.743</td>
<td>LmidFG</td>
<td>9</td>
<td>B/A</td>
</tr>
<tr>
<td>-4</td>
<td>21</td>
<td>47</td>
<td>9.823</td>
<td>L Medial FG</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>-26</td>
<td>50</td>
<td>10.504</td>
<td>R Postcentral G</td>
<td>40, 2</td>
<td></td>
</tr>
</tbody>
</table>
Table 7: Stereotaxic coordinates of ROIs demonstrating significant Before / After X time interactions in Experiment 4 compared to several neuroimaging studies of interference resolution, short-term maintenance of semantic representations and short-term memory load.

<table>
<thead>
<tr>
<th>Hamilton, Exp 4 - ROI 1</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>deviation in mm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manipulations of STM Load</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jonides et al (1997) – 3-back task</td>
<td>-44</td>
<td>8</td>
<td>27</td>
<td>2 1 1</td>
</tr>
<tr>
<td>Martin, Wu et al (2003) Load4 v. Load1</td>
<td>-40</td>
<td>9</td>
<td>27</td>
<td>6 0 1</td>
</tr>
<tr>
<td><strong>Semantic Selection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson-Schill et al. (1997)</td>
<td>-49</td>
<td>8</td>
<td>30</td>
<td>3 1 2</td>
</tr>
<tr>
<td><strong>Stroop Interference (Inongruent&gt;Congruent)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leung et al. (2004)</td>
<td>-46</td>
<td>8</td>
<td>31</td>
<td>0 1 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hamilton, Exp 4 - ROI 2</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>deviation in mm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tasks of interference resolution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jonides et al. (1998)</td>
<td>-48</td>
<td>21</td>
<td>9</td>
<td>7 18 0</td>
</tr>
<tr>
<td>Jonides et al. (2000)</td>
<td>-51</td>
<td>21</td>
<td>11</td>
<td>10 18 2</td>
</tr>
<tr>
<td>Badre &amp; Wagner (2005)</td>
<td>-51</td>
<td>21</td>
<td>6</td>
<td>10 18 3</td>
</tr>
<tr>
<td><strong>Semantic STM Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crosson et al.(1999) ROI 1</td>
<td>-50</td>
<td>19</td>
<td>25</td>
<td>9 20 16</td>
</tr>
<tr>
<td>Crosson et al. (1999) ROI 2</td>
<td>-49</td>
<td>32</td>
<td>-1</td>
<td>8 7 10</td>
</tr>
<tr>
<td>Devlin, Matthews, &amp; Rushworth (2003)</td>
<td>-46</td>
<td>20</td>
<td>-3</td>
<td>5 19 12</td>
</tr>
<tr>
<td>McDermott et al (2003)</td>
<td>-43</td>
<td>39</td>
<td>0</td>
<td>2 0 9</td>
</tr>
<tr>
<td>Shivde &amp; Thompson-Schill (2004)</td>
<td>-37</td>
<td>22</td>
<td>0</td>
<td>4 17 9</td>
</tr>
<tr>
<td><strong>Semantic “Processing”</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poldrack et al. (1999) Sem &gt; Phon</td>
<td>-42</td>
<td>40</td>
<td>8</td>
<td>1 1 1</td>
</tr>
<tr>
<td>Abstract/concrete judgment &gt; syllable counting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Patient ML accuracy in the Open and Closed versions of the Recent Negatives Task – a.) Closed condition, b.) Open condition
Fig 2: Anterior Posterior Channels used for Frontal N400 (FN400) analyses
Fig 3: Experiment 2: FN400 component – Recent Negative v. Nonrecent Negatives in Open and Closed conditions. The temporal window used for analysis of FN400 was 300-500ms

a. Left & Right Channels- FN400 effect

b. Left Channels
c. Right Channels
Figure 4: Experiment 2 - FN400 – Open vs. Closed X Recent vs. Nonrecent. Statistically significant differences between Recent and Nonrecent Negatives were observed in the Closed condition, but not the Open condition.
Figure 5: Experiment 2 - FN400 effect in Closed condition with positive trials. Visual inspection reveals Recent negative trials appear more similar to positive trials compared to Nonrecent negative trials.

Left + Right - Closed

b. Left Channels
c. Right Channels
Figure 6: Experiment 2 - FN400 effect in Open condition with positive trials. No significant differences in the FN400 between Recent negatives and Nonrecent negatives was observed in the Open condition.
Figure 7: Negative and Positive Trials in Closed Condition
Figure 8: FN400 for Open condition: Interaction of Trial Type X Laterality
Figure 9. Experiment 2 - Late Positive Component (550ms-650ms) Open and Closed conditions. Recent Nonrecent negatives show a greater LPC than Recent negatives, but only in the Open condition.
Figure 10: Experiment 2 - Posterior Superior Channels used in analyses of Late Positive Component
Fig 11. Experiment 2: Late Positive Component- 550ms-650ms: Set Size X Recency Interaction
Figure 12: Experiment 2: Late Positive Component for Recent vs. Nonrecent negatives in Open condition

a. Late Positive Component – 500 – 800 ms

b. Left Channels
c. Right Channels

The rusty old red pail was lying in

Figure 2. On-line processes and working memory representations involved in comprehension of sentence with prenominal adjectives. def = definite; NP = noun phrase; N = noun; S = sentence; VP = verb phrase; det = determiner; adj = adjective; aux = auxiliary.
Figure 14. Schematic of short-term memory representations necessary in the “After” condition in Experiment 3. Adapted from Martin & Romani (1994).

propositions: def(pail) exist(pail) old(pail) red(pail) rusty(pail) past(exist)

syntactic structure:
- NP: det N
- VP: V AdjP Adj Adj Conj Adj
- AdjP: Adj Adj Conjunction Adj

lexical semantic/syntactic features:
- the def(pail) (definite) (container) (handle)
- pail (exist) (past) (worn) (color) (also) (oxidized) (orangish)
- was old red and rusty
- phonological: /ðæ/ /peyl/ /wæz/ /old/ /red/ /ænd/ /ʌstɪ/

The pail was old red and rusty

Figure 3. On-line processes and working memory representations involved in comprehension of sentence with postnominal adjectives. def = definite; NP = noun phrase; S = sentence; N = noun; VP = verb phrase; V = verb; AdjP = adjective phrase; Adj, adj = adjective; Conj, conj = conjunction; det = determiner.
Figure 15. Sentence Anomaly Task (a.) reaction time data for control Ss (b.) accuracy data for semantic STM patient AB (c.) accuracy data for semantic STM patient ML (d.) accuracy data for phonological STM patient EA.
Figure 16: Before vs. After X Sensibility interaction with visual presentation of sentences. Sensible vs. Anomalous differences in N400 in a.) Before and b.) After conditions. Adapted from Yang, Martin and Potts (2002). Differences between N400s in response to Sensible vs. Anomalous conditions are larger in the After condition compared to the Before Condition.
Figure 17: Experiment 3: N400 effects at parietal channels. A 200 – 600 ms window was selected for extraction of statistics for analysis. There was a significant Before/After X Laterality interaction, whereby N400s in the Before condition were smaller than N400s to the After condition, but only in the left hemisphere (b.). Channels selected for analysis appear in d.

a. Left & Right Channels- N400 effect

b. Left Channels

c. Right Channels

d. selected channels for N400 analysis
Figure 18: Experiment 3: Main effect of Sensibility
Figure 19: Experiment 3: After/Before X Laterality interaction. N400s were smaller in the Before condition, but only over the Left hemisphere.
Figure 20: Twelve ROIs defined by $t > 11.32$. 
Figure 21: ROIs 1, 2 and 4
Figure 22: ROI 3

ROI 3 / -52 10 11 / L BA 44
AFTER

ROI 3 / -52 10 11 / L BA 44
BEFORE
Figure 23: ROI 6 - Before/After X Sensible/Anomalous X Time interaction

ROI 6 / -1 10 44 / L Medial Frontal
AFTER

ROI 6 / -1 10 44 / L Medial Frontal
BEFORE
Figure 24: ROI 10

ROI 10 / -46 11 37 / L MFG

% Signal Change

TR

BEF_AFT

AFTER

BEFORE
Figure 25: ROI 17

ROI 17 / 38 17 7 / R Insula

% Signal Change

TR
Figure 26: ROI 18

ROI 18 / 50 13 15 / R Inferior Frontal / BA 44

% Signal Change

TR
Figure 27: ROI 25. Right Inferior Pareital Lobe
Figure 28: ROI 28

ROI 28 / -45 -49 43 / L Inferior Parietal
Figure 29: Two ROIs in left middle temporal gyrus,
Figure 30: a) ROI 28 (BA 40) in Experiment 4 b) parietal region identified by Becker et al. (1999) as anatomical locus of phonological store. Yellow areas identify BA 40. Orange areas identify BA 39.
Appendix A: Hamilton & Martin (2005)
Dissociations Among Tasks Involving Inhibition: A Single-Case Study

A. Cris Hamilton and Randi C. Martin

Rice University

Address correspondence to: Randi C. Martin, Psychology Department, MS-25, Rice University, P.O. Box 1892, Houston, Texas 77251, phone: 713-348-3417, e-mail: rmartin@rice.edu
Abstract

Recent theories of working memory have emphasized the role of inhibition in suppressing irrelevant information. Moreover, psychometric studies have reported that several inhibition tasks with very diverse requirements load on a single inhibition factor. A patient with left inferior frontal damage, patient ML, previously reported to have a semantic short-term memory deficit (Martin and He, 2004), showed evidence of difficulty with inhibition on short-term memory tasks. We investigated whether he would show evidence of inhibition difficulty on two verbal tasks (Stroop task and a recent negatives task) and two nonverbal tasks (nonverbal spatial Stroop task and antisaccade task). Patient ML was impaired on both verbal tasks but performed normally on the nonverbal tasks. ML’s data also represent a dissociation between Stroop and antisaccade performance, two tasks that load on a single factor in factor analytic studies. Implications of these data for theories of inhibition and executive function are discussed.
Dissociations Among Tasks of Inhibition: A Single-Case Study

Executive function is associated with complex planning and the ability to organize behavior. A core component of executive processes is inhibition - a term referring to the suppression of irrelevant information and overlearned or prepotent responses. Tasks commonly associated with inhibition measure a seemingly diverse set of abilities and one might question whether all such tasks measure the same cognitive processes. For example, it is not obvious that the processes necessary to inhibit reflexive eye-movements to sudden stimulus onsets in an antisaccade task are the same as those necessary to inhibit the tendency to name the word rather than the color in the Stroop task. Rather than a global inhibitory process, it may be necessary to further fractionate the processes involved in inhibition. The present study begins to address these issues concerning the nature of inhibitory processes by examining whether these processes dissociate in a brain-damaged patient with left inferior frontal damage.

Previous research with normal participants has suggested that several different tasks involving inhibition do share common processes. In an effort to delineate the components of executive function, Miyake, Friedman, Emerson, Witzki, Howerter and Wager (2000) collected data from a dozen tasks commonly assumed to involve executive processes and performed a latent variable analysis. This analysis yielded three factors which Miyake et al. identified as shifting, updating and inhibition. According to Miyake et al., shifting represents the ability to shift cognitive control between different tasks or routines. Updating is the ability to update and monitor representations in working memory. Finally,
inhibition is described as the ability to inhibit irrelevant information or prepotent responses. Tasks that loaded on the inhibition factor included the antisaccade task, stop signal task and Stroop task. Miyake et al. made no explicit predictions as to whether these factors may be associated with distinct and separable neural substrates. However, one might predict that a patient who performed poorly on one inhibition task would show a deficit on other tasks loading on the inhibition factor, but would not necessarily have difficulty with tasks loading on the other factors. In fact, Miyake et al. recommended that their results be carefully evaluated in neuropsychological populations in order to provide converging evidence. The present study takes this approach.

In addition to its fundamental role in executive function, inhibition and the ability to resist interference have been the subject of an increasing number of studies in the memory literature. Recent theories of working memory have emphasized resistance to interference as accounting for much of the individual variability in memory performance among normal individuals (see Rosen and Engle, 1997; Whitney, Arnett, Driver & Bud, 2001; Zacks & Hasher, 1994). Among college age participants, Engle and colleagues have also demonstrated a relationship between working-memory capacity and susceptibility to interference in a number of different paradigms (Rosen and Engle, 1997, 1998). In the aging literature, poorer performance for older adults on working memory tasks has been attributed to difficulty inhibiting irrelevant information. Studies by Hasher and colleagues (Chiappe, Hasher, and Siegel, 2000; May, Hasher & Kane, 1999) have shown that the typically poorer working memory performance of older adults
compared to younger adults could be eliminated if procedures were used that minimized proactive interference. Similarly, others have attempted to use susceptibility to proactive interference, specifically a susceptibility to intrusions of previously presented material, as a behavioral marker of various neurological deficits that serve to undermine memory (Rouleau, Imbault, Laframboise and Bédard, 2001).

Relevant to the question of working-memory capacity and susceptibility to interference is a series of recent studies employing a probe task (formulated by Monsell, 1978) to induce proactive interference. This task, subsequently referred to as the “recent negatives” task, consists of some trials in which the probe item matches an item in the preceding list but not on the current list. These studies (D’Esposito, Postle, Jonides, and Smith, 1999; Jonides et al., 2000; Jonides, Smith, Marshuetz, Koepppe and Reuter-Lorenz, 1998) have identified unique brain areas involved in overcoming proactive interference. Specifically, Jonides et al. (1998) have demonstrated increased activation in the left inferior prefrontal cortex, in the region of Brodmann’s Area (BA) 45. This activation in BA 45 is uniquely associated with items presumably affected by proactive interference. Moreover, Jonides et al. (2000) have reported that elderly participants show greater susceptibility to proactive interference when compared to younger adults and a corresponding absence of cortical activation in BA 45 during trials that promote such interference. In addition, Thompson-Schill et al. (2002) have demonstrated that a patient with damage to BA 45 showed exaggerated effects of proactive interference on this task.
A very different inhibition paradigm reported to be associated with working memory was first reported by Roberts, Hager and Heron (1994) and has been employed by Kane, Bleckley, Conway, and Engle (2001). This paradigm is notable because it employs a nonverbal task with no obvious memory requirements. Participants are required to maintain fixation and are then presented with a brief distractor appearing in the periphery of the display. Participants must inhibit the tendency to make a reflexive saccade to this distractor in order to successfully make an eye movement to a brief target presented on the opposite side. Reported correlations between working memory capacity and performance on the antisaccade task have prompted many to propose a role for attention in working memory capacity (see Kane et al. 2001, Roberts et al. 1994 and Mitchell, Macrae & Gilchrist, 2002). Specifically, working memory is thought to involve an ability to inhibit prepotent responses.

Given the reported relationships among inhibition and working memory performance, we have asked whether patients with short-term memory deficits will demonstrate deficits on measures of inhibition. Although our research program has previously emphasized the role of language representations in short-term memory (STM) (e.g., Martin and Lesch, 1996), we have recently become interested in a possible role for inhibition in conceptualizing some types of STM. Next we describe our model of STM organization.

Distinct patterns of patient performance on various tasks of STM support a dissociation between semantic and phonological capacities in STM. For example, patients with semantic STM deficits typically demonstrate no advantage of words
relative to nonwords (no lexicality effect), while patients with phonological short-term memory deficits fail to demonstrate normal phonological similarity effects. Patients with semantic short-term memory deficits perform better on a rhyme probe task relative to category probe task, while phonological patients are better on the category probe than rhyme probe task (Martin, Shelton & Yaffee, 1994). It is believed that patients with semantic short-term memory deficits are unable to accurately maintain lexical-semantic representations in STM and must therefore rely on phonological representations. On the other hand, cases with phonological STM deficits are unable to adequately maintain phonological information and depend upon lexical-semantic information during recall. Related patterns for brain-damaged patients have been reported by N. Martin and Saffran (1997).

Of relevance to the present study, Martin and Lesch (1996) reported a particularly interesting observation in patients with semantic STM deficits. These patients made numerous intrusions of previously presented items during serial recall tasks. This effect is surprising given that if patients’ short-term memory deficits are attributed to reduced capacity or an abnormally fast decay rate, one might expect fewer intrusions from previous lists. Interestingly, this paradoxical effect of excessive intrusions in patients with greatly limited STM has only been observed in cases of semantic short-term memory deficits. Here, we will address these intrusions as an apparent failure to inhibit processing of previously presented items, which produces exaggerated effects of proactive interference.

Anatomically, semantic STM deficits are associated with lesions involving inferior frontal areas (Martin & Freedman, 2001; Romani & Martin, 1999). This
observation is intriguing given recent fMRI data implicating inferior frontal areas with inhibitory functions such as semantic selection (Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997) and resolution of proactive interference (Jonides et al. 1998). Moreover, other interference effects possibly attributable to failures of inhibition have been reported – for example, Freedman, Martin and Biegler (2004) reported that patients with semantic STM deficits demonstrated greatly exaggerated interference effects when producing a conjoined noun phrase to describe two semantically related items compared to two unrelated items. These observations implicate a possible role for impaired inhibitory function in semantic STM deficits.

Given the prominent role of inhibition in many contemporary models of working memory and the aforementioned patient performance suggesting deficits of inhibitory function, the present study examines the performance of a semantic STM patient on tasks reported to be associated with inhibition. We included two tasks identified by Miyake et al. (2000) as loading on an inhibition factor – the standard Stroop task and the antisaccade task. We also employed a probe task designed to elicit proactive interference (Jonides, et al. 1998; Monsell, 1978). Finally, we included a nonverbal spatial Stroop task as a nonverbal analogue to the Stroop task.

The present study attempts to address several questions. First, would a patient who showed a semantic STM deficit and intrusions of previous list items on serial recall show a deficit on inhibition tasks not involving STM? If so, would this deficit on inhibition tasks include a deficit on the antisaccade task, a
nonverbal task that has been associated with working memory function? Second, would such a patient show an exaggerated effect on the recent negatives task that manipulates proactive interference? Third, would the patient generally show a difference in performance between verbal and nonverbal inhibition tasks? Given that the previously reported deficits of this patient are language-related (e.g., Martin & He, 2004) and that his lesion is in the left hemisphere, it is possible that any noted deficits with inhibition might be limited to the verbal domain.

PARTICIPANT BACKGROUND INFORMATION

Patient ML

Patient ML is a 62-year-old right-handed male with a left-hemisphere lesion resulting from a cerebrovascular accident (CVA) in 1990. He completed two years of college coursework and was employed as a draftsman prior to his CVA. His lesion includes the left frontal and parietal operculum, with atrophy noted in the left temporal operculum and mild diffuse atrophy.

ML has a verbal STM deficit as his accuracy of list recall is 77% lists correct for two word lists and 10% correct for three word lists (for lists recalled in order). As reported in several studies by Martin and colleagues (Freedman and Martin, 2001; Martin & He, 2004; Martin & Lesch, 1996), a number of features of ML’s short-term memory deficit support the conclusion that he has a deficit in semantic retention but shows better retention of phonological information. For instance, he shows no advantage for word recall over nonword recall whereas patients with a phonological retention deficit typically show a large advantage for words over nonwords (Martin & Lesch, 1996). It should be noted, however, that
ML's retention of phonological information does not appear to be normal, as his nonword span is reduced. However, his semantic retention is more impaired than his phonological retention. Freedman and Martin (2001) computed a composite z-score for performance on a number of measures tapping semantic and phonological retention. ML's semantic retention composite score was −2.59 whereas his composite phonological retention score was −.23. (In contrast, patient EA, with a phonological retention deficit, obtained a semantic composite score of 3.86 and a phonological composite score of −3.95).

ML demonstrates good comprehension of conversational speech on clinical exam but his narrative production is plagued by pausing, hesitations, word-finding difficulties and reduced phrase length. He demonstrates no apraxia of speech and his repetition of single words is excellent (96% correct).

ML's difficulties with semantic retention on STM tasks and with spontaneous speech cannot be attributed to difficulty in comprehending word meanings or producing individual words. As reported by Martin and Lesch (1996), ML scored above the mean for control participants on the Peabody Picture Vocabulary Task (Dunn & Dunn, 1981), a standardized test of word comprehension, using norms for 40-year-old participants (the highest age for which norms are available). On the Philadelphia Naming task (Roach, Schwartz, Martin, Grewal, & Brecher, 1996) which involves naming pictured objects, ML scored 98% correct, which was above the mean for control participants (96% correct). As reported by Martin and He (2004), he performed at a normal level of accuracy on unspeeded and speeded tasks examining living-nonliving judgments
and category judgments. His reaction times on the living-nonliving judgment were just outside the normal range, but on the category judgments were substantially longer than controls. However, as argued by Martin et al. (1994), the category judgments place some demand on semantic STM as the subject must retain the category label while deciding if the exemplar is a member of the category.

Despite ML’s excellent performance on most semantic tasks, the question is often raised as to whether his deficit in reaction times on some semantic tasks might reflect a disruption of semantic knowledge representations that underlies his poor short-term memory performance. It should be noted, however, that ML’s short-term memory performance contrasts substantially with that of patients who do have severe semantic deficits, that is, semantic dementia patients who have a degenerative disorder that selectively affects semantic knowledge (see Hodges, Patterson, Oxbury, & Funnell, 1992). In contrast to the primarily frontal damage for ML, these patients have progressive focal atrophy of the infero-temporal cortex, which is typically more pronounced in the left hemisphere (Snowden, Goulding, & Neary, 1989). Several such patients, who performed far below the normal range on semantic tasks such as picture naming and word comprehension, have been assessed on their short-term memory performance (e.g., Jefferies, Bateman, & Lambon Ralph, in press; Patterson, Graham, & Hodges, 1994; McCarthy & Warrington, 2001). Despite their severe semantic deficits, these patients’ spans appear to be larger than ML’s. For example, the semantic dementia patient AM reported by Knott, Patterson and Hodges (1997) scored
more than 10 standard deviations below the control mean in accuracy of picture
naming and word comprehension, yet scored 56% lists correct for four-item word
list recall. ML scored 0% lists correct for four-item lists composed of similar
materials. Also, the semantic dementia cases routinely perform better on word
than nonword lists, at least when the word lists are composed of words that the
patient still comprehends (see Jeffries et al., in press, for discussion). Although it
certainly would be valuable to compare ML and semantic dementia cases on the
same materials, the existing evidence suggests that it would be difficult to
attribute ML’s short-term memory pattern to a disruption of semantic knowledge.

Control Participants

All control participants were selected from a pool of older adults that
regularly participate in experiments at Rice University’s Brain and Language Lab.
Ages ranged from 53 years – 68 years with a mean age of 59 years (SD= 4.6).
Not all control participants were available for testing in all experiments. All
control participants had attended at least some college.

EXPERIMENTAL TASKS

Stroop Task

Stroop interference is a classic measure of interference and executive
control. We administered a Stroop task consisting of a single mixed block of
neutral, congruent and incongruent trials. We also assessed ML’s naming of
colors outside the Stroop context to examine the possibility that repeated naming
of colors might elicit additional interference for patient ML. This was necessary
given other data suggesting that semantically related stimuli also induce
interference in naming for patient ML (Freedman et al., 2004). A similar interference effect during repeated naming of semantically related exemplars has also been reported in healthy participants (Damian, Vigliocco, and Levelt, 2001; Kroll and Stewart, 1994; Pickard, Brandon, Hodgson, Schwartz and Thompson-Schill, 2003). Thus, we argue that ML’s reaction times for color naming outside the Stroop task provide a useful comparison to his naming on the neutral trials within the Stroop task. Such a comparison provides evidence on whether repeated sampling from the same semantic set causes interference in producing the appropriate color name. We also assessed whether patient ML demonstrated anomalous effects of facilitation. This was in accord with recent studies suggesting that facilitation may also reflect poor performance in the Stroop task (Kane and Engle, 2003). Moreover, given past objections to the use of traditional difference scores in measuring Stroop interference in the aging literature (see Verhaeghen and De Meersman, 1998), we also employed a more conservative measure of Stroop interference using log transforms.

Method

Materials and procedure. A computerized Stroop task (Stroop, 1935) provided by Akira Miyake (University of Toronto- St. George) was used. Participants named aloud the color of each stimulus (color words or rows of asterisks) as quickly as possible. A voice key was used to measure voice onset latencies. The participants saw 72 neutral trials consisting of asterisks appearing in either red, green, blue, orange, yellow, or purple. Sixty incongruent trials consisted of color words (red, green, blue, orange, yellow, or purple) that appeared in a different
color (red, green, blue, orange, yellow or purple). Additionally, 12 congruent trials were comprised of color words appearing in same color (red, green, blue, orange, yellow or purple). Trials were presented in a fixed pre-randomized order. All voice onset latencies beyond 2.5 standard deviations above the mean for each condition for each subject were removed for analysis. However, we also submitted untrimmed data to analysis using log transforms.

Results

Interference was defined as the mean difference in voice onset time between incongruent trials and neutral trials. Data for patient ML and 10 control participants appear in Table 1. At a group level, controls demonstrated the expected Stroop interference effect ($M=197\text{ms}$, $SD=62\text{ms}$), $t(9)=-10.0$, $p<.0001$. ML showed dramatically greater interference (969ms) than the controls, which was 12.4 standard deviations above the mean interference effect for controls and well outside the range (101-279 ms). Using items as a random factor, ML’s interference effect was statistically significant, $t(111)=-3.38$, $p<.001$.

Facilitation was defined as the mean difference in voice onset time between neutral trials and congruent trials. Control participants showed no facilitation effect, $t(9)=-1.21$, $p=.26$. In fact, control participants were 51ms slower to respond to congruent trials relative to neutral trials. Patient ML, however, displayed substantial facilitation of 581ms. This facilitation effect failed to reach statistical significance, $t(71)=1.27$, $p=.21$; however, there were only 11 congruent trials in this task, thereby limiting statistical power.
We also calculated interference by subtracting mean onset latencies for congruent trials from mean onset latencies for incongruent trials. Here the difference for patient ML was 1550ms, 9.18 standard deviations above the mean of 10 controls ($M=146\text{ms},\ SD=153\text{ms}$) and beyond controls’ range (range= - 176ms-330ms).

Given the concerns about comparing difference scores for individuals who are showing large differences in mean reaction times (Verhaeghen and De Meersman, 1998), interference was also calculated using log transforms of untrimmed data. These data are presented in Table 2. Using transformed onset latencies, ML’s interference effect was .17 which was 1.9 standard deviations above the mean interference for control participants ($M=.086,\ SD=.043$). Furthermore, ML’s interference was outside the range of controls (range= -.009 -.14). Using the same method for the incongruent-congruent effect, ML’s interference effect was .258 and 3.55 standard deviations above the control mean ($M=.057,\ SD=.057$). This effect was also beyond the range of the 10 control participants (range= -.06-.156). However, using log transforms, patient ML showed no facilitation effect. This was owing to the fact that log transforms used untrimmed data and ML had a single onset latency of over 11000ms in the untrimmed data in the congruent condition, which skewed the mean. This trial was considered an outlier (greater than 2.5 standard deviations above the mean) in the conventional analysis of untransformed data.

Accuracy data are also reported in Table 1. The difference in accuracy between neutral and incongruent trials for controls was statistically significant,
\[ t(9)=3.09, p=.01, \] but the difference between neutral and congruent trials was not, 
\[ t(9)= -2.07, p=.068. \] There were no statistically significant differences in ML’s accuracy across conditions.

Color Naming outside Stroop Task

Relative to naming in other tasks, ML’s naming of colors was unusually long in the Stroop task, even on neutral trials, which merely require naming of a colored row of asterisks. We suspected that the repeated naming of six color words might be sufficient to produce interference in production for ML. That is, all color words are semantically related and one might expect that repeated naming of words from this semantic category would lead to a high level of semantic activation for each and difficulty in selecting the correct response from competitors. This hypothesis is supported by the observation that ML shows interference during production when naming pairs of semantically related pictures relative to unrelated pictures (Freedman et al., 2004). Related effects have been reported by Kroll and Stewart (1994) and Damian et al. (2001) for non-brain damaged participants in naming a series of pictures when the pictures were either from the same semantic category or from different categories (randomly intermixed). For example, Kroll and Stewart (1994) found that healthy participants were 36 ms slower to name categorized lists of pictures relative to randomized lists of pictures. Damian et al. (2001) replicated this effect after controlling for possible confounding variables such as visual similarity. These investigators have attributed the category interference effect to difficulty in selecting the correct response when closely related competitors have been highly
activated through their recent production in the categorized lists. Selection from competitors might plausibly involve the inhibition of the competitors (Wheeldon & Monsell, 1994). Consequently, a difficulty with inhibition for ML might lead to long times even on the neutral trials in the Stroop paradigm. In order to investigate this possibility, we compared ML's naming of color patches in the context of naming line drawings with his color naming in the Stroop task. This allowed comparison of ML's naming in a context that does not require repeated production of words from the same category.

*Method*

*Materials and Procedure.* Six colors (red, green, blue, orange, yellow or purple) were randomly presented with 30 other pictures taken from five additional categories (totaling 6 categories). The five other categories were fruits/vegetables, animals, clothes, musical instruments and transportation vehicles. There were six exemplars from each category for a total of 36 pictures. Each image was presented on a computer monitor and a voice key recorded voice onset latency. With exception of the color patches, all pictures were line drawings. ML and five additional control participants were tested on this task.

*Results*

Data for the naming task appear in Table 3. ML was 1370ms faster to name colors in the naming task ($M=1100\text{ms}$, $SD=175.4$, range=$850\text{ms}-1284\text{ms}$) compared to naming colors in the Stroop task ($M=2470\text{ms}$, $SD=1269$, range=$1075\text{ms}-5712$), $t(149)=2.396$, $p=.02$. In contrast, five control participants showed no difference between naming neutral trials in the Stroop task ($M=818\text{ms}$,
range = 662ms-1110ms) and naming color patches in the context of a naming task (M = 804ms, range = 636ms-902ms), t(4) = .24, p = .82.

Discussion: Stroop Effect and Color Naming

Using the traditional method to calculate Stroop interference (mean voice onset for Incongruent trials minus mean voice onset for Neutral trials), ML showed an exaggerated interference effect almost five times greater than the mean for controls. However, given past objections to the use of traditional difference scores in measuring Stroop interference in the aging literature (see Verhaeghen and De Meersman, 1998), we chose to further examine ML’s Stroop performance using a more conservative log transform. Using log transformed data, ML’s interference effect was still beyond the range of all control participants. Thus, this patient with a semantic STM deficit did show evidence of a deficit in inhibition on the Stroop task.

In addition to his exaggerated Stroop effect, ML also differed from controls in showing much longer reaction times to name colors in the Stroop task (neutral trials) than to name colors when color patches were mixed with line drawings. Controls showed no difference between naming colors in the Stroop task and naming colors mixed with line drawings in a naming task. As discussed earlier, similar interference effects for the production of semantically related words in close succession have been reported for ML (Freedman et al., 2004) and for young normal subjects (see Damian et al., 2001; Kroll and Stewart, 1994). ML’s much longer times for the neutral condition in the Stroop task than in naming color patches mixed with other items could also be attributed to a
difficulty with inhibition – that is, difficulty inhibiting competitors from the same semantic category which are strongly activated because they have been recently produced on other trials.

To summarize, relative to controls, ML demonstrated an exaggerated interference effect in the traditional Stroop task. Moreover, ML showed much longer onset latencies for color naming in the Stroop task when compared to naming of colors outside the Stroop task. Controls did not show this effect. The longer color naming in the context of the Stroop task could also be plausibly attributed to a difficulty with inhibition, that is, difficulty inhibiting other color names when they are all highly activated during the Stroop task.

Nonverbal Spatial Stroop Task

A nonverbal analogue to the Stroop task was designed to further assess ML’s performance in tasks that require resolution of response conflict and presumably inhibition and to determine if he would show similar difficulty with inhibition in verbal and nonverbal domains. For a similar task, see Clark and Brownell (1975) and Experiment 3 of Lu and Proctor (1994).

Method

Materials and Procedure. Participants saw arrows pointing either left or right and were asked to press a key with the middle finger (for left) or index finger (for right) of their left hand corresponding to each direction. The left hand was used because of ML’s mild hemiplegia on the right. The arrows appeared in one of three positions – either at the left of the screen, at the right of the screen or directly in the middle of the screen. The arrows and a fixation point appeared in a
rectangular box 9.25" wide and 1.75" in height. Similar to the Stroop task, there were congruent trials (a right pointing arrow on the right side of the display), neutral trials (either right or left pointing arrows appearing in the middle of the display) and incongruent trials (a left pointing arrow appearing on the right side of the display). There were 80 trials for each condition for a total of 240 trials. Interference was calculated by subtracting reaction times of neutral trials from reaction times to incongruent trials. ML was administered this task twice with several months intervening between the two administrations. Data were combined, resulting in 480 total trials. Fifteen control participants performed this task.

Results

Results for patient ML and 15 control participants on the nonverbal spatial Stroop task appear in Table 4. Control participants showed a 75 ms interference effect for incongruent trials relative to neutral trials, a difference which was significant at the group level, \( t(14)= 10.41, p<.0001, SD = 27.99, \) range= 27ms-129ms. Patient ML showed a significant interference effect of 106ms, \( t(309)=7.66, p<.0001 \). This interference effect was greater than the control mean, but only 1 standard deviation above the mean for controls and well within their range. ML’s mean reaction time for both incongruent and neutral trials were also within the range of healthy controls. There were no differences in accuracy for either controls or ML.

Although ML clearly performed at a normal level on this spatial Stroop task, the reaction time difference for controls between the conflict and neutral
conditions was less than half than that of the verbal Stroop task, raising the possibility that the spatial Stroop task is simply easier. However, the standard deviation of the difference scores between the two conditions was also substantially smaller in the spatial than in the verbal Stroop task. Effect sizes for reaction times for the control participants were very large for both experiments, and fairly similar in magnitude (Cohen’s $d = 3.33$ for the verbal Stroop and 2.78 for the spatial Stroop). (See Rosnow & Rosenthal, 2003, for calculation of Cohen’s $d$ for repeated measures.) Thus, in terms of effect size, there is little evidence that the spatial task was substantially easier for controls.

**Discussion: Nonverbal Spatial Stroop Task**

On this task, selected as a nonverbal analogue to the classic Stroop task, ML showed performance well within range of control subjects. Thus, resolving response conflict in this nonverbal task did not prove to be as difficult as resolving conflict in the verbal Stroop task. While some may argue that the task used here differs fundamentally from the classical Stroop task in many ways, it is important to note that similar tasks have traditionally been considered to represent a variation of Stroop interference (see MacLeod, 1991, for a review). The difference in ML’s performance relative to controls for the standard Stroop and the nonverbal spatial Stroop suggests that ML’s deficit in inhibition may be limited to the verbal domain. His performance on the next task, the antisaccade task, provided another means of testing this possibility. Also, given that the antisaccade task is generally a difficult task, with accuracy in the range of 79-88% for undergraduate participants (Roberts et al., 1994), data from this task should
provide further evidence on whether the ML’s performance on inhibition tasks relates to the difficulty of the task or to whether the task is in the verbal or nonverbal domain.

Antisaccade Task

The antisaccade task used here was provided by Akira Miyake and was adapted from Roberts et al. (1994). Given previous studies reporting correlations between antisaccade performance and working memory function, this task was of particular interest for patient ML.

Method

Materials and procedure. Each trial began with a fixation point that was presented for a variable length. Specifically, the fixation appeared for one of nine lengths spaced at 250ms intervals between 1500 and 3500ms. Next, a cue appeared on one side of the screen, 3.4” from fixation for 175ms (this was more brief than the cue in Miyake et al. 2000, which was 225ms). The cue was followed by a target appearing on the opposite side of the screen, 3.4” from fixation, for 150ms. The cue was a small black square, while the target was an arrow that pointed to the left, up or right. After 150ms the target was replaced by cross-hatching to mask after-images which might aid participants in identifying the target. Correct identification of the target required the subject to press a key corresponding to left, right, and up. In the antisaccade task, participants must resist making a reflexive saccade to the initial cue in order to detect the briefly presented target on the opposite side of the screen. If the participant makes an initial saccade to the cue, the target is difficult, if not impossible, to identify due
to the brief presentation of the target. The task began with 22 practice trials, followed by 90 target trials presented in a fixed, pre-randomized order. Patient ML was administered the antisaccade task twice over a 6 month span and combined data from these two administrations are reported below.

We also administered a prosaccade task with the same parameters described above. However, in the prosaccade task, the cue always predicted the target location. Thus, the participant is to direct an eye movement toward the cue in order to detect the target. Different control participants were tested in the antisaccade and prosaccade tasks.

Results

As expected, 12 controls were substantially less accurate on the antisaccade (72% correct) than on the prosaccade task (97% correct). On the antisaccade task, ML’s accuracy was above the mean for controls (ML = 80%, control M= 72%, control SD = 11.8, control range= 59%-94%). His mean reaction time (M=724ms) was well within the range of the controls (control M=771ms, SD=236, control range=492ms-1394ms). ML’s accuracy on the prosaccade task (Table 4) was 93% and within the range of tested controls (M=97%, range= 93%-100%). ML’s mean reaction time on the prosaccade task was 521ms, which was within range of controls (M=485ms, SD=80, range=386ms – 608ms). Thus, ML performed in a normal fashion on these tasks.

Discussion: Antisaccade Task

As with the nonverbal spatial Stroop, ML performed at a normal level on the antisaccade task. As anticipated, the antisaccade task was a quite difficult task
for controls, with mean accuracy of only 72% compared to 97% for the prosaccade task. This 27% difference in accuracy between the two conditions was substantially larger than that between the incongruent and neutral conditions (6.3%) and between the incongruent and congruent conditions (4.3%) in the verbal Stroop task. The reaction time difference of 207 ms between the antisaccade on prosaccade tasks was also somewhat larger than that observed between the incongruent and neutral conditions (197 ms) and between the incongruent and congruent conditions (126 ms) in the verbal Stroop task. Consequently, it is highly unlikely that ML’s excellent performance on the antisaccade task could be attributed to its being a generally easier task than the standard verbal Stroop task. Thus, the results from both the antisaccade and the spatial Stroop task compared with the standard Stroop task provide evidence that ML’s difficulty with inhibition is limited to the verbal domain. This hypothesis will be further tested through use of the fourth task, the recent negatives task.

ML’s normal level of performance on the antisaccade task would seem to go against findings from normal participants indicating a relation between antisaccade performance and working memory capacity (Kane et al., 2001). Further discussion of this point will be delayed until the General Discussion.

Proactive Interference: Recent Negatives Task

To further assess patient ML’s ability to inhibit irrelevant information in the verbal domain, we administered a task designed to elicit proactive interference in a test of short-term memory. This task is of great interest, given the aforementioned experiments that have demonstrated correlations between short-
term memory performance and susceptibility to proactive interference in other memory paradigms (e.g., Chiappe et al., 2000, Rosen and Engle, 1998).

A memory probe paradigm developed by Monsell (1978) was employed. In this recent negatives task, the subject sees a list of items presented serially, followed by a probe. The subject responds “yes” or “no” according to whether the probe appeared in the list. A recent negative probe trial is one in which the negative probe did not appear in the present list, but appeared in the list presented immediately before the present list. For the other negative probe trials, the negative probe did not appear in the previous list, but appeared three lists back – this was considered a nonrecent negative probe. The same recent v. nonrecent manipulation was applied to the positive probes - a recent positive trial included a probe that appeared in the present list as well the previous list. In a nonrecent positive trial, the probe appeared in both the present list and in a list three trials back. The principal effect of interest in these manipulations is the contrast of recent negative v. nonrecent negative probes - reaction times are expected to be longer and accuracy lower for the recent negatives compared to nonrecent negatives. In other words, it should take longer to correctly reject a negative probe if it appeared in an immediately previous list.

Studies using Monsell’s recent negatives paradigm (D’Esposito et al., 1999; Jonides et al. 1998) have attributed these longer reaction times to involuntary persistence of previously presented lists or proactive interference. Jonides et al. (1998) assume that an inhibitory mechanism must be necessary to overcome this involuntary persistence of items presented in previous lists that
match a negative probe in the current list. The time needed to resolve this conflict is argued to result in the longer reaction times for the recent negative probes. In the past, we have attributed semantic STM deficits to overly rapid decay of semantic representations (Martin & Lesch, 1996). If so, then one would expect a patient with a semantic STM to show less interference than normal from items in a previous list. On the other hand, if the deficit is in inhibiting irrelevant representations, then an exaggerated interference effect would be predicted.

**Method**

*Materials and procedure.* Three words were presented serially, followed by a probe word. Each item was drawn from a set of 16 words. Each word was presented for 750 ms followed by a 100 ms inter-stimulus interval. The third word was followed by 100 ms interval and then a row of "****" was presented for 400 ms followed by a probe word presented for 600 ms. The probe word was immediately followed by blank screen which remained until the subject responded. Essentially, the probe followed the last word in the series by 500 ms. The participant was instructed to respond as quickly as possible after the probe was presented. After the participant’s response, the next trial began 250 ms later with the presentation of the fixation point. There were 18 trials in each condition (recent and nonrecent conditions for both positive and negative trials). Stimuli were presented in a fixed pre-randomized order to form the recent and nonrecent conditions. Sixteen words were substituted for the 16 letter stimuli used in previously published research with this paradigm. (Words would ultimately allow a greater number of manipulations to be explored in this paradigm. For example,
in other unpublished studies we have manipulated the number of unique stimuli presented in the task.)

**Results**

All reaction times beyond 2.5 standard deviations above the mean for each condition for each subject were removed from analysis of control participants and patient ML. The data for ML and fourteen controls appear in Table 5. In an analysis of group data, control participants demonstrated a significant interference effect for reaction time (M=91ms), t(13)=3.135, p=.008, SD=108ms. In addition, control participants were significantly less accurate on recent negative trials than on nonrecent negative trials (94.7% vs. 98.9% correct), t(13)=-2.97, p=.01. ML demonstrated a substantial (731ms) interference effect in reaction time for recent vs. nonrecent negatives. Although this difference failed to reach significance, t(32)=1.36, p=.18, ML’s interference effect was 5.9 standard deviations above the mean interference effect for controls and substantially outside their range (-74 ms to 337 ms). ML did show a statistically significant interference effect in accuracy. He was much worse on recent negative trials (accuracy =62.5%) than on nonrecent negative trials (accuracy = 87.5%). This 25% difference was statistically significant, t(46)=-2.04, p=.046, and far outside the range for controls (0% - 13%).

We calculated interference effects using a more conservative log transform of untrimmed data. Using this method, ML’s interference effect (.145) was 3.12 standard deviations above the mean interference for controls (.037) and well beyond the range of controls (-.027 -.09).
Discussion: Recent Negatives Task

Although exaggerated interference effects in a patient with impaired STM at first appear paradoxical, Thompson-Schill et al. (2002) have reported another such patient. Based on neuroimaging findings from Jonides et al. (1998, 2000) and D'Esposito et al. (1999) implicating BA 45 in resolution of proactive interference, Thompson-Schill et al. (2002) tested a patient with a lesion that included BA 45. This patient showed exaggerated effects of proactive interference, similar to the effect reported in this paper. Also of note, we have examined other patients with memory spans comparable to patient ML. In these patients, most primarily with phonological STM deficits, we have failed to find exaggerated interference effects. One hypothesis (currently being tested by our laboratory) is that phonological short-term memory deficits are better characterized as rapid decay of representations in short-term memory, whereas semantic short-term memory deficits involve failures to inhibit representations that have been activated.

GENERAL DISCUSSION

While executive function has been a topic of much discussion in the neuropsychological literature, fewer patient studies have attempted to better delineate the components of this complex cognitive capacity. Chief among these components is inhibition. The principal concern of this paper was whether several tasks that are commonly assumed to require inhibition – specifically, the Stroop task, antisaccade task, nonverbal spatial Stroop task, and a proactive interference task – are tapping similar inhibition processes. Given the factor
analytic studies by Miyake et al. (2000) and Friedman and Miyake (2004), one might expect that a patient showing a deficit on one inhibition task, due specifically to problems with inhibition, would show deficits on other tasks loading on the inhibition factor. We were particularly interested in the antisaccade task, given that previous research has demonstrated a correlation between antisaccade performance and working memory/STM performance in normal participants.

Data presented in this paper demonstrate dissociations among these tasks. Here we presented a patient, ML, with a semantic short-term memory deficit, who showed exaggerated effects of interference on the standard color naming Stroop task and recent negatives task, while demonstrating normal performance on a nonverbal spatial Stroop task and on the antisaccade task. While the verbal/nonverbal dissociation suggested here must be considered with caution, two important points must be made. First, ML’s observed deficits are not likely attributable to disordered semantic representations. As detailed in the patient’s background information, patient ML performs remarkably well on word processing tasks. His single word comprehension and production are at a normal level in terms of accuracy, and even his reaction times on semantic tasks are near normal for some tasks, though not when the task has a short-term memory component. Thus, while the present data suggest that ML may have particular difficulty with maintenance and manipulation of verbal representations, these representations are largely intact. His performance contrasts with semantic
dementia cases who show much worse semantic processing but typically better short-term memory (Jefferies et al., in press).

A second issue is whether the degree of dissociation between verbal and nonverbal tasks for ML is much greater than might be observed in the normal population given that the correlations among inhibition tasks for control subjects tend to be small (e.g., .18 - .20 in the Miyake et al. study). Of course, if the correlation in the normal population were zero, there would be little justification for assuming a common underlying inhibition factor in the first place. Let us assume that the correlation between the standard verbal Stroop effect and the nonverbal spatial Stroop effect is .2 for controls matched to ML in age and education. Given that ML scored 1 standard deviation above the mean in terms of his nonverbal spatial Stroop and assuming N=15, the 95% confidence interval about the predicted value for the standard Stroop task, in terms of standard scores,

\[ \text{.2} \pm (2.14) \times \text{standard error of the predicted scores} \]

\[ = \text{.2} \pm (2.14)(1.01) \]

\[ = [-1.97, 2.37] \]

ML's verbal Stroop effect in terms of raw reaction times was 12.4 standard deviations above the mean - clearly outside of this range. In terms of log transforms, although his effect was within this interval when computed in terms of the difference between incongruent and neutral trials (1.9 standard deviations above the mean), it was outside this interval when computed in terms of the difference between incongruent and congruent trials (3.5 standard deviations
above the mean). Assuming a .2 correlation between the standard Stroop effect
and antisaccade performance, similar computations in predicting the verbal Stroop
effect from antisaccade performance would show that ML's verbal Stroop effect
was even farther outside the 95% confidence interval, given that his reaction time
difference between the pro- and antisaccade task was only slightly greater than the
mean for controls and his error difference was smaller than the mean for controls.
Furthermore, the probability that a normal individual would show the pattern
shown by ML would be substantially smaller than any of these individual
computations when one took into account the probability of showing normal
performance on both of the nonverbal tasks and performance outside the normal
range on both of the verbal tasks.

Finally, the present data raise the intriguing possibility that semantic short-
term memory deficits may uniquely involve deficits of inhibition. Whereas
phonological short-term memory deficits may indeed result from rapid decay of
phonological representations, semantic short-term memory deficits may be
uniquely characterized by failures of inhibition in the verbal domain.
Experiments presently being conducted in our laboratory attempt to establish to
what extent this difficulty with inhibition applies to all verbal materials or perhaps
only stimuli with a semantic component. Similar experiments are to be
conducted with patients with phonological short-term memory deficits to
determine whether further dissociations may be established.

Implications for the Relation of Antisaccade Performance and Working Memory
Patient ML’s performance on the antisaccade task is especially noteworthy given the previous research of Kane et al. (2001), Roberts et al. (1994) and Mitchell et al. (2002) relating working memory ability to antisaccade performance. Specifically, Kane et al. (2001) have claimed that attentional control is central to working memory performance and that performance on the antisaccade task reflects individuals’ attentional control ability. Mitchell et al. (2002) describe the relation in a somewhat opposite fashion – namely, that working memory operations play a role in suppressing the prepotent behavioral response of attending toward the cued location. In either case, one might expect that a patient with reduced working memory capacity would be impaired on the antisaccade task. Although patient ML shows deficits on verbal short-term memory tasks, his performance was normal on the antisaccade task, which would seem to complicate these authors’ claims. It might be argued, however, that ML’s normal performance on the antisaccade task is not relevant to claims about working memory as his span deficit could be attributed purely to a STM deficit rather than to a working memory deficit. That is, in the working memory theory postulated by Engle, Tuholski, Laughlin and Conway (1999), working memory is composed of STM storage plus central executive function and patient ML’s deficit is restricted to the STM storage component. Certainly ML’s poorer retention of semantic than phonological information on span tasks suggests that he does not have an over-arching deficit in working memory that affects all working memory tasks equally.
However, Kane and Engle and colleagues have argued that both the antisaccade task (Kane, Bleckley, Conway, & Engle, 2001) and the Stroop task (Kane & Engle, 2003) tap attentional control that is vital to the executive component of working memory. The present results showed that ML performed normally on the antisaccade task but showed abnormally large interference on the verbal Stroop task. The findings for ML suggest that attentional control is not a unitary mechanism localized to dorsolateral prefrontal cortex as proposed by Kane and Engle (2002). The present data suggest that the correlations reported in these individual differences studies are not attributable to a shared neural substrate, but rather some other factor. This point is discussed further at the end of the General Discussion.

Implications for Organization of Executive Function and Inhibition

Miyake et al. (2000) left open the question of whether the three target factors identified in their study would generalize to neurologically impaired participants. They also left open the possibility that the target functions could be decomposed into more basic component processes. The present study suggests that Miyake’s inhibition factor can indeed be further decomposed into finer-grained functions, as the Stroop and antisaccade tasks loaded on the same factor in the Miyake et al. (2000) study. Of further interest, in another study using similar factor analytic methodology and a number of additional inhibition tasks to examine the possibility of multiple inhibition factors, Friedman and Miyake (2004) reported that both the antisaccade task and Stroop task load on the same
factor, which they term "prepotent response inhibition". Here, again, data from our case study show dissociations of tasks loading on the same factor.

In addition to this dissociation among tasks that load on the same factor in Miyake et al. (2000) and a dissociation among antisaccade performance and STM performance, Patient ML also shows seemingly paradoxical interference effects on tasks of proactive interference. Given his restricted memory span of 2.5 items one might expect ML to show very little proactive interference. That is, if deficits of short-term memory are conceptualized as rapid decay of items in a STM or WM buffer, he should show little proactive interference from preceding trials. The larger proactive interference effect suggests that ML's short-term memory deficit may be better conceptualized as a problem of persisting activation resulting from deficient inhibitory processes. Once representations are activated in ML's short-term memory, he appears to have difficulty in subsequently suppressing these representations. This is also consistent with the intrusions that ML produces during serial recall tasks.

The dissociation between verbal and nonverbal tasks reported here are consistent with a recent fMRI study demonstrating dissociable networks for control of verbal and visuospatial tasks (Stephan, Marshall, Friston, Rowe, Ritzl, Zilles and Fink, 2003). In this study, control processes in a verbal letter detection task were reported to preferentially activate a network in the left hemisphere, while control processes involved in a visuospatial task (using the same stimuli) preferentially activated a network in the right hemisphere. More importantly, the network activated in the verbal letter task involved a left region of the anterior
cingulate and the left inferior frontal gyrus. In contrast, the visuospatial task activated a network involving the right anterior cingulate and intraparietal sulcus. The authors suggest that both content and control processes for any task might prove to be lateralized in the same hemisphere. Thus, patient ML’s damage to the left inferior frontal gyrus undermines control processes for verbal information, while his spared right hemisphere leaves his performance of visuospatial tasks unimpaired.

The dissociations presented here also support recent neuroimaging data by Nelson, Reuter-Lorenz, Sylvester, Jonides and Smith (2003) indicating different areas of activation for inhibition of previously executed motor responses and inhibition of previously presented mnemonic stimuli. Our data would appear to support such a dissociation. For example, ML shows no deficit on the nonverbal spatial Stroop task which requires resolution of motor response conflict, while showing exaggerated interference on the recent negatives task which requires inhibition of previously presented items in memory. This further suggests the need for finer-grained distinctions in models of executive function.

While the present data serve as interesting tests of factor analytic studies of executive function, the behavioral dissociations correspond with much of the neuroimaging literature examining these specific tasks. First, it is interesting to note that ML’s lesion is located in the inferior frontal gyrus, including the frontal and parietal opercula. This area is very near the site of a lesion reported in Patient RC in Thompson-Schill et al. (2002), whom showed similar effects of exaggerated proactive interference. Also, this corresponds with the lesions of
patients with semantic STM deficits reported by Martin and colleagues (Martin & Freedman, 2001; Romani & Martin, 1999). This leads to the speculation that perhaps semantic STM deficits may be uniquely characterized by deficits of inhibition and interference.

Neuroimaging data are also potentially useful in explaining the behavioral dissociations among the antisaccade and proactive interference tasks reported in this paper. For example, as mentioned previously, imaging studies have reported activations in BA 45 for the proactive interference task (D’Esposito et al. 1999; Jonides et al., 1998, 2000). However, a recent neuroimaging study using the antisaccade task (Curtis and D’Esposito, 2003) reported activity in the pre-supplementary eye fields in the preparatory period before execution of saccades away from a cue. This area is located along the medial wall of the frontal cortex near the interhemispheric fissure- well superior and medial to the areas reported in neuroimaging tasks of proactive interference task (Jonides et al., 1998), as well as ML’s lesion in the inferior frontal gyrus. Based on neuroimaging data, it seems less than surprising that we observed the dissociations reported in the present study (see also Jonides et al. 2002).

The present data illustrate the necessity of recognizing distinctions among tasks assumed to tap “inhibition” and the utility of patient studies in complementing neuroimaging data in cognitive neuroscience. While carefully designed neuroimaging studies have proven invaluable in elucidating the neural substrate of specific cognitive processes, patient studies are useful in corroborating conclusions concerning the functional significance of these data.
Discrepancies Between Factor Analytic Studies and Patient Data

Finally, we address why factor analytic studies and patient studies might yield different conclusions with regard to the organization of executive function. One alternative would be to simply claim that the data from one subject, despite the degree of dissociation, do not constitute strong enough data to invalidate the results from a large number of normal subjects. However, we offer this data as a first attempt to assess the generalizability of the Miyake et al. (2000) data to “neuropsychological populations” – as was suggested by Miyake et al. (p. 91). Of course, further case studies that replicated this dissociation would be important to obtain. In fact, we do not intend to claim that the present data are incompatible with the factor analytic studies. We do suggest, however, that the present data elucidate the possible source of shared variance among these tasks. To be explicit, based on patient ML’s data, shared variance does not appear to be due to all inhibition tasks sharing a common neural substrate. Although Miyake et al. (2000) and Friedman and Miyake (2004) do not claim that a common neural substrate is the source of shared variance in their study, we feel that it is a reasonable possibility to address.

If a shared neural substrate is not the source of the shared variance, then what might be? One possibility is that variations in performance among normal participants result from variations in the level of neurotransmitters such as dopamine and norepinephrine. Neurotransmitter function might affect the operation of a number of different frontal brain areas. Thus, even though different cortical areas are involved in different executive function tasks, the
shared variance among tasks observed for normal participants might relate to activity of a small number of neurotransmitters. Consequently, tasks supported by diverse cortical areas may load on the same factor because of intercorrelations stemming from shared dependence on neurotransmitter function. The different neural substrates involved in the different tasks could be differentially affected by brain damage – thus, explaining the dissociations reported here. This hypothesis is appealing, given deficits of executive function have often been reported with Parkinson’s disease and schizophrenia, both conditions characterized by abnormal dopamine function. Although dopamine function and cognitive performance have been examined in both humans and nonhuman primates, the relationship between dopamine and prefrontal functioning is complex. Effects may differ depending upon the subreceptors examined and tasks performed (for review, see Arnsten and Robbins, 2002). However, psychopharmacological studies examining the cognitive effects of dopamine agonists and antagonists in humans could begin to address these issues in the future (see Kimberg and D’Esposito, 2003 for a recent example of a psychopharmacological approach).
Refernces


Proceedings of National Academy of Sciences, 95, 8410-8413.

controlled-attention view of working-memory capacity. Journal of 
Experimental Psychology: General, 130 (2), 169-183.

Kane, M.J. & Engle, R.W. (2002). The role of prefrontal cortex in working-
memory capacity, executive attention, and general fluid intelligence: an
individual-differences perspective. Psychonomic Bulletin and Review, 9
(4), 637-671.

attention: The contributions of goal neglect, response competition, and
task set to Stroop interference. Journal of Experimental Psychology:
General, 132, 1, 47-70.

receptor agonist pergolide. Neuropsychologia, 41, 1020-1027.

Knott, R., Patterson, K., & Hodges, J. R. (1997). Lexical and semantic binding
effects in short-term memory: Evidence from semantic dementia.
Cognitive Neuropsychology, 14, 1165-1218

naming: Evidence for asymmetric connections between bilingual memory

dimension as a function of the relevant stimulus dimension. Journal of
Experimental Psychology: Human Perception and Performance, 20, 286-298.


Philadelphia naming test: Scoring and rationale. Clinical Aphasiology, 24, 121-134.


Authors’ Notes

This research was supported by NIH grant DC-00218 to Randi Martin at Rice University. The authors wish to acknowledge the assistance of Meredith Knight and Jill Henderson in collecting control data for this project and Naomi Friedman and Akira Miyake for providing their experimental tasks.
Footnotes

1. The failure to find longer times in the neutral condition of the Stroop task than in the naming of color patches mixed with other items for the controls could not be considered a failure to replicate the findings of Kroll and Stewart (1994) or Damian et al. (2001) with older participants. There were many differences between the two conditions in terms of the stimuli presented (rows of colored XXXs vs. color patches), numbers of practice trials, numbers of items, etc.

2. It should be noted that some previous research using the antisaccade task has employed other measures of performance – namely, eye-movement data collected by eye tracking equipment. Although we have not incorporated eye-movement data into this study, we feel confident in our assessment of patient ML’s antisaccade performance based on target identification. Previous research by Roberts et al. (1994) has demonstrated that target identification data parallels data from eye movements and that both measures are similarly affected by demands of dual task methodologies. Therefore, we suspect that ML would show similar performance relative to controls if eye-movement data were collected. In fact, ML’s accuracy is comparable to the mean of young controls reported in Table 2 of Roberts et al. (1994). This was the case even though we used a shorter cue, a modification made by Miyake to increase variance, and which presumably made the task more difficult.

3. A reviewer commented that ML could be tested on some spatial working memory task to provide evidence on whether he has a generalized
working memory deficit. We predict that he would do well on such a task, but such testing has yet to be carried out.
Table 1

Reaction Time (RT, in Milliseconds), Interference (in Milliseconds) and Accuracy (% Correct) for Stroop Task – Controls and Patient ML

<table>
<thead>
<tr>
<th>Participant</th>
<th>Congruent</th>
<th>Neutral</th>
<th>Incongruent</th>
<th>Interference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>828</td>
<td>777</td>
<td>974</td>
<td>197*</td>
</tr>
<tr>
<td>% Correct</td>
<td>98</td>
<td>100</td>
<td>93.7</td>
<td></td>
</tr>
<tr>
<td>Patient ML</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>1889</td>
<td>2470</td>
<td>3449</td>
<td>969*</td>
</tr>
<tr>
<td>% Correct</td>
<td>92</td>
<td>89</td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

* *p < .001
Table 2

Log Transformed Reaction Time Data for Stroop Task – Controls and Patient ML

<table>
<thead>
<tr>
<th>Participant</th>
<th>Congruent</th>
<th>Neutral</th>
<th>Incongruent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>2.907</td>
<td>2.878</td>
<td>2.964</td>
</tr>
<tr>
<td>Patient ML</td>
<td>3.258</td>
<td>3.346</td>
<td>3.516</td>
</tr>
</tbody>
</table>
Table 3

Onset Latencies for Naming Task (in Milliseconds) – Controls and Patient ML

<table>
<thead>
<tr>
<th>Participant</th>
<th>Colors</th>
<th>Animals</th>
<th>Fruits/Vegetables</th>
<th>Clothes</th>
<th>Musical Instruments</th>
<th>Transportation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>804</td>
<td>739</td>
<td>753</td>
<td>751</td>
<td>811</td>
<td>848</td>
</tr>
<tr>
<td>ML</td>
<td>1100</td>
<td>900</td>
<td>964</td>
<td>1119</td>
<td>879</td>
<td>1048</td>
</tr>
</tbody>
</table>
Table 4
Reaction Times (in Milliseconds), Interference (in Milliseconds) and Accuracy (% Correct) on Nonverbal Spatial Stroop Task – Controls and Patient ML

<table>
<thead>
<tr>
<th>Participant</th>
<th>Congruent</th>
<th>Neutral</th>
<th>Incongruent</th>
<th>Interference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>593</td>
<td>611</td>
<td>686</td>
<td>75</td>
</tr>
<tr>
<td>Range</td>
<td>(566-993)</td>
<td>(476-880)</td>
<td>(566-993)</td>
<td>(27-129)</td>
</tr>
<tr>
<td>SD</td>
<td>99</td>
<td>94</td>
<td>101</td>
<td>28</td>
</tr>
<tr>
<td>Patient ML</td>
<td>585</td>
<td>556</td>
<td>662</td>
<td>106</td>
</tr>
</tbody>
</table>

Accuracy

<table>
<thead>
<tr>
<th>Participant</th>
<th>Congruent</th>
<th>Neutral</th>
<th>Incongruent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>99.7</td>
<td>99.8</td>
<td>99.8</td>
</tr>
<tr>
<td></td>
<td>(97.5-100)</td>
<td>(98.75-100)</td>
<td>(98.75-100)</td>
</tr>
<tr>
<td>Patient ML</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 5
Reaction Time (RT, in Milliseconds), Log Transformed RT Data and Accuracy (% Correct) for Recent Negatives Task - Controls and Patient ML

<table>
<thead>
<tr>
<th>Participant</th>
<th>Recent Negative</th>
<th>Nonrecent Negative</th>
<th>Recent Positive</th>
<th>Nonrecent Positive</th>
<th>Interference (Recent - Nonrecent Negatives)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>1006</td>
<td>915</td>
<td>873</td>
<td>872</td>
<td>91</td>
</tr>
<tr>
<td>ML</td>
<td>2905</td>
<td>2174</td>
<td>1474</td>
<td>1416</td>
<td>731</td>
</tr>
</tbody>
</table>

Log Transformed RT Data

<table>
<thead>
<tr>
<th>Participant</th>
<th>Controls</th>
<th>ML</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.990</td>
<td>3.445</td>
</tr>
<tr>
<td></td>
<td>2.953</td>
<td>3.300</td>
</tr>
<tr>
<td></td>
<td>2.933</td>
<td>3.223</td>
</tr>
<tr>
<td></td>
<td>2.932</td>
<td>3.136</td>
</tr>
<tr>
<td></td>
<td>.037</td>
<td>.145</td>
</tr>
</tbody>
</table>

Accuracy

<table>
<thead>
<tr>
<th>Participant</th>
<th>Controls</th>
<th>ML</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>94.7*</td>
<td>62.5*</td>
</tr>
<tr>
<td></td>
<td>98.9</td>
<td>87.5</td>
</tr>
<tr>
<td></td>
<td>98.9</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>98.9</td>
<td>96</td>
</tr>
</tbody>
</table>

*Recent Negative < Nonrecent Negatives, p<.05