

Characteristics of Syntactic Processing:
An Examination Utilizing Behavioral and fMRI Techniques

by

Evan Chen

Submitted to the Harvard-MIT Division of Health Sciences and Technology,
Speech and Hearing Biosciences and Technology Program,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy of Speech and Hearing Biosciences and Technology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 2004

[February 2004]

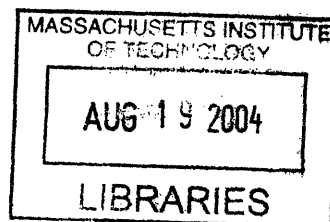
© 2003 Evan Chen. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly
paper and electronic copies of this thesis document in whole or in part.

Signature of Author: _____
Harvard-MIT Division of Health Sciences and Technology
Speech & Hearing Biosciences & Technology Program
9 January 2004

Certified by: _____
David Caplan, M.D., Ph.D.
Professor of Neurology
Harvard Medical School
Thesis Supervisor

Accepted by: _____
Martha L. Gray, Ph.D.
Edward Hood Taplin Professor of Medical and Electrical Engineering
Co-Director, Harvard-MIT Division of Health Sciences and Technology



ARCHIVES

Characteristics of Syntactic Processing: An Examination Utilizing Behavioral and fMRI Techniques

by

Evan Chen

Submitted to the Harvard-MIT Division of Health Sciences and Technology,
Speech and Hearing Biosciences and Technology Program,
on 9 January 2004,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy of Speech and Hearing Biosciences and Technology

Abstract

This thesis explores two important factors that constrain the syntactic parser of the sentence processing mechanism, syntactic storage costs and plausibility information. It uses behavioral methods to explore the characteristics of the two factors and neuroimaging to explore the underlying neurological substrates associated with these aspects of syntactic processing. Experiment 1 behaviorally demonstrated the presence of syntactic storage costs for predictions of verbs, filler-gaps, and subcategorized prepositional phrases. It is argued that the data support the Dependency Locality Theory (Gibson, 2000) supposition of stored predicted heads as well as a theory of syntax that includes empty categories. Experiment 2 demonstrated brain regions associated with storage and integration cost demands in the contrast of subject-object (SO) and object-subject (OS) sentence structures. The results indicate that the inferior parietal cortex is part of a larger network of cortex, including inferior frontal perisylvian areas, that is involved in the processing of SO vs. OS sentences. However, the involvement is not identical to that of the inferior frontal areas and has a distinct hemodynamic character. Experiment 3 explored regions of the brain involved in the resolution of the main verb / reduced relative (MV/RR) ambiguity. Activation was seen in portions of the angular gyrus and the middle temporal gyrus for a contrast in subject noun plausibility, but not structure ambiguity, indicating that the MV interpretation was still considered even in unambiguously relative clause sentence structures. The unexpected results could imply that syntax is not the only factor that determines θ -role assignment and ultimately provide evidence about the brain regions involved in the process of plausibility information resolution in sentence interpretation.

Thesis Supervisor: David Caplan

Title: Characteristics of Syntactic Processing: An Examination Utilizing Behavioral and fMRI Techniques

Acknowledgements

I would first like to thank the members of my committee – David Caplan, Edward Gibson, and Stefanie Shattuck-Hufnagel – for their insightful guidance and advice, without which this research would not be possible. I would especially like to thank my advisor, David, and committee chair, Ted, for their continued patience, support and encouragement.

Many others have contributed immensely to my growth as a researcher as well. In particular, I am grateful to Caroline West, Gina Kuperberg, and Jennifer Melcher for their indispensable help in my times of need. I am also indebted to the many incredibly talented individuals across all of Boston with whom I have had the pleasure to work with, including the wonderful people of TedLab at MIT, the Language Science Lab at Boston University, the A. Martinos Imaging Center at Charlestown, MA, the Division of Health Sciences and Technology, and the Speech and Hearing Biosciences and Technology Program.

Additional thanks go to the innumerable friends who have somehow managed to keep me relatively sane throughout the years, including, but not limited to, Michael Qin, Laura Diao, Ritwik Chatterjee, Eric Stuckey, Patrick Purdon, Kinu Masaki, Asa Masaki, Sarah Wang, and LeeAnn Henn. Special thanks go to my U of I college buddies Alex Wang, Peter Chou, David Eng, and Richard Chen.

Finally, and most importantly, I would like to express my profound gratitude to my family – my mom, Anita, my dad, Haydn, and my brother, Chris – to whom I dedicate this thesis, for their never-ending love and support.

Table of Contents

1	Introduction	9
1.1	How do We Process and Comprehend Language?	9
1.2	The Constraint-based Framework	13
1.3	Important Constraints	15
1.4	The Dependency Locality Theory (DLT)	19
2	Techniques	21
2.1	Behavioral	21
2.2	Functional Magnetic Resonance Imaging (fMRI)	23
3	Behavioral Studies on Syntactic Storage Costs	25
3.1	Introduction	25
3.2	Experiment 1a: Verb Predictions	32
3.2.1	Methods	34
3.2.2	Results	37
3.2.3	Discussion	44
3.3	Experiment 1b: Wh-trace Predictions	46
3.3.1	Methods	49
3.3.2	Results	51
3.3.3	Discussion	54
3.4	Experiment 1c: Verb Subcategorization Predictions	56
3.4.1	Methods	59
3.4.2	Norming Methods	60
3.4.3	Norming Results	62
3.4.4	Results	63
3.4.5	Discussion	69
3.5	General Discussion	71
3.6	Future Work: An Application of a Model of Storage Costs	75
3.6.1	Serial & Parallel Accounts of Ambiguity Resolution	76
3.6.1.1	Serial Models	76
3.6.1.2	Parallel Models	77
3.6.2	State of the Research	78
3.6.2.1	The Garden Path Model & Constraint-based Approaches	78
3.6.2.2	Modularity	79
3.6.2.3	A Recent Study	79
3.6.3	Alternative Constructions in the Ambiguity Resolution	82
4	Neuroimaging Studies on Syntactic Storage Costs	87
4.1	Introduction	87
4.2	Experiment 2: The SO/OS Contrast	95
4.2.1	Methods	96
4.2.2	Results	103
4.2.3	Discussion	117
4.3	Experiment 3: Wh-trace Predictions	122
4.3.1	Behavioral Methods	123

4.3.2	Behavioral Results.....	124
4.3.3	Discussion	126
5	Neuroimaging Experiment on Ambiguity Resolution.....	127
5.1	Introduction	127
5.2	Experiment 4: The MV/RR Ambiguity.....	135
5.2.1	Imaging Methods.....	135
5.2.2	Imaging Results.....	139
5.2.3	Behavioral Methods	153
5.2.4	Behavioral Results.....	153
5.3	Discussion	156
6	Conclusion.....	163
	References.....	167
	Appendices	177

Chapter 1

Introduction

1.1 How do We Process and Comprehend Language?

In our everyday experience, understanding a discourse we hear or read sometimes seems so easy and natural that, to our minds, it is often not the words that we remember, but the ideas and the messages behind them. Language is the powerful conduit through which we communicate our thoughts and feelings to each other and, it is a uniquely human skill. While many animals can communicate to each other, signaling danger or a source of food, none possess a fully developed language in quite the same way as humans. None have a language that can combine discrete elements, in an infinite manner, to form new and totally novel elements, nor have a language that can displace the communication to refer to events remote in time and space or even to refer to imaginary worlds. Yet, the apparent effortless with which we process language belies the true complexity of the system.

Our ability to understand even a single sentence is truly amazing when we consider the intricacies involved, from the rules of syntax governing the possible constructions to the elaborate syntactic dependencies and semantic relationships between individual words to the fact that processing all occurs within a short time frame, on an immediate, moment-by-moment basis. One of the key elements in sentence comprehension is the processing of syntactic structures. Syntactic

structures are mental representations that allow the sentence processor to link the meaning of the words together to reconstruct aspects of a proposition such as who did what to whom, what a pronoun refers to, etc. Thus, the syntactic processor, or parser, allows us to understand that in the English sentence, “The boy kissed the girl”, it is the boy doing the kissing and not the girl. Syntax also governs the relationships of reflexives such as “himself” and pronouns such as “him”, to the antecedents they refer to.

In sentence (1a) for instance, it is our intuition that “him” cannot refer to the possible antecedent John. Instead it must refer to some other male, not mentioned in the sentence, who got hurt. On the other hand, in sentence (1b), “himself” must refer to John and no one else.

(1)

- a. [John hurt him].
- b. [John hurt himself].

The candidate antecedents are not just the first word of a sentence though. Changing the sentences and thus the underlying syntactic structures means that in sentence (2a), “him” can now mean John but cannot refer to Bill. Meanwhile, in sentence (2b) “himself” must refer to Bill and cannot mean John. In the case of sentence (2c), “himself” must refer to the brother and cannot refer to John, unlike sentence (1b).

(2)

- a. John believes that [Bill hurt him].
- b. John believes that [Bill hurt himself].
- c. John’s [brother hurt himself].

Thus, it seems that the syntax of the sentence defines a relationship such that within a region (enclosed in square brackets), called the ‘governing category’, a pronoun like “him” cannot refer to an antecedent within the region while a reflexive like “himself” must refer to an antecedent within the region. Furthermore, because of examples like sentence (2c), it appears that the syntactic representations governing the relationships are complex and involve more than just a simple sentence or clause. It is the purpose of this thesis to explore some of the characteristics of the parser as it processes syntactic structures.

We begin by approaching the sentence processing mechanism within a constraint-based framework¹. The constraint-based framework is an interactive approach in which the syntactic parser, in the case of a temporarily syntactically ambiguous sentence for example, constructs the multiple possible syntactic structures of a sentence simultaneously, in parallel, and the correct one is chosen based on competition between the possible alternatives in order to best satisfy the relevant constraints described below. With competition, the parser encounters difficulty when there is either inconsistent information and the various alternatives are equally supported, or when new information is at odds with the previously biased and preferred alternative.

Researchers have discovered many factors that can constrain the parser when it processes both syntactically ambiguous sentences as we have just discussed, or when processing even syntactically unambiguous sentences². Some factors include the frequency of co-occurrence for lexical items³, the plausibility of the propositions⁴, the contextual situation in which the sentence is presented, the prosody, and the computational memory resources (for a brief review, see Gibson &

¹ See Section 1.2 for a brief description.

² See Section 1.3 for a further discussion of some relevant constraints.

³ Frequencies of co-occurrence are based on an examination of a corpus of written text and determining the number of occurrences of each lexical item with its counterpart. Greater frequency is assumed to imply familiarity and easier processing.

⁴ Plausibility is the possibility of the proposition happening in the real world and is somewhat related to how expected a word is. It can be quantified by participant ratings or such metrics as the Cloze frequency. Cloze frequencies are measured by presenting a fill-in-the-blank sentence and counting the number of times a given word is used to complete the sentence. In this case, greater Cloze frequency implies greater plausibility.

Pearlmutter, 1998). This thesis concentrates specifically on investigating the characteristics of computational memory resource use and the importance of plausibility in sentence processing. For our investigation of computational memory resource use, we focused primarily on the syntactic storage costs component of Gibson's (2000) Dependency Locality Theory (DLT) using both behavioral and neuroimaging techniques.

In Experiment 1, we used behavioral techniques to determine whether the storage of verb, filler-gap, or verb subcategorization predictions generates measurable increases in reading times. In Experiment 2, we examined the subject-object / object-subject (SO/OS) contrast⁵, relevant to the question of storage and integration costs and with an extensive presence in the neuroimaging literature, as an initial look into the issues of neuroimaging syntactic processes and to establish the anatomical regions that are relevant. Next, we planned to use neuroimaging techniques in Experiment 3 to locate the neural substrates of storage costs, in particular the storage of filler-gap predictions. Identifying the neurological basis of syntactic storage costs will be important to further our understanding of the underlying cognitive processes in the sentence comprehension mechanism. However, it was determined that the effects of our stimuli were not robust enough to warrant expending expensive neuroimaging resources. Experiment 4 was a step away from the study of syntactic storage costs and explored the neurological correlates involved in the resolution of the main verb / reduced relative (MV/RR) ambiguity⁶, an important fundamental finding and a robust effect in the behavioral literature. It begins an examination of the role of plausibility, another important constraint that can influence the parser. Finally, we will discuss a potential behavioral application of our understanding of syntactic storage costs and ambiguity that could add insight to the age-old question of the serial vs. parallel nature of sentence processing⁷.

⁵ To be discussed in Chapter 4.

⁶ To be discussed in Chapter 5. See also Section 1.3 for a brief description.

⁷ A serial sentence processing mechanism computes only one preferred interpretation of a sentence at a time while a parallel processing mechanism considers several possible interpretations simultaneously. Making sense of this issue could potentially have an impact on our understanding of cognitive processing in general.

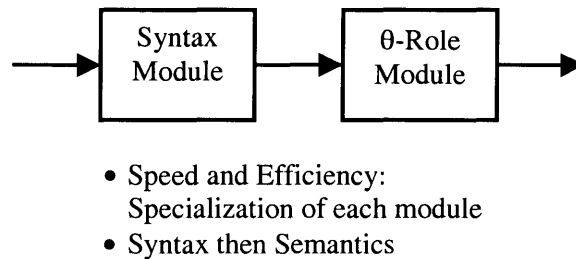
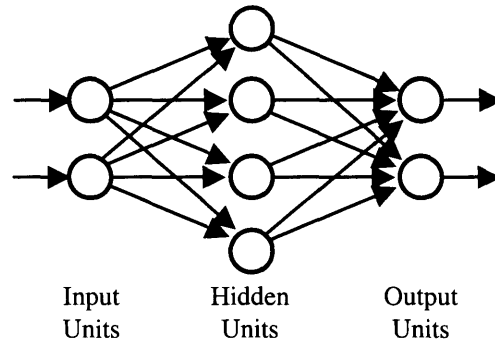


Figure 1-1. The Garden Path model is an example of a modular approach to describing the sentence comprehension mechanism. It contains two stages, one dedicated to syntactic processing and the second to semantic processing.

1.2 The Constraint-based Framework

What is the framework within which computational resource use and plausibility information operate? They are viewed by Gibson & Pearlmutter (1998) as components of a constraint-based approach to sentence comprehension. While considering the characteristics of the sentence processing mechanism, two broad approaches have been developed: an informationally encapsulated approach and a constraint-based approach (for a brief review, see Tanenhaus & Trueswell, 1995).

The informationally encapsulated approach typically refers specifically to the Garden Path model (Figure 1-1) in which the sentence processing mechanism is divided into two modules, a syntactic processor module and a thematic-role (θ -role) processor module (Ferreira & Clifton, 1986; Frazier & Fodor, 1978). The modules are organized one after the other and operate serially on a single sentence at a time, making repairs or re-parsing the entire sentence as errors are detected. The syntax module is first and its processing takes precedence and plays a primary role. The premise behind the approach is that the sentence comprehension mechanism achieves its speed and ease of processing from having each of its modules extremely specialized for its domain (Fodor, 1983).



- Speed and Efficiency:
Immediate access to all information
- Syntax and Semantics together

Figure 1-2. This is a simple example of a neural network, which is possibly a realistic neurological instantiation of the constraint-based approach to describing the sentence comprehension mechanism.

The syntactic module for instance, has access only to syntactic information, such as grammatical category (e.g. noun or verb), and can only use syntactically based rules and heuristics of processing, such as Minimal Attachment and Late Closure⁸ (Rayner et al., 1983; Frazier & Rayner, 1982; Frazier, 1978). The syntactic module has no access to such semantic information as the plausibility of the propositions or the corpus frequency of the words. The Garden Path model offers a unified account of many types of attachment ambiguities, but does not fully capture the apparent immediate influence of non-syntactic information on the parser (Altmann & Steedman, 1988; Spivey-Knowlton et al., 1993; MacDonald, 1994a; Trueswell et al., 1994; Garnsey et al., 1997).

The constraint-based approach (Figure 1-2) is a strongly interactive architecture with multiple constraints that must be satisfied and competition among incompatible alternative syntactic structures. One instantiation of the approach is lexically-mediated syntactic processing (MacDonald, 1994b; MacDonald et al., 1994). In MacDonald's model, each lexical entry representation contains

⁸ Minimal Attachment states that the parser must build the simplest structure consistent with the rules of grammar for the language and Late Closure states that the parser prefers to make attachments to the phrase currently being processed if the alternatives are equally simple.

the local syntactic structure of the word as well as semantic information such as the meaning, the corpus frequency, etc. Syntactic processing then is the combination of the local syntactic structures of each word, as it appears, to the existing sentence structure, and the parser is naturally influenced by the non-syntactic components of the lexical entry. Like the Garden Path model, this approach accounts for many of the attachment ambiguities in the literature. However, the constraint-based approach suffers from underspecification and is not a model per se. It merely states that all possible types of information are available immediately and can influence the parser. Despite its underspecification, it must be noted that the constraint-based approach is in line and compatible with neural network models of sentence comprehension, a possibly realizable neurological model of the brain.

1.3 Important Constraints

What types of constraints are involved in syntactic processing? In addition to storage and integration costs, the constraints that have been shown to influence the syntactic parser include: (1) lexical, such as grammatical category and corpus frequency of the word as well as verb subcategorization and frequency bias of the verb; (2) contextual, such as the local state of world knowledge and plausibility as well as referential context in the discourse, both current and prior; (3) phrase-level contingent frequency; and (4) prosody, such as the position and character of pauses, intonation, etc.

Constraints at the lexical level, such as the preferred subcategorization of a verb, can influence sentence-level syntactic processing. For instance, in example (3), the two possible syntactic structures for the ambiguous initial words “The defendant examined” are presented. After the verb “examined”, the MV structure of sentence (3a) is disambiguated by the direct object noun phrase (NP) “the documents” while the RR structure of sentence (3b) is disambiguated by the prepositional phrase (PP) “by the lawyer”. By disambiguated, we mean that the processing mechanism now has enough information to construct the proper syntactic structure.

(3)

a. Main Verb (MV) Structure

The defendant examined the documents.

b. Reduced Relative (RR) Structure

The defendant [examined by the lawyer] was unreliable.

Typically, the MV structure is empirically preferred, however it has been shown that the subcategorization preference of the verb “examined” is also important. In the MV structure, “examined” is in its simple past tense form while in the RR structure, “examined” is actually in its past participle form. It has been shown that the frequency bias of the verb (i.e. which verb form occurs more often in a corpus count or generation task) can influence the preferred interpretation of the ambiguity (Trueswell et al., 1994).

Another influence on syntactic processing is the context within which the sentence occurs. One ever-present context is the real world. For sentences to be sensible to participants, they must be plausible and satisfy the participant’s knowledge of the world. In the case of example (3) for instance, it is quite plausible and even typical for “defendants” to examine things. Thus, the MV structure is preferred, both because of the verb form bias of “examined” and because of the plausibility of “defendants” examining things. However, in the case of example (4), “evidence” cannot examine anything, but is most often examined by someone, making the RR structure more preferred.

(4)

a. MV Structure

*The evidence examined the documents.

b. RR Structure

The evidence examined by the lawyer was unreliable.

As Trueswell et al. (1994) showed, this kind of plausibility information does indeed influence the syntactic processor.

A third constraint that can influence the syntactic processor is phrase-level contingent frequency. Unlike the frequencies we've been discussing up till now, phrase-level contingent frequencies are not over individual lexical items, but rather over phrase combinations in certain syntactic contexts. In example (5), the word "that" is in the sentence initial position. For sentence (5a) it is used as an article, which is its preferred interpretation in this position, while in sentence (5b) it is used as a complementizer.

(5)

a. Sentence Initial, Article (preferred)

That cheap hotel was clean and comfortable to our surprise.

b. Sentence Initial, Complementizer

That cheap hotels were clean and comfortable surprised us.

On the other hand, in example (6), the word "that" is in a post-verbal position. For sentence (6a) it is used as an article, but this time the preferred interpretation is as a complementizer in sentence (6b).

(6)

a. Post-verbal, Article

The lawyer insisted that cheap hotel was clean and comfortable.

- b. Post-verbal, Complementizer (preferred)

The lawyer insisted that cheap hotels were clean and comfortable.

As a result, it is clear that the preference for the interpretation of “that” depends on its location in the overall sentence structure and is also distinct from its overall lexical frequency preference for the complementizer interpretation (Juliano & Tanenhaus, 1994).

Finally, prosody is an important constraint that can also influence syntactic processing. For instance, in sentence (7a) the NP “the door” is attached as the object of the verb “checks”, while in sentence (7b) “the door” is the subject of the verb phrase “is locked”.

(7)

- a. Object Clause, No Clause Boundary

Whenever the guard checks the door, it’s locked.

- b. Subject Clause, Major Clause Boundary

Whenever the guard checks, the door is locked.

Syntactically, the sentences are ambiguous until after “the door”, but with an auditory presentation, the speed of comprehension was fastest when the locations of the clause boundaries, in the form of pauses, intonation markers, etc., was consistent with the syntax (i.e. boundaries at the commas). Furthermore, there is no evidence of a preference for the syntactic rule of Late Closure when the prosody does not support it (Speer & Bernstein, 1992).

The constraints that are described above are examples of some of the factors in addition to computational resource use that can influence the syntactic parser. Below, we will briefly describe the specific model of computational resource use that this thesis explores.

1.4 The Dependency Locality Theory (DLT)

The DLT is a resource use model that describes the relationship between the sentence processing mechanism and available computational resources proposed by Gibson (2000). Evidence for the model was derived from studies of both ambiguous and unambiguous structures. In essence, the DLT claims that some processing differences associated with differences in syntactic relations are determined by the two resource-related factors of integration costs and storage costs.

Integration costs are a measure of the difficulty of integrating a new word to the appropriate head established previously in the sentence. The measure is distance- or locality-based such that the further away the input word is from its head, as measured in the number of new intervening referents, the greater the costs. Several experiments have investigated integration cost effects already and seem to show increased reading times at the predicted points at which new words are integrated to prior heads (see Gibson, 1991, 1998).

Storage costs, on the other hand, are associated with the number of predictions or dependencies the processing system needs to maintain in order to complete the current sentence fragment as a grammatical sentence. As a preliminary assumption, the DLT presupposes that only a subject noun and a verb are minimally required to create a simple grammatical sentence (e.g. “The man slept”). The less studied storage cost component is of particular concern for this research program, and many of the experiments described below are part of an effort to characterize the role of storage costs in sentence processing both in terms of behavioral reading times and neuroimaging localization.

Chapter 2

Techniques

To examine the characteristics of storage costs and the influence of plausibility information, we planned to utilize two techniques, self-paced moving windows for the behavioral reading times and functional Magnetic Resonance Imaging (fMRI) for the neuroimaging activation maps. As mentioned in the Introduction, Experiment 1 is the initial behavioral examination of storage costs while Experiment 2 is a neuroimaging experiment to establish the anatomical and theoretical foundations of storage and integration costs for Experiment 3. Unfortunately, a behavioral pretest of the stimulus items in Experiment 3 did not warrant a subsequent neuroimaging experiment. Finally, Experiment 4 is a behavioral and neuroimaging exploration into the influence of plausibility on syntactic parsing and ambiguity resolution.

2.1 Behavioral

The technique used most often because of its ease of implementation, low cost, and proven ability to obtain interpretable results is self-paced reading in a moving-window presentation (Figure 2-1). In this method, the entire sentence form is presented to the participant on the screen with dashes replacing all of the characters except for the spaces. The participant then presses a button and the letters of the first word appears, replacing the dashes. With the next press of a button, the first word

-----,	A -----,	- moving -----,
t=0	t=1	t=2
----- window -----,	----- display.	Comprehension question?
t=3	t=4	t=5

Figure 2-1. An example of a moving window display for the sentence “A moving window display”.

is replaced with dashes again and the next word is uncovered, and so on until the end of the sentence. The time it takes for the participant to press the button is measured as the reading time for each word and is thought to reflect online processing by the sentence processing mechanism. Longer reading times for words controlled for in length, frequency, and other such confounding factors are assumed to represent more effortful processing as a result of greater confusion or load experienced by the processor. After each sentence, a question testing the participants' comprehension of the material is presented, typically a yes or no question. For the experiments in this thesis, the self-paced moving window task was implemented with Linger 1.7, written by Doug Rhode (Boston, MA), a multi-platform program capable of running on Macintosh, Microsoft Windows, and Unix operating systems. The details of the items and methods are described in the appropriate sections of each behavioral experiment.

2.2 Functional Magnetic Resonance Imaging (fMRI)

Because of the lack of a comparable animal model, studies of the brain structures associated with sentence processing are necessarily constrained by the ethics of conducting human studies. Neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI) provide a powerful non-invasive means of examining neural activity in the brain. The first studies localizing the neurological substrates of sentence processing began with deficit-lesion examinations of patients. These studies are crucial for determining areas of brain that are necessary for processing sentences. However, as with all techniques there are limitations. First, brains with lesions are necessarily abnormal systems and the plasticity of the brain, its ability to recover lost functions, becomes relevant. Second, damage often includes both cortical and sub-cortical areas and is thus not well localized. Even within cortical areas, damage is often spread across a relatively wide extent and not limited to specific regions (on the order of cm) and spatial resolution is therefore relatively poor. Finally, in addition to the difficulty of determining the localization of the damage, it is also difficult to classify the aphasias and determine specific aspects of syntactic and sentence processing deficits. Despite these limitations however, researchers have certainly learned much about the basic functional organization of the brain in sentence processing, including the importance of such regions of the brain as Broca's and Wernicke's areas.

Neuroimaging techniques such as Event-Related Potentials (ERP), Positron Emission Tomography (PET), and fMRI can complement deficit-lesion studies by examining the normal processing of sentences and finding the all the areas that are relevant in the processes. For fMRI, temporal resolution is on the order of seconds, which while much finer than the temporal resolution of deficit-lesion studies or PET, is not yet on the hundreds-of-milliseconds level of measuring language processing that is possible with ERP. Spatial resolution for fMRI is in the millimeters, making very detailed analyses of the neurological organization of sentence processing a possibility. Furthermore, fMRI allows for event-related designs in which items can be presented in a random

order and later combined by condition for analysis (Dale & Buckner, 1997). Compared to block designs in which items of a given condition must be presented together in a block, event-related designs reduce the possibility of strategy effects by the participants and present a slightly more natural reading environment, though they are statistically less powerful than block designs.

What does fMRI measure? Functional MRI is typically a measure of the Blood Oxygenation Level Dependant (BOLD) contrast. With an increase in the mean electrical activity of a population of neurons, there is a local increase in metabolism. Subsequently, the body responds by increasing the regional cerebral blood flow (rCBF) in order to supply the energy and waste removal needs of the cells. The increase in the rCBF in turn lowers the local deoxygenated hemoglobin concentrations. Deoxygenated hemoglobin, because of the paramagnetic properties of the free electron in the protein's structure when there is no oxygen bound to it, effectively decreases the intensity of the local MR signal. Thus, with a lower concentration of local deoxygenated hemoglobin as a result of greater neural activity, the local MR signal intensity increases. In this way, the MR signal that we measure is several steps removed from the electrical neural activity we are most interested in, and the signal is dependent on both the temporal and spatial characteristics of the underlying vasculature. Details of the procedures used in the neuroimaging experiments are described in the Methods section of each neuroimaging experiment.

Chapter 3

Behavioral Studies on Syntactic Storage Costs

3.1 Introduction

Before we can begin any neuroimaging examination of syntactic storage costs, we will need to have a good psychological theory of the relevant behavioral phenomena. It is well known that nested (or center-embedded) syntactic structures are harder to understand than their right- or left-branching counterparts (Yngve, 1960; Chomsky & Miller, 1963). For example, the right-branching English structure in (8a) is easier to understand than the nested structure in (8b), in which the greater difficulty is indicated by the “#”, and the left-branching Japanese structure in (9a) is easier to understand than its nested version in (9b):

(8)

- a. Mary met the senator who attacked the reporter who ignored the president.
- b. # The reporter who the senator who Mary met attacked ignored the president.

(9)

a. [[imooto-ga naita to] bebiisitaa-ga itta to] obasan-wa omotteiru.

sister-nom cried that babysitter-nom said that aunt-top think

“My aunt thinks that the babysitter said that my sister cried”

b. # obasan-wa [bebiisitaa-ga [imooto-ga naita to] itta to] omotteiru.]

aunt-top babysitter-nom sister-nom cried that said that think

“My aunt thinks that the babysitter said that my sister cried”

The difficulty of understanding nested structures occurs despite the fact that each nested sentence has the same propositional content and lexical items as its right- or left-branching counterpart. Although the phenomenon is known, the details of the cause for the difficulty of processing nested structures are not yet known. A number of researchers have hypothesized that nested structures require more syntactic memory or storage space to process than left- or right-branching structures (Chomsky & Miller, 1963; Gibson, 1991, 1998; Abney & Johnson, 1991; Lewis, 1996; Stabler, 1994), but the accounts differ on which aspects of syntactic structure are associated with storage costs. Yngve (1960) and Chomsky & Miller (1963) proposed that syntactic storage at a particular parse state is quantified in terms of the number of partially processed phrase structure rules like $S \Rightarrow NP VP$. More generally, syntactic storage may also be required for all obligatorily required syntactic dependency positions, including incomplete dependencies that are not represented within a single phrase structure rule, such as filler-gap dependencies (Wanner & Maratsos, 1978; Gibson, 1998). Other proposals for units that require syntactic storage include incomplete clauses (Kimball, 1973), incomplete thematic role assignments (Hakuta, 1981; Gibson, 1991), incomplete case-assignments (Lewis, 1996; Stabler, 1994) and predicted syntactic heads such as verbs that are needed to complete sentences grammatically (Gibson, 1998).

Consider how the incomplete dependency account explains the contrast between sentences (8a) and (8b). The most complex state in processing the nested structure (8b) in terms of syntactic storage occurs at the point of processing the most embedded subject noun phrase (NP) “Mary”. There are five incomplete syntactic dependencies at this point: three subject-verb dependencies: (1) the NP “the reporter” requires its verb “ignored”; (2) the NP “the senator” requires its verb “attacked”; and (3) the NP “Mary” requires its verb “met”; and two filler-gap dependencies: (1) the first instance of the relative clause (RC) pronoun “who” which is eventually linked with the object position of the verb “attacked”; and (2) the second instance of the RC pronoun “who” which is eventually linked with the object position of the verb “met”. In contrast, the storage requirements for the right-branching structure in (8a) are at most two incomplete dependencies at the point of processing the determiner “the”, following “met”. At this point, the determiner requires a noun, and the verb “met” also requires a noun. In terms of storage costs, the right-branching structure is less complex than the nested version.

Note that a syntactic storage account that is quantified in terms of predicted syntactic heads makes the same predictions as the incomplete dependency account on these and many other syntactic structures of English. In particular, there are five predicted syntactic heads at the most complex state of processing sentence (8b) – three predicted verbs and two predicted empty categories associated with the wh-fillers. In contrast, there are at most two predicted syntactic heads in processing sentence (8a), and the contrast in processing difficulty follows similarly. The reason that both the incomplete dependency and the predicted syntactic head accounts of syntactic storage make identical predictions in (8) is that in head-initial languages such as English, dependencies between heads branch to the right such that each incomplete dependency necessarily has an upcoming predicted syntactic head. However, for head-final languages such as Japanese, the dependencies branch to the left and there can be situations in which several incomplete dependencies all predict the same syntactic head (see Section 3.5 for further discussion).

Although the intuitive contrast in difficulty between sentences like (8a) and (8b) has been verified using off-line measures such as intuitive acceptability (e.g. Miller & Isard, 1964; Stolz, 1967; Gibson, 1998; Gibson & Thomas, 1997), the presence of this contrast does not provide evidence for the existence of syntactic storage costs in on-line sentence processing. Furthermore, there are at least two other factors in addition to storage costs that are involved in the comparison between sentences (8a) and (8b). One such confounding factor is integration cost. Gibson (1998) has hypothesized that another reason that nested structures are more complex than non-nested structures is that nested structures always necessitate longer distance integrations between syntactic heads than non-nested structures. Integration distance also accounts for the observation that object-extracted RCs such as sentence (10a) are harder to process than subject-extracted RCs as in sentence (10b) (Ford, 1983; Holmes & O'Regan, 1981; King & Just, 1991; Gibson, 1998).

(10)

a. Object-Extracted RC:

The reporter_i [who_i the senator attacked _____i] ignored the president.

b. Subject-extracted RC:

The reporter_i [who_i _____i attacked the senator] ignored the president.

The integration difference between sentences (10a) and (10b) comes as a result of integrating the wh-pronoun with the gap associated with the embedded verb. In the case of the object-extracted RC in sentence (10a), this integration crosses the NP “the senator”, whereas there is no intervening material in sentence (10b). Similarly, integration distances are longer in the nested structure in sentence (8b) than in the right-branching structure in sentence (8a).

Another factor relevant to comparing sentence (8a) with (8b) involves potential processing differences between modifying the subject of a verb and modifying an object of a verb. Gibson et al.

(2002) observed that RCs modifying subjects are processed faster than RCs modifying objects (cf. Holmes, 1973, for related results). Gibson et al. attributed the observed reading time differences to differences in the information flow status of restrictive RCs (the type of RC in examples (8) and (10) above) together with the position of an NP in a sentence. Restrictive RCs are generally used to refer to objects that are background information, having already been mentioned or implied in the current discourse. In addition, background information is read faster when it modifies a subject rather than an object. Thus restrictive RCs are read faster in subject position, where they are compatible with being background information, than in object position, where they are less compatible with being background information. In order to avoid this issue when investigating on-line storage costs, the critical region clause to be compared should be in the same syntactic position.

Four recent English self-paced reading studies by Gibson and colleagues investigated the existence of on-line syntactic storage costs. In the first, Gibson et al. (2002) circumvented the potential confounds described above by comparing sentences like (11a) to the same sentences embedded within the sentential complement (SC) of a noun like “fact”, as in (11b):

(11)

- a. **The reporter who the senator attacked on Tuesday ignored the president.**
- b. The fact [that **the reporter who the senator attacked on Tuesday ignored the president**] bothered the editor.

The dependency between the subject NP “the fact” and its verb “bothered” is pending when people process the embedded sentence “the reporter who the senator attacked on Tuesday ignored the president” in (11b). Consistent with the hypothesis that maintaining this incomplete dependency consumes processing resources, Gibson et al. (2002) found slower RTs during the processing of the embedded sentence in (11b) as compared with the same regions in (11a). However, there is a further

complication with Gibson et al.'s design: greater syntactic storage is confounded with sentence position. That is, if people tend to read slower, later in sentences (perhaps only in self-paced reading), then sentence position could also account for the observed data pattern. Gibson et al. present correlational evidence against this hypothesis, but the design of their experiment makes the alternative theory difficult to rule out conclusively.

The three other experiments by Gibson and colleagues that are relevant to the existence of on-line storage costs in English involve comparisons between RC structures and verbal SC structures like those in example (12) (Grodner, Gibson & Tunstall, 2002; Gibson & Warren, 1999; Warren & Gibson, 2002):

(12)

a. RC:

The witness [who **the evidence [that was examined by the lawyer]** implicated ____] was lying.

b. Verbal SC:

The witness thought [that **the evidence [that was examined by the lawyer]** implicated his next door neighbor].

During the processing of the embedded subject NP region in bold, the structures, and hence the integrations, are the same. However, two extra incomplete dependencies need to be stored during the processing of the bold region in the RC structure (12a): (1) the subject-verb dependency associated with the NP “the witness”; and (2) the filler-gap dependency associated with the RC pronoun “who”. The additional storage should lead to longer reading times in sentences like (12a) when compared to (12b), and this is the pattern of data that was observed. Gibson & Warren (1999) and Warren & Gibson (2002) presented similar comparisons with similar results. The pattern of data

is consistent with a storage cost theory based on incomplete dependencies, incomplete phrase structure rules, or predicted syntactic heads. The only theory that does not predict this pattern of data is Kimball's (1973) principle of two sentences. Kimball's theory predicts a steep increase in complexity in regions where three verbs are expected, but makes no predictions with respect to the conditions that were compared here, because only at most two verbs are needed in either partial continuation.

The three parts of Experiment 1 test a number of predictions of storage cost theories using self-paced reading. First, the results that have been reported in the literature thus far show effects of a difference in a single incomplete subject-verb dependency: one vs. two, or two vs. three. A prediction of the theories is that there should be an incremental cost for one vs. two vs. three incomplete subject-verb dependencies, within the same set of items. Experiment 1a tests this prediction. Second, all of the experiments that show evidence of on-line syntactic storage cost compare conditions in which the high-storage conditions involve more incomplete subject-verb dependencies / predicted verbs than the low-storage conditions. It is therefore possible that storage cost is restricted to subject-verb dependencies (i.e. partially processed clauses). Experiments 1b and 1c test this prediction by examining two other kinds of incomplete dependencies: a relative clause pronoun dependency in Experiment 1b, and a PP argument of a verb in Experiment 1c.

3.2 Experiment 1a: Verb Predictions

Experiment 1a was designed to test whether predicted verbal heads are associated with incremental storage costs depending on the number of predicted verbs. To test for effects of storage costs on reading times, we used materials containing sentential complement (SC) - biased verbs (e.g. “realized” and “implied”) and their nominalizations (e.g. “realization” and “implication”) in order to manipulate the number of predicted verbs across a critical embedded clause. The materials were constructed in a 2x2 design, crossing the syntactic category of the first SC-taking word (Verb 1 or Noun 1) with the category of the second SC-taking word (Verb 2 or Noun 2). An example of the four conditions that were tested is given in (13).

The critical region in this design consists of the embedded clause “the company planned a layoff”, in bold. Because this clause has the same structure in all conditions, integration costs are identical across the four conditions. Furthermore, this clause is in a similar serial word order position in all conditions. Finally, this clause is in a similar syntactic position in each condition, the complement of a verb or noun. Thus, any observed differences in reading difficulty during this region cannot be attributed to integration cost differences, or word or syntactic position differences.

(13)

a. Zero predicted verbs (Verb 1 / Verb 2):

The employee realized that the boss implied that **the company planned a layoff** and so he sought alternative employment.

b. One late predicted verb (Verb 1 / Noun 2):

The employee realized that the implication that **the company planned a layoff** was not just a rumor.

c. One early predicted verb (Noun 1 / Verb 2):

The realization that the boss implied that **the company planned a layoff** caused a panic.

d. Two predicted verbs (Noun 1 / Noun 2):

The realization that the implication that **the company planned a layoff** was not just a rumor caused a panic.

For the zero predicted verbs condition in (13a), the critical material “the company planned a layoff” is embedded as the SC of the verb “implied” which is itself part of a clause embedded as the SC of the matrix verb “realized”. As a verb in the second clause, “implied” is labeled as Verb 2 while “realized”, as a verb in the first, matrix clause, is labeled as Verb 1. Because both verbs “implied” and “realized” are encountered immediately after their respective subject nouns, no additional verbs are predicted after the critical embedded clause. For the one late predicted verb condition in (13b), the verb “implied” is nominalized to “implication” (Noun 2) with the result that the critical clause is a SC of the noun “implication”. The change to the embedded subject NP “the implication” results in the prediction for an additional verb during the processing of the critical region. Similarly, for the one early predicted verb condition in (13c), the matrix verb “realized” is nominalized to “realization” (Noun 1) and the embedded SC “the boss implied that the company planned a layoff” is an argument of the matrix subject NP “the realization”. Once again, a verb is required after the critical region. Finally, for the two predicted verbs condition in (13d), both the verbs “realized” and “implied” are nominalized (Noun 1 and Noun 2, respectively) and two verbs are required following the critical region. Thus, if storage costs are proportional to the number of incomplete verb dependencies or predicted verbs held in memory, then the zero predicted verbs condition should be read fastest of the four conditions in the critical region. The one late and one early predicted verb conditions should be read more slowly, and the two predicted verbs condition should be read the slowest of the four.

On-line syntactic storage cost hypotheses also predict reading time differences in other regions of the target sentences. First, there is an extra predicted verb in the one early condition

relative to the zero predicted verbs condition during the second clause “the boss implied that”. Second, there is an extra predicted verb in the two predicted verbs condition relative to the one late condition during the region “the implication that”. Third, there is an extra predicted verb during the region “was not just a rumor” (following the critical region) for the two predicted verb condition relative to the one late condition. Reading times are predicted to be longer for each of the conditions with more predicted verbs.

Although both the one late and the one early conditions require the maintenance of one predicted verb over the critical region, the expectation is initiated at different points in the two sentences. If the amount of time that an incomplete dependency is held in memory increases the load, then reading times in the critical region of the one early condition should be longer than in the one late condition. Such an hypothesis was made in Gibson’s (1998) Syntactic Prediction Locality Theory (SPLT). In contrast to this theory, results reported in Gibson et al. (2002) failed to provide evidence for a syntactic storage cost metric that increases over distance (using a different manipulation). This evidence provided some of the motivation for Gibson’s (2000) more recent hypothesis, the Dependency Locality Theory (DLT). Because of the lack of distance-based storage effects in the earlier experiments, we do not anticipate such a difference in reading times between the one late and one early conditions for the critical region.

3.2.1 Methods

Participants. Forty-eight participants from the MIT and Boston University communities were paid for their involvement. All were native speakers of English and were naïve as to the purposes of the study.

Materials & Design. Forty sets of sentences were constructed, each with the four conditions as exemplified in (13). Each item in a set consisted of a matrix subject NP (e.g. “the employee”),

followed by a SC-biased verb (e.g. “realized”) or its nominalized form (e.g. “realization”). The next region consisted of the first SC, which began with the word “that” and was followed by an embedded subject NP (e.g. “the boss”). The following region consisted of an embedded SC-biased verb (e.g. “implied”) or its nominalization (e.g. “implication”), followed by the most embedded SC (e.g. “that the company planned a layoff”), which made up the critical material. Finally, various materials, which diverged for the four conditions but did not influence the theoretical predictions because they were post-critical, completed the sentences of each condition grammatically and plausibly.

The target sentences were split into four lists of 40 sentences each, balancing all factors in a Latin-Square design. Each list was combined with 60 fillers of various types. Appendix A provides a complete list of the stimuli. The stimuli were pseudo-randomized separately for each participant so that at least one filler item intervened between two targets.

Procedure. The task was self-paced word-by-word reading, using a moving window display (Just, Carpenter & Woolley, 1982). The software used to run the experiment was Linger 1.7 by Doug Rohde (Boston, MA). Each trial began with a visual display of a series of dashes marking the length and position of the words in the sentences, printed approximately one third of the way down the screen. A single line displayed up to 100 characters. Participants pressed the spacebar to reveal each successive word of the sentence. As each new word appeared, the preceding word returned to dashes. The amount of reading / processing time the participant spent on each word was recorded as the time between key-presses. After the final word of each item was a comprehension task in which the participant was presented with a fill-in-the-blank statement, which asked about information contained in the preceding sentence. This question was presented on the screen as a whole, along with a “?”. There were four choices provided as possible answers. Example comprehension questions for the items in example (13) are provided in (14):

(14)

- a. The employee realized that the boss implied that the company planned a layoff and so he sought alternative employment.

Question: The _____ planned a layoff.

Choices: company, boss, secretary, manager

Correct Answer: company

- b. The employee realized that the implication that the company planned a layoff was not just a rumor.

Question: The employee thought that the _____ planned a layoff.

Choices: manager, boss, company, executive

Correct Answer: company

- c. The realization that the boss implied that the company planned a layoff caused a panic.

Question: The _____ caused a panic.

Choices: realization, manager, company, employee

Correct Answer: realization

- d. The realization that the implication that the company planned a layoff was not just a rumor caused a panic.

Question: The implication had not been _____.

Choices: a rumor, unintentional, a threat, on purpose

Correct Answer: a rumor

Participants pressed one of four keys to respond. After an incorrect answer, the word "INCORRECT" flashed briefly on the screen. No feedback was given for correct responses. Correct answers were balanced across the four possible response positions. Participants were asked to read sentences at a natural rate and to be sure that they understood what they read. They were told to

answer the questions as quickly and accurately as they could and to take wrong answers as an indication to read more carefully.

Before the main experiment, a short list of practice items and comprehension tasks was presented in order to familiarize the participant with the task. The experiment averaged 30 minutes for each participant. For most participants, this experiment was followed by an unrelated experiment using the same self-paced reading procedure. Participants were given a short break between the two experiments.

3.2.2 Results

Comprehension Task Performance. Each participant had comprehension task performances greater than 67% accuracy overall. On average, the comprehension tasks for the experimental items in Experiment 1a were answered correctly in 85.2% of the trials. The percentages of correct answers for each condition are presented in Table 3-1.

An omnibus 1-way ANOVA (0, 1-Late, 1-Early, and 2 predicted verbs) showed a significant effect of comprehension task accuracy both by participants and by items ($F(3, 141) = 49.63$, $MS_{\text{within}} = 0.0096$, $p < 0.0001$; $F(3, 93) = 14.16$, $MS_{\text{within}} = 0.0281$, $p < 0.0001$). A 2x2 ANOVA (Noun or Verb in the first clause / NP region (N1/V1) x Noun or Verb in the second clause / NP region (N2/V2)) showed a significant main effect of the syntactic category of the first clause both by participants and by items ($F(1,47) = 45.39$, $MS_{\text{within}} = 0.0070$, $p < 0.0001$; $F(1,39) = 7.827$, $MS_{\text{within}} = 0.0337$, $p < 0.01$), a significant main effect of the syntactic category of the second clause both by participants and by items ($F(1,47) = 39.70$, $MS_{\text{within}} = 0.0116$, $p < 0.0001$; $F(1,39) = 11.50$, $MS_{\text{within}} = 0.0333$, $p < 0.002$), and a significant interaction both by participants and by items ($F(1,47) = 63.75$, $MS_{\text{within}} = 0.0102$, $p < 0.0001$; $F(1,39) = 31.82$, $MS_{\text{within}} = 0.0171$, $p < 0.0001$). These effects were carried by the fact that response accuracies to the two predicted verbs condition were worse than to any of the other conditions ($F_s > 16$; $p_s < 0.0004$).

Table 3-1

Number of Predicted Verbs			
0	1-Late	1-Early	2
88.33 (1.34)	90.21 (1.47)	91.87 (1.14)	70.42 (2.06)

Table 3-1. Mean (standard error) comprehension task accuracy for Experiment 1a in percent correct as a function of the number and type of predicted verbs, by participants.

Reading Times in the First Two Regions. We analyzed the reading time (RT) data in the same way across all three experiments presented in this chapter. To adjust for differences in word length as well as overall differences in participants' reading rates, a regression equation predicting RTs from word length was derived for each participant, using all filler and target items (Ferreira & Clifton, 1986; see Trueswell, Tanenhaus & Garnsey, 1994, for discussion). At each word position, the RT predicted by the participant's regression equation was subtracted from the actual measured RT to obtain a residual RT. Thus, a negative value for a residual RT denotes that the measured RT was faster than predicted by the regression. Furthermore, a less negative residual RT indicates a slower RT than a more negative residual RT. The statistical analyses gave the same numerical patterns for analyses of raw RTs as residual RTs and tables present both residual and raw RTs for each experiment. In order to save space, we restrict the presentation of statistical tests to the residual RTs.

Only items with correctly answered comprehension questions were analyzed. Furthermore, reading time data points that were greater than 3 standard deviations from the mean for a word position within a condition were also excluded from the analysis, affecting less than 1.3% of the data for Experiment 1a. For the purpose of analysis, we divided the materials into five regions, as shown below in (15):

(15)

The employee realized that / the boss implied that / the company planned a layoff / and so / (REST)

The realization that / the implication that / / was not /

/ / / caused a /

Clause 1 | NP1 / Clause 2 | NP2 / Critical clause / two words /

The first (1) region consisted of the material up to the first SC-taking word followed by the complementizer “that”. The second (2) region consisted of the material up to the second SC-taking word followed by the complementizer “that”. The third (3) region was the critical region. The fourth (4) region consisted of the first two words following the critical region. We analyzed only the first two words in this region because some items ended on the third word for some conditions and not others. Including the sentence-final word would have resulted in end-of-sentence effects confounding any possible effects of storage costs. The fifth (5) region (which was not analyzed because of large item differences between conditions in this region) consisted of the remainder of material in each item. Figure 3-1 presents the mean residual RTs per word (msec / word) for these regions across the four conditions in this experiment and Table 3-2 presents the mean RTs per word in both residual and raw forms.

In the First Clause / NP region, a 2x2 ANOVA (N1/V1 x N2/V2) revealed a main effect of the syntactic category of the first SC-taking word (N1/V1; i.e. “realization” / “realized”) both by participants and by items ($F(1,47) = 9.224$, $MS_{\text{within}} = 2747$, $p < 0.005$; $F(1,39) = 4.641$, $MS_{\text{within}} = 5619$, $p < 0.05$), in which clauses with the nominalized form of the verb had faster RTs than clauses with the verb itself, an unexpected result for models of storage costs. One possible explanation for this effect is that more interpretation can take place at a verb than at a subject NP, such as thematic role assignment (e.g. “the employee realized” vs. “the realization”). But this explanation is difficult to evaluate because the materials are necessarily different in this comparison as we are comparing

Figure 3-1

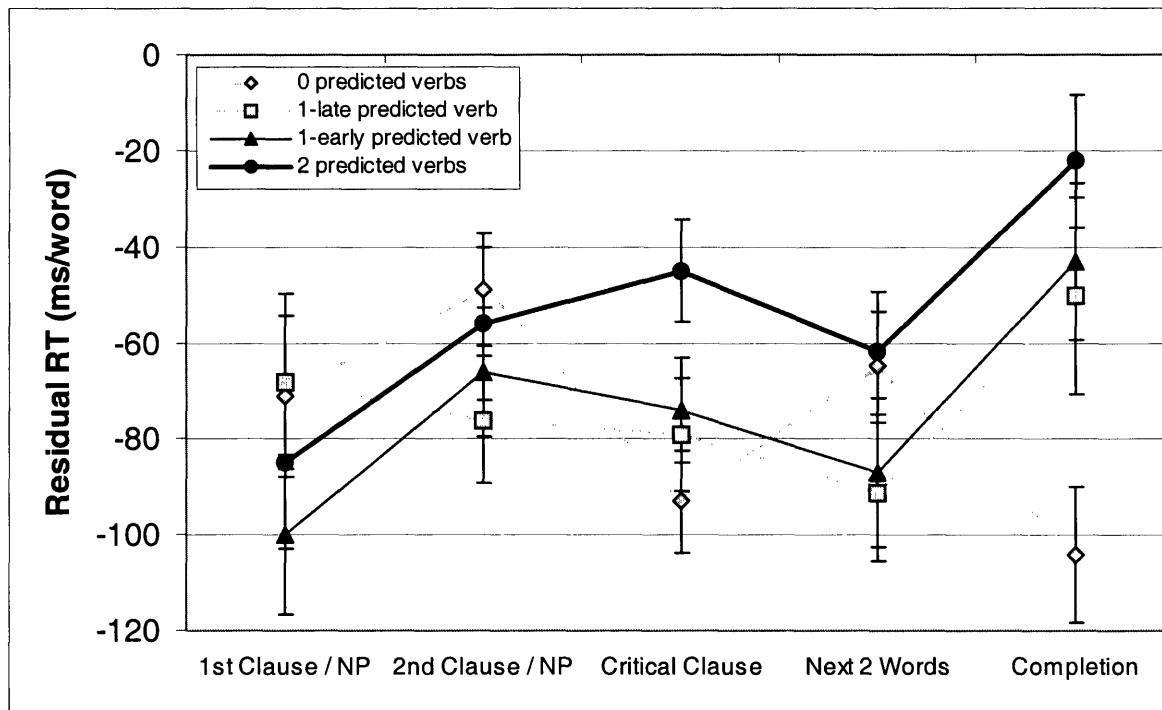


Figure 3-1. Plot of mean (standard error) residual RTs per word for each region of Experiment 1a as a function of the number and type of predicted verbs, by participants.

nouns with verbs. As expected, there was no effect of the category of the second SC-taking word (N2/V2; i.e. “implication” / “implied”) in the first region ($F_s < 1$), nor was there any interaction between the factors ($F_s < 1.1$). The lack of N2/V2 category effects makes sense because the materials do not yet differ along this dimension in the first region.

In the Second Clause / NP region, a second 2x2 ANOVA (N1/V1 x N2/V2) revealed no significant main effects of either the category of N1/V1 or of the category of N2/V2 ($F_s < 1.2$). There was, however, a significant interaction between the factors by participants and marginally so by items ($F_1(1,47) = 5.449$, $MS_{within} = 2958$, $p < 0.05$; $F_2(1,39) = 3.421$, $MS_{within} = 4122$, $p = 0.072$). One possible interpretation of this interaction is as an interference effect (cf. Gordon, Hendrick & Johnson, 2001). The two conditions that are slow in this region are the Verb 1 / Verb 2 (zero

predicted verbs) condition and the Noun 1 / Noun 2 (two predicted verbs) condition, as compared to the Verb 1 / Noun 2 (one late) and the Noun 1 / Verb 2 (one early) conditions. While this effect is not expected in models of storage costs, it could be that having similar types of NPs or VPs following one another leads to some slowdown at the second SC-taking word. In particular, the Noun 1 / Noun 2 condition consists of two SC-taking nouns in a row, whereas the Verb 1 / Verb 2 condition consists of two SC-taking verbs in a row. This similarity could have led to processing difficulty at the second instance of the category.

Recall that the storage cost theories predict that the second region should be read slower when the first region was nominalized. This prediction was not confirmed, but the presence of the interference effect in the second region may have masked any effects of structural storage cost here. In particular, the one early condition was predicted to be slower than the zero predicted verbs condition in the Second Clause / NP region, but the reverse numerical pattern appeared in the residual RTs ($F(1,47) = 3.821$, $MS_{\text{within}} = 1739$, $p = 0.057$; $F_2 < 1$), possibly because of the interference of the two verbs in the zero predicted verbs (Verb 1 / Verb 2) condition. The two predicted verbs condition was slower than the one late condition, marginally significant both by participants and by items ($F(1,47) = 2.910$, $MS_{\text{within}} = 3301$, $p = 0.095$; $F_2(1,39) = 3.489$, $MS_{\text{within}} = 2908$, $p = 0.069$), as predicted by the storage cost theories, but this effect may have been due to increased interference between the two nouns in the two predicted verbs (Noun 1 / Noun 2) condition. In this way, interference effects makes the storage cost hypotheses difficult to assess in the Second Clause / NP region.

Reading Times in the Critical Region. A 2x2 ANOVA (N1/V1 x N2/V2) revealed a significant main effect of N1/V1 ($F(1,47) = 16.14$, $MS_{\text{within}} = 2073$, $p < 0.0003$; $F_2(1,39) = 23.85$, $MS_{\text{within}} = 1981$, $p < 0.0001$) and a significant main effect of N2/V2 ($F(1,47) = 8.850$, $MS_{\text{within}} = 2369$, $p < 0.005$; $F_2(1,39) = 10.25$, $MS_{\text{within}} = 1903$, $p < 0.003$), but no significant interaction ($F_s < 1.3$). These

Table 3-2

	Number of Predicted Verbs			
	0	1-Late	1-Early	2
First Clause / NP	-71 (478)	-68 (484)	-100 (434)	-85 (446)
Second Clause / NP	-49 (491)	-76 (451)	-66 (474)	-56 (471)
Critical Clause	-93 (422)	-79 (441)	-74 (445)	-45 (471)
Next two words	-65 (391)	-91 (432)	-87 (448)	-62 (458)
Completion	-104 (418)	-50 (527)	-43 (514)	-22 (522)

Table 3-2. Mean residual RTs (msec / word) for each region of Experiment 1a as a function of the number and type of predicted verbs, by participants (raw RTs in parentheses).

main effects are the targeted syntactic storage effects: the noun conditions are slower than the verb conditions during the critical clause. Note that the interference effect that was present in the Second Clause / NP region is no longer present in the critical region. In particular, having two similar SC-taking verbs in the first two regions as in the zero predicted verbs condition did not cause processing difficulty in the critical region. Instead, the zero predicted verbs condition was the fastest condition to process. The lack of interference effects in the critical region is probably because the verb in the critical region is quite different from those in the first two clauses, which were always verbs that take SCs. Thus, the residual RT effects of storage costs in the critical region are not likely attributable to interference effects.

In the planned pair-wise comparisons of residual RTs in the critical region, the zero predicted verbs condition was significantly faster than the one early condition ($F(1,47) = 5.618$, $MS_{within} = 1500$, $p < 0.05$; $F(1,39) = 8.471$, $MS_{within} = 1624$, $p < 0.006$) and the two predicted verbs condition ($F(1,47) = 18.49$, $MS_{within} = 2905$, $p < 0.0002$; $F(1,39) = 28.86$, $MS_{within} = 2208$, $p < 0.0001$). Furthermore, the two predicted verbs condition was significantly slower than the one late condition ($F(1,47) = 9.888$, $MS_{within} = 2819$, $p < 0.005$; $F(1,39) = 14.04$, $MS_{within} = 2573$, $p < 0.0007$).

and the one early condition ($F(1,47) = 10.02$, $MS_{\text{within}} = 1954$, $p < 0.003$; $F(1,39) = 6.826$, $MS_{\text{within}} = 2677$, $p < 0.02$). The only predicted difference that did not reach statistical significance was the contrast between the zero predicted verbs condition and the one late condition. In this comparison, the zero predicted verbs condition was numerically but not reliably faster than the one late condition ($F(1,47) = 1.579$, $MS_{\text{within}} = 2660$, $p = 0.22$; $F(1,39) = 2.697$, $MS_{\text{within}} = 1442$, $p = 0.11$). This was a comparison that fared better in the analyses of raw RTs, in which the participants effect was marginally significant ($F(1,47) = 3.771$, $MS_{\text{within}} = 2442$, $p = 0.06$; $F(1,39) = 2.049$, $MS_{\text{within}} = 2182$, $p = 0.16$). Also, the comparison of residual RTs between the zero predicted verbs condition and one predicted verb condition (grouping together both the one late and one early conditions) was marginally significant by participants and significant by items ($F(1,47) = 3.635$, $MS_{\text{within}} = 1707$, $p = 0.063$; $F(1,39) = 6.129$, $MS_{\text{within}} = 1050$, $p < 0.02$). The same comparison in raw RTs was significant both by participants and items ($F(1,47) = 7.164$, $MS_{\text{within}} = 1574$, $p < 0.02$; $F(1,39) = 4.723$, $MS_{\text{within}} = 1231$, $p < 0.04$). Finally, the one late and one early conditions were not significantly different from each other ($F_s < 1.8$).

In summary, the pair-wise comparisons show us a 3-tiered pattern in which the condition with zero predicted verbs across the critical region displays the fastest reading times, the conditions with one predicted verb display slower reading times no matter when they are initiated, and the condition with two predicted verbs gives rise to the slowest reading times. These results are compatible with any of the syntactic storage accounts described previously.

Reading Times in the Post-critical Region. In the analysis of the post-critical region, it was discovered that the second word was actually the sentence-final word in 5 items of the one late condition. Excluding these items from the analysis produced a residual (raw) RT, by participants, of -103 (411) msec/word for the one late condition and -73 (444) msec/word for the two predicted verbs condition. The one late condition had significantly faster residual RTs than the two predicted

verbs condition in the post-critical region, both by participants and by items ($F(1,47) = 4.464$, $MS_{\text{within}} = 4750$, $p < 0.05$; $F(1,34) = 6.460$, $MS_{\text{within}} = 3053$, $p < 0.02$). This effect was as predicted by the storage cost theories. No other comparisons were made in this or later regions because of substantial item differences.

3.2.3 Discussion

As predicted by an on-line storage cost theory, RTs in the critical region increased proportionally to the number of predicted verbs that the sentence processor needed to maintain: zero predicted verbs fastest, followed by one predicted verb, with two predicted verbs slowest. Furthermore, the amount of time that the dependencies were stored did not affect the measured reading times: there was no significant difference between the one late and one early conditions, corroborating similar evidence from Gibson et al. (2002). The results in the post-critical region also ratified the prediction of the storage cost theory. However, one prediction of the storage cost theory was not ratified: that RTs should be slower for the one early condition as compared with the zero predicted verbs condition in the region preceding the critical region. Instead, RTs for the zero and two predicted verbs conditions were both slower than the one early and one late conditions, respectively. The observed effect could be accounted for by verbal interference in this region: in the zero predicted verbs condition there are two similar SC-taking verbs back-to-back, which may have resulted in slower processing. A similar interference confound between two SC-taking nouns makes the comparison between the two predicted verbs condition and the one late condition also difficult to assess in this region. In contrast, the effects in the critical region cannot be accounted for by interference because the verb there does not take a SC and is therefore considerably different from the verbs in the previous two regions. Thus, the results from Experiment 1a establish the measurable presence of storage costs across the critical region from storing verb predictions. However, among the theories presented in the introduction, the results do not differentiate which aspects of syntax are being stored, from predicted

syntactic heads to incomplete phrase structure rules. Experiment 1b begins to tease apart the theories by examining the filler-gap dependency, another kind of predicted head / incomplete dependency for which some of the theories make different predictions.

3.3 Experiment 1b: Wh-trace Predictions

In Experiment 1a, we presented evidence demonstrating a measurable increase in reading times for incomplete syntactic dependencies or storing predictions of expected verbs. These results are consistent with a storage cost theory that is sensitive to all kinds of incomplete dependencies, or with a storage theory that is sensitive only to incomplete clausal dependencies, such as Kimball's (1973) principle of two sentences. Experiment 1b was designed to examine whether incomplete dependencies other than predicted verbs also sustain storage costs. In particular, we focus on the filler-gap dependency between a wh-pronoun and its expected wh-trace. To test the possible processing costs of storing incomplete wh-dependencies / predicting wh-traces, participants were presented with ambiguous and unambiguous forms of noun-modifying sentential complement (SC) and relative clause (RC) structures. An example of the four conditions is given in (16).

(16)

a. SC, Ambiguous:

The claim [that **the cop who the mobster attacked** ignored the informant] might have affected the jury.

b. SC, Unambiguous:

The claim alleging [that **the cop who the mobster attacked** ignored the informant] might have affected the jury.

c. RC, Ambiguous:

The claim [that **the cop who the mobster attacked** ignored ____] might have affected the jury.

d. RC, Unambiguous:

The claim [which **the cop who the mobster attacked** ignored ____] might have affected the jury.

In sentences (16a) and (16b), a SC “that the cop who the mobster attacked ignored the informant” is attached to the matrix subject NP “the claim”. Sentences (16c) and (16d) have a RC “that / who the cop who the mobster attacked ignored” attached to the matrix subject NP instead. Critically, an RC includes a wh-dependency between its specifier (“who” or “that”) and a gap position inside the RC, in this case, the object of the verb “ignored”. An SC includes no such dependency. Thus, if storing the incomplete wh-dependency of the RC has an associated cost, reading times for the embedded subject noun phrase “the cop who the mobster attacked” – the critical region of interest for this experiment – are expected to be slower in the unambiguous RC condition in (16d) than in the unambiguous SC condition in (16b). Note that the critical region has the same structure and is in the same position in all four conditions, so that integration costs and other factors are controlled.

One confounding factor that is potentially different between the unambiguous RC condition and the other three conditions is that the unambiguous RC condition may be interpreted non-restrictively. The wh-pronoun “which” often initiates a non-restrictive RC, although usually with a comma, which is not present in these items, as well. In contrast, the other three conditions will very likely be interpreted restrictively because they are initiated by “that” or a verb SC sequence such as “alleging that”. Although the restrictiveness is a possible difference between the conditions, it is not problematic because such a difference only works against the predictions of the storage cost theories. In particular, non-restrictive RCs are processed more quickly in a null context than restrictive RCs (e.g. Grodner, Gibson & Watson, 2001), possibly because they implicate smaller discourse structures than restrictive RCs in a null context (Crain & Steedman, 1985; Altmann & Steedman, 1988). Thus, if we find evidence in support of the storage cost hypothesis, it will be in spite of differences in restrictiveness across the conditions.

Note that the predicted-head storage cost hypothesis makes different predictions depending on whether or not there are empty categories (wh-traces) mediating wh-filler-gap dependencies (cf.

Pickering & Barry, 1991; Gibson & Hickok, 1993). In particular, if there are such empty categories, then a predicted-head storage cost theory predicts increased complexity for the RC structure compared to the SC control structure in the critical region because of the additional predicted empty category which will eventually be placed in the object position of the embedded verb “ignored”. But if there are no such empty categories, then the predicted-head storage theory predicts no difference between the two structures in the critical region. In contrast to the predicted-head storage cost theory, the incomplete-dependency storage theory predicts increased complexity over the critical region of the RC structure whether or not there are empty categories mediating wh-dependencies because there is always an extra dependency associated with the wh-filler. Although the theories make potentially different predictions, we will simplify the discussion of the predictions of storage cost theories by restricting our attention initially to a storage cost theory based on predicted heads. We will therefore initially assume that there are empty categories associated with wh-fillers. We return to potential ways of distinguishing these kinds of theories in the general discussion.

The temporarily ambiguous forms of the SC and RC constructions like sentences (16a) and (16c) were presented as controls for the unambiguous RC structure. They are useful controls because the unambiguous SC structure includes an extra verb between the head noun and the SC (e.g. “alleging”). The temporarily ambiguous structures serve as controls that are more similar to the unambiguous RC version because these do not include the additional verb. In sentences with “that” clauses attached directly to matrix subject nouns (like those in (16a) and (16c)), there is an empirical and theoretical bias for the SC construction (Pearlmutter & Mendelsohn, 1999). These investigators have hypothesized that both the SC and RC interpretations are pursued in parallel for the ambiguous constructions, but such parallel parsing may be restricted to simple versions of our materials (e.g. without the additional RC embedding “who the mobster attacked”). In doubly-nested versions like the material in our Experiment 1b, Gibson & Thomas (1997) have shown that there is a strong advantage for the SC interpretation, suggesting that the RC reading is not followed when the

construction is quite difficult. Thus, we believe that in the critical region of the ambiguous constructions, people will be following the SC interpretation only.

In summary, if predicting an empty category gap following a wh-filler incurs a storage cost, then in the critical region “the cop who the mobster attacked”, unambiguous RC structures like (16d) should have longer reading times than unambiguous SC structures like (16b). The unambiguous RC structure is also predicted to have greater reading times in the critical region than either ambiguous structures (16a) or (16c), because we expect the processor to maintain only the less resource intensive SC completion in the ambiguous constructions. One consequence for the processor of maintaining only the SC completion is higher integration costs on the disambiguating material immediately following the embedded verb “ignored” for the ambiguous RC condition (i.e. the matrix verb phrase “might have affected the jury”) compared to the ambiguous SC condition (i.e. the embedded direct object “the informant”). However, because the material in these two regions is necessarily different, no direct comparisons on this prediction can be made for these items.

3.3.1 Methods

Participants. Forty-eight participants from the MIT community who did not take part in Experiment 1a were paid for their involvement. All were native speakers of English and were naïve as to the purposes of the study.

Materials & Design. Twenty sets of sentences were constructed with 4 conditions each (structure type (SC / RC) crossed with ambiguity (ambiguous / unambiguous)) following the form of (16). These items were constructed based on items from Pearlmuter & Mendelsohn (1999), with an additional RC modifying the embedded subject NP in order to make the critical region longer and more difficult. Each item in a set began with the same noun phrase, which consisted of the

determiner “the” followed by a noun (e.g. “claim”), which could take an SC. The word “that” followed the noun phrase for the ambiguous forms of the SC and RC structures, “implying that” or some equivalent for the unambiguous SC, and “which” for the unambiguous RC. The embedded clause followed, consisting of an embedded subject (e.g. “the cop”) and an RC modifying this NP (e.g. “who the mobster attacked”). The SC versions continued with a transitive verb and its object (e.g. “ignored the informant”), whereas the RC versions continued with the transitive verb alone completing the top level RC (e.g. “ignored”). Finally, a matrix verb phrase (e.g. “might have affected the jury”) completed the sentences.

The selected ambiguous SC / RC - taking nouns were somewhat biased on average towards the SC in both sentence completion and corpus frequency norms conducted by Pearlmuter & Mendelsohn (1999), resulting in the likelihood that participants would follow the SC interpretation in the temporarily ambiguous conditions. Appendix B provides a complete list of the stimuli along with the individual SC-completion percentage for each noun. Although there was an overall bias for the SC interpretation over all the items (62.6% SC completions on average), there was a wide range of biases in the individual items, between 44% and 93%.

The target sentences were split into four lists of 20 sentences each, balancing all factors in a Latin-Square design. Each list was combined with 112 fillers of various types, including 32 sentences from an experiment which was an early version of Experiment 1a. The stimuli were pseudo-randomized independently for each participant such that at least one filler item separated any two targets.

Procedure. The procedure was the same as in Experiment 1a, except that the comprehension task that followed each item was a simple yes / no question about the contents of the preceding sentence.

Table 3-3

	SC	RC
Ambiguous	76.34 (3.09)	77.35 (2.78)
Unambiguous	77.28 (2.88)	78.68 (2.47)

Table 3-3. Mean (standard error) comprehension question performance for Experiment 1b in percent correct as a function of clause type and ambiguity, by participants.

3.3.2 Results

Four participants' data were omitted from all analyses because of poor comprehension question performance (< 67% accuracy overall).

Comprehension Question Performance. Overall, the comprehension questions for the experimental items in Experiment 1b were answered correctly in 77.4% of the trials. The percentages of correct answers for each condition are presented in Table 3-3. An omnibus 2x2 ANOVA (SC/RC clause type x ambiguous/unambiguous) revealed no main effects and no interactions (all $F_s < 1$).

Reading Times in the Critical Region. Only items with correctly answered comprehension questions were analyzed. Reading time (RT) data points that were greater than 3 standard deviations from the mean were excluded from the analysis, affecting less than 1.8% of the data points in Experiment 1b overall. Figure 3-2 presents the mean residual RTs (msec / word) across the four conditions in this experiment. The critical region in the storage cost analysis consisted of the embedded subject noun phrase "the cop who the mobster attacked". Table 3-4 presents the mean RTs per word for all regions, in both residual and raw forms.

Figure 3-2

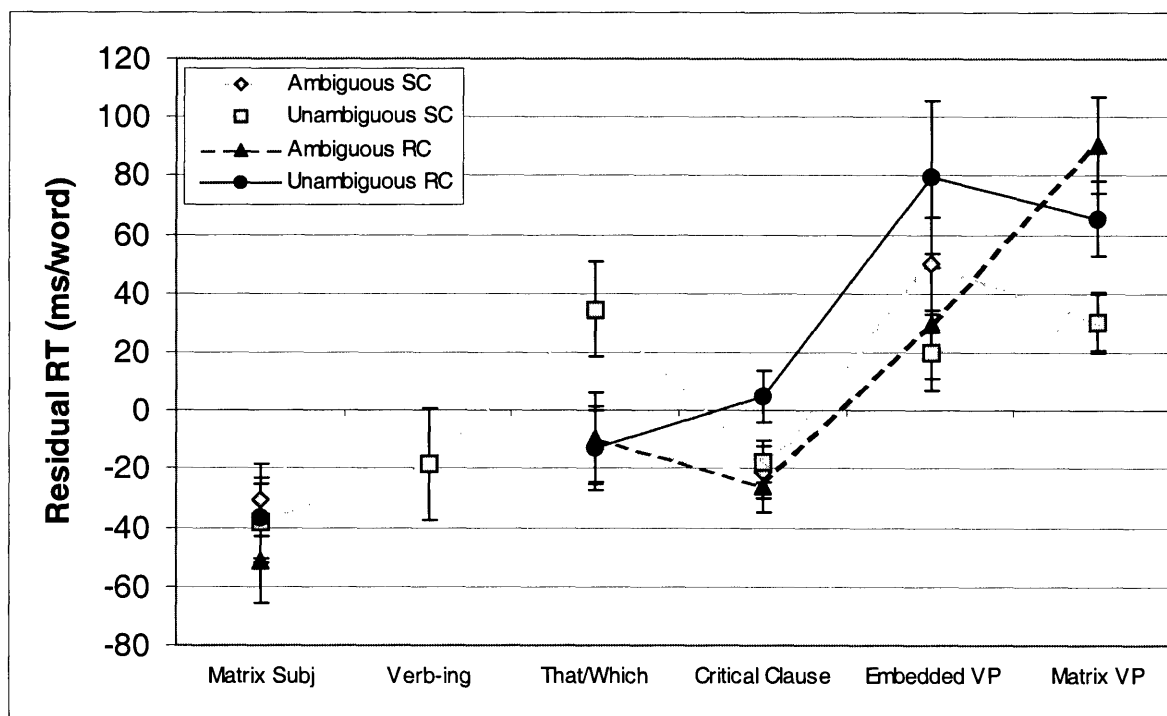


Figure 3-2. Plot of mean (standard error) residual RTs per word for each region of Experiment 1b as a function of clause type and ambiguity, by participants.

Across the critical material “the cop who the mobster attacked”, residual RTs were slower for the unambiguous RC condition than any of the other three conditions. Most importantly, unambiguous RC sentences had significantly longer residual RTs than the unambiguous SC sentences ($F(1,43) = 5.136$, $MS_{within} = 2214$, $p < 0.03$; $F(1,19) = 9.593$, $MS_{within} = 1385$, $p < 0.006$). In addition, unambiguous RC sentences were significantly slower than ambiguous SC sentences ($F(1,43) = 6.537$, $MS_{within} = 2364$, $p < 0.02$; $F(1,19) = 13.97$, $MS_{within} = 1061$, $p < 0.002$) as well as ambiguous RC sentences ($F(1,43) = 7.332$, $MS_{within} = 2893$, $p < 0.01$; $F(1,19) = 18.77$, $MS_{within} = 789.9$, $p < 0.0005$). For all other comparisons, $F_s < 1$. These results are as predicted by the storage cost theories.

Table 3-4

	SC		RC	
	Ambiguous	Unambiguous	Ambiguous	Unambiguous
Matrix Subj	-31 (348)	-39 (348)	-51 (338)	-37 (349)
Verb-ing		-18 (402)		
That/Which	-12 (355)	35 (400)	-9 (359)	-13 (371)
Critical Clause	-21 (358)	-18 (360)	-26 (353)	5 (382)
Embedded VP	50 (437)	20 (411)	30 (430)	80 (491)
Matrix VP	30 (409)	30 (397)	91 (470)	66 (445)

Table 3-4. Mean residual RTs (msec / word) for each region of Experiment 1b as a function of clause type and ambiguity, by participants (raw RTs in parentheses).

One potential alternative explanation of these effects is that people may process the RC continuation more slowly because of the lexical bias favoring the SC interpretation in the nouns that were used. That is, the difficulty that people experienced in processing the RC structure could have been due to a preference to have an SC following the head noun, rather than an RC, an effect that could potentially persist through the RC itself. In order to test this potential explanation, we divided the items into two groups: those with a strong bias toward SC completions (9 items, averaging 74.8% SC completions) and those with no bias toward either SC or RC completions (11 items, averaging 52.5% SC completions). If the lexical-preference explanation of the RC difficulty is correct, then we should expect a stronger effect in the items that are more strongly SC-biased and we should see little to no effect in the equi-biased set of items.

Contrary to this prediction, we found the opposite to be true. Despite having half as much data overall, and half as many items, the effects that were observed with all items still remained in the equi-biased subset of 11 items. In particular, the unambiguous RC condition was read slower than the unambiguous SC condition in this set of 11 items ($F(1,40) = 3.110$, $MS_{\text{within}} = 4844$, $p = 0.085$; $F(1,10) = 6.613$, $MS_{\text{within}} = 1935$, $p < 0.03$), and were slower than either of the ambiguous conditions (vs. ambiguous SC: $F(1,40) = 3.756$, $MS_{\text{within}} = 4469$, $p = 0.060$; $F(1,10) = 4.814$, MS_{within}

= 1954, $p = 0.053$; vs. ambiguous RC: $F1(1,40) = 5.032$, $MS_{\text{within}} = 6241$ $p < 0.05$; $F2(1,10) = 16.63$, $MS_{\text{within}} = 1011$, $p < 0.003$). (Note that the data of three subjects had to be omitted in these analyses because of insufficient correct responses in at least one condition). In contrast, many of the comparisons in the more SC-biased item set were not quite significant, although the means were numerically in the predicted direction. For example, the comparison between unambiguous SC and unambiguous RC was numerically present, but was not significant by participants or by items ($F_s < 1.7$). These results provide strong evidence against the lexical-bias interpretation of the effects in the critical region.

We did one further analysis in order to test the lexical-bias explanation of the elevated RTs of the unambiguous RC condition. We tested for a correlation between (1) the SC-bias of an item, as provided by the completion norms in Appendix B, and (2) the residual RT difference score for each item, as determined by subtracting the mean residual RT in the critical region of the unambiguous SC condition from that of the unambiguous RC condition. Contrary to the prediction of the lexical-bias explanation, there was a negative correlation of 0.27 ($r^2=0.07$; $p>.24$) between SC-bias and the RT difference score in the region. This correlation was non-significant and was in the reverse direction to that predicted by the lexical-bias hypothesis.

3.3.3 Discussion

Across the critical region of the unambiguous RC condition, online RTs were greater when compared to the same region of the ambiguous and unambiguous SC conditions as well as the ambiguous RC condition. These results are exactly what are expected if predicting a gap following a wh-filler incurs a storage cost.

The results of Experiment 1b are anticipated by a storage cost theory which associates cost with predicted syntactic heads, as long as there are empty categories mediating wh-filler-gap dependencies. The results are also predicted by a storage cost theory that associates cost with each

incomplete dependency. The results are not predicted by a storage cost theory that associates cost only with predicted verbs or incomplete clauses.

3.4 Experiment 1c: Verb Subcategorization Predictions

Experiments 1a and 1b demonstrate that there is a storage cost associated with incomplete verbal and wh-filler dependencies. Within a predicted-head theory, these results are consistent with there being storage costs for predicted verbs and predicted empty categories in relative clause (RC) constructions. An interesting natural prediction of syntactic storage theories is that there should be greater processing load in head-initial languages like English at heads that require a greater number of arguments. For example, there should be greater processing load at a verb that generally predicts two arguments to the right (e.g. “give”) as compared to a verb that generally predicts only one argument to the right (e.g. “publish”). Some preliminary evidence in support of this prediction is provided by Shapiro, Zurif & Grimshaw (1987), who demonstrated that reaction times to a cross-modal lexical decision task were larger at ditransitive verbs like “donate” than at transitive verbs like “secure”. Unfortunately, only a small number of verbs were tested, and the method that Shapiro et al. used to classify verbs according to their argument structures was not well documented. For example, it does not appear that they used corpus searches or sentence completion norms (two current standard methods) to classify the verbs. In addition, neither Rayner & Duffy (1986) nor Schmauder, Kennison & Clifton (1991) found any complexity effects in their reading experiments that examined different verbal argument structures. But a storage cost effect may be small enough such that it is difficult to detect on just a verb, without an extended critical region.

In the current experiment, we wanted to use self-paced reading to test whether there is a measurable storage cost for some obligatory arguments of verbs. We therefore tested a lengthy region over which storage costs would differ because a short region might not result in measurable effects. As a result, we could not compare transitive and intransitive verbs in English, because the direct object of a verb must generally occur immediately adjacent to its role-assigning verb. Lengthening out the region between a verb and its object would result in materials that are poorly formed, independent of potential storage cost differences. However, there is no adjacency constraint

on prepositional phrase (PP) arguments of verbs, so we designed items to compare obligatorily ditransitive verbs to obligatorily transitive verbs. The materials tested in this experiment had the form of (17):

(17)

a. Obligatory Transitive Verb:

Mary published **a book which had impressed some critics who worked for a magazine**.

b. Optional Ditransitive Verb:

Mary read **a book which had impressed some critics who worked for a magazine** to a young child.

c. Obligatory Ditransitive Verb:

Mary gave **a book which had impressed some critics who worked for a magazine** to a young child.

In the obligatory ditransitive condition in (17c) the complex NP “a book which had impressed some critics who worked for a magazine” is the object of the verb “gave”. When an inanimate NP like “a book” follows the verb “gave”, a PP goal argument is required further downstream. In contrast, the verb “read” in (17b) is only optionally ditransitive. That is, it can take a PP argument following the NP argument “the book”, but such a PP is not required. If there is a storage cost associated with all obligatory syntactic predictions, then the obligatory ditransitive condition should be read more slowly during the critical region “the book which had impressed some critics who worked for a magazine”. A second control condition is provided in (17a), the obligatory transitive condition. No PP argument of a transitive verb like “published” is possible following the object NP “a book”. Reading times for the object NP “the book which had impressed some critics who worked for a magazine” should therefore be faster for the obligatory transitive condition as

compared with the obligatory ditransitive condition, according to storage cost theories. Reading times are not predicted to differ between the obligatory transitive and optionally ditransitive conditions because there are the same number of predicted syntactic heads / incomplete obligatory dependencies over this region in each.

We included two further conditions in this experiment to test a prediction of the DLT about integration costs. The dependency between the verb “gave” and its second argument, the PP “to a young child”, connects across a long region, the complex NP “the book which had impressed some critics who worked for a magazine”, in sentence (17c). Shortening this region by removing the RCs is predicted by the DLT to speed RTs at the PP. Consequently, we tested two further conditions, as in (18): one with a single RC between the PP and the verb, and a second with a local dependency between the PP and the verb:

(18)

a. Medium Integration Distance: one RC between PP and verb:

Mary gave a book which had impressed some critics **to a young child** who was very grateful for the present.

b. Short Integration Distance: local dependency between PP and verb:

Mary gave a book **to a young child** who was very grateful for the present and read it quickly.

The integration distance component of the DLT predicts the shortest RTs for the PP “to a young child” in (18b). RTs on the same region should be slower in (18a), and slowest in the longest integration condition, (17c). This is a potentially interesting comparison, because integration distance effects have only been demonstrated at the point of processing a verb in English, by manipulating pre-verbal material. The comparison here varies post-verbal material, such that integration costs should arise at a category other than a verb (cf. Konieczny, 2000; Vasishth, 2002).

3.4.1 Methods

Participants. Forty-eight participants from the MIT community who did not take part in Experiments 1a or 1b were paid for their involvement. All were native speakers of English and were naïve as to the purposes of the study.

Materials & Design. Twenty sets of sentences were constructed with 5 conditions each, as exemplified in (17) and (18). Each item in a set began with a proper name (e.g. “Mary”) followed by a verb which was either obligatorily ditransitive (e.g. “give”), optionally transitive / ditransitive (e.g. “read”), or obligatorily transitive (e.g. “publish”). An object NP consisting of a determiner (sometimes definite “the”, sometimes indefinite “a” or “some”) and a head noun occurred immediately following the verb. The subcategorizations of the verbs were verified using a completion study and the plausibility of the transitive and ditransitive conditions (the most crucial comparison here for storage cost theories) was matched in an off-line plausibility norming study, both described below in Section 3.4.2. For the storage conditions described in (17), the direct object NP was followed by two RCs, the first modifying the object NP (e.g. “which impressed some critics”), and the second modifying the object NP in the preceding RC (e.g. “who worked for a magazine”). The obligatory transitive condition (17a) ended at this point. The ditransitive conditions (17b) and (17c) continued with a PP argument of the main verb (e.g. “to a young child”). The PP was always a goal thematic role argument and none of the intervening verbs allowed goal arguments, so there was no ambiguity of attachment at this point. The remaining two conditions – the short and medium length integration conditions – omitted either both or the second of the RCs separating the direct object and the PP argument of the ditransitive verb. These items then continued in uncontrolled ways so as to make most versions of an item roughly the same word length overall.

The target sentences were split into five lists of 20 sentences each, balancing all factors in a Latin-Square design. Each list was combined with 72 fillers of various types. The stimuli were

pseudo-randomized independently for each participant such that at least one filler item separated any two targets.

Procedure. The procedure was the same self-paced word-by-word paradigm as in the previous two experiments. The comprehension task that followed each item consisted of simple yes / no questions.

3.4.2 Norming Methods

Participants. One hundred and twenty participants from the MIT community (60 for the plausibility rating study and 60 for the completion study) who did not take part in Experiments 1a, 1b or 1c were compensated for their involvement. All were native speakers of English and were naïve as to the purposes of the study.

Materials & Design. For the plausibility rating survey, three conditions were tested for each target item, corresponding to the obligatory ditransitive, optionally ditransitive, and obligatory transitive conditions exemplified in (17) above. The plausibility norming items consisted of the subject NP together with the verb and its complements, without any of the RCs, as shown in (19):

(19)

a. Obligatory Transitive Verb:

Mary published a book.

b. Optional Ditransitive Verb:

Mary read a book to a young child.

c. Obligatory Ditransitive Verb:

Mary gave a book to a young child.

The target sentences were separated into three lists of 20 sentences each, such that no participant saw two conditions of the same item. In addition to the target sentences, thirty fillers of various types were presented in the survey. The fillers included short and long transitive sentences (e.g. “Luke poured the juice” and “Paul woke his brother with a loud scream”) as well as short and long preposition mediated transitive sentences (e.g. “Sebastian talked to the mailman” and “Ruth agreed with the woman on the committee”). Half of the filler items were implausible while all of the target items were plausible. Each of the three lists was presented with two different pseudo-randomizations. Participants were asked to rate the sentences on a scale from 1 to 5 with 1 being very unnatural and 5 being very natural.

The same three conditions were also tested for the sentence completion survey – obligatory ditransitive, optionally ditransitive, and obligatory transitive – in order to ensure that the verbs that we used in each condition were categorized correctly. In this study, participants were presented with sentence fragments consisting of the initial proper noun, the verb and the object noun phrase for each target item, as shown in (20):

(20)

a. Obligatory Transitive Verb:

Mary published a book

b. Optional Ditransitive Verb:

Mary read a book

c. Obligatory Ditransitive Verb:

Mary gave a book

Participants were overtly asked to determine if the fragment was a complete or incomplete sentence as is. If the fragment was considered incomplete, the participants were asked to complete the item in

the first natural way that came to mind. The percentage of items with verb phrase (VP) - attached PP completions was recorded for each item.

The target items were separated into three lists of 20 items each, such that no participant saw two conditions of the same item. No filler items were included. Each of the three lists was presented with two different pseudo-randomizations.

3.4.3 Norming Results

The mean plausibility ratings and percentages of VP-attached PP completions for each condition are presented in Table 3-5. One participant was omitted from all analyses of the completion data because s/he did not follow the directions.

The plausibility rating of the obligatory ditransitive condition were greater than the optional ditransitive condition, significant by participants and marginal by items ($F(1,59) = 18.74$, $MS_{within} = 0.1522$, $p < 0.001$; $F(1,19) = 4.217$, $MS_{within} = 0.2242$, $p = 0.054$). The obligatory ditransitive condition ratings were less than the obligatory transitive condition, significant by participants but insignificant by items ($F(1,59) = 23.70$, $MS_{within} = 0.0879$, $p < 0.001$; $F(1,19) = 2.296$, $MS_{within} = 0.3415$, $p = 0.146$). Finally, the optional ditransitive condition plausibility ratings were significantly less than the obligatory transitive condition by both participants and items ($F_s > 16$, $p_s < 0.0001$). Meanwhile, the completion percentages were significantly different from each other by participants and by items for all pair-wise comparisons ($F_s > 37$, $p_s < 0.0001$).

The results of the completion study were as we expected when we constructed the items: the obligatory ditransitive conditions generally need a further PP attached to the verb phrase (VP) in order to sound complete, whereas this was much less true for the optionally ditransitive and transitive conditions. But there were two items of the twenty for which this was not the case. In these two items, the ditransitive condition had a VP-attached PP percentage of 35%. We therefore excluded these two items from later analyses. We also excluded two further items: one because of

Table 3-5

	Obligatory Transitive	Optional Ditransitive	Obligatory Ditransitive
Plausibility	4.41 (0.07)	3.84 (0.08)	4.14 (0.07)
Completion	8.57 (1.52)	26.09 (3.10)	80.23 (2.69)

Table 3-5. Mean (standard error) plausibility rating on a scale of 1 (very unnatural) to 5 (very natural) for Experiment 1c as a function of verb subcategorization, by participants. Mean (standard error) percentage of items with VP-attached PP completions as a function of verb subcategorization, by participants

low obligatory ditransitive plausibility rating, and another because of both low obligatory ditransitive plausibility rating and unclear subcategorization. The completion norms and plausibility ratings for each item are listed in Appendix C. The mean plausibility ratings and percentages of VP-attached PP completions for the remaining 16 items are presented in Table 3-6.

In this subset of items, the completion percentages were significantly different from each other by participants and by items for all pair-wise comparisons, as desired ($F_s > 28$, $p_s < 0.0001$). Importantly for the plausibility ratings, the obligatory ditransitive condition is non-significantly different from the obligatory transitive condition both by participants and by items ($F_s < 1$). The optional ditransitive condition ratings are significantly different from the other two conditions both by participants and by items ($F_s > 10$, $p_s < 0.0001$). Hence comparisons involving this condition are somewhat difficult to interpret.

3.4.4 Results

Overall, four items were omitted from all analyses based on the norming results above. Four participants' data were omitted from all analyses because of poor comprehension question answering performance (< 67% accuracy overall).

Table 3-6

	Obligatory Transitive	Optional Ditransitive	Obligatory Ditransitive
Plausibility	4.34 (0.08)	3.82 (0.08)	4.30 (0.07)
Completion	8.53 (1.66)	25.85 (3.17)	87.99 (2.23)

Table 3-6. After omitting four items for poor plausibility rating and/or poor subcategorization performance. Mean (standard error) plausibility rating on a scale of 1 (very unnatural) to 5 (very natural) for Experiment 1c as a function of verb subcategorization, by participants. Mean (standard error) percentage of items with VP-attached PP completions as a function of verb subcategorization, by participants

Comprehension Question Performance. On average, the comprehension questions for the experimental items in Experiment 1c were answered correctly in 89.4% of the trials. The percentages of correct answers for each condition are presented in Table 3-7.

Pair-wise comparisons of the task accuracy for the conditions revealed some potentially significant differences. The accuracy percentage for the short ditransitive condition was numerically better than the medium ditransitive condition, marginally significant by participants ($F(1,43) = 3.124$, $MS_{\text{within}} = 0.0227$, $p = 0.084$; $F(1,15) = 2.428$, $MS_{\text{within}} = 0.0080$, $p = 0.140$). The accuracy percentage for the short ditransitive condition was significantly better than for the long ditransitive condition by participants, though not by items ($F(1,43) = 6.438$, $MS_{\text{within}} = 0.0177$, $p < 0.02$; $F(1,15) = 2.720$, $MS_{\text{within}} = 0.0171$, $p = 0.120$) and significantly more accurate than the optional ditransitive condition both by participants and by items ($F(1,43) = 10.69$, $MS_{\text{within}} = 0.0240$, $p < 0.003$; $F(1,15) = 7.307$, $MS_{\text{within}} = 0.0141$, $p < 0.02$). There was a trend for the short ditransitive condition to be more accurate than the transitive condition, but this trend was not significant in the items analysis ($F(1,43) = 2.917$, $MS_{\text{within}} = 0.0243$, $p = 0.095$; $F(1,15) < 1.8$). For all other comparisons, $F_s < 2$.

Table 3-7

Obligatory Transitive	Optional Ditransitive	Obligatory Ditransitive (Long)	Obligatory Ditransitive (Med)	Obligatory Ditransitive (Short)
89.58 (2.66)	84.47 (3.05)	88.07 (2.69)	89.58 (2.79)	95.27 (1.87)

Table 3-7. Mean (standard error) comprehension question performance for Experiment 1c in percent correct as a function of verb subcategorization (and integration distance), by participants.

Reading Times. Only items with correctly answered comprehension questions were analyzed. Residual reading time (RT) data points that were greater than 3 standard deviations from the mean were excluded from the analysis, affecting less than 1.6% of the data in Experiment 1c.

Storage Costs:

Figure 3-3 presents the mean residual RTs per word (msec / word) across the three conditions in the verb subcategorization experiment. Table 3-8 presents the mean RTs per word for six regions: (1) the initial proper noun and matrix verb; (2) the object NP (the first component of the critical region); (3) the first RC (the second component of the critical region); (4) the second RC; (5) the first three words of the PP for the ditransitive and optionally ditransitive conditions (except for the sentence-final word in items #1, 6, 10, and 15; see Appendix C); and (6) the remaining material completing the ditransitive and optionally ditransitive sentences (omitted from Figure 3-3 for clarity).

Figure 3-3

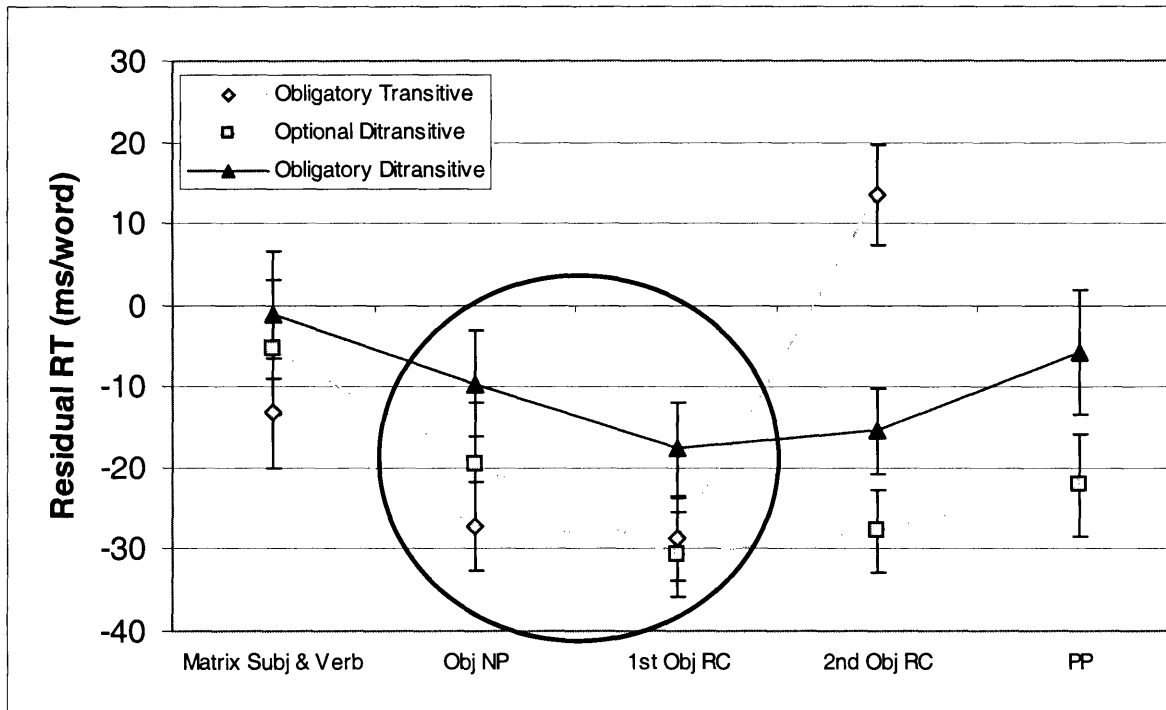


Figure 3-3. Plot of mean (standard error) residual RTs per word for each region of Experiment 1c as a function of verb subcategorization, by participants.

Table 3-8

	Verb Subcategorization		
	Obligatory Transitive	Optional Ditransitive	Obligatory Ditransitive
Matrix Subj & Verb	-13 (306)	-5 (313)	-1 (315)
Object NP	-27 (283)	-20 (289)	-10 (298)
1st Object RC	-29 (288)	-31 (286)	-18 (298)
2nd Object RC	14 (329)	-28 (284)	-15 (298)
PP		-22 (291)	-6 (306)
Completion		11 (327)	14 (335)

Table 3-8. Mean residual RTs (msec / word) for each region of Experiment 1c as a function of verb subcategorization, by participants (raw RTs in parentheses).

Reading Times in the Critical Region. The critical region for the verb subcategorization storage cost experiment consisted of the direct object NP together with the first following RC. The second RC was not included as part of the critical region because it is the sentence-final clause in the transitive condition. In the planned pair-wise comparisons of residual RTs for the critical region, the obligatory ditransitive condition was significantly slower than the obligatory transitive condition by participants and marginal by items ($F(1,43) = 5.542$, $MS_{\text{within}} = 764.1$, $p < 0.03$; $F(1,15) = 3.181$, $MS_{\text{within}} = 486.6$, $p = 0.095$). The obligatory ditransitive condition was significantly slower than the optional ditransitive condition, both by participants and by items ($F(1,43) = 4.629$, $MS_{\text{within}} = 722.3$, $p < 0.04$; $F(1,15) = 5.527$, $MS_{\text{within}} = 186.1$, $p < 0.04$). RTs for the optional ditransitive and obligatory transitive conditions were not significantly different from each other ($F_s < 1$). Thus, in the critical region, the obligatory ditransitive condition had the slowest reading times, while the optional ditransitive and obligatory transitive conditions had comparably fast reading times, as predicted by the storage cost hypothesis.

Reading Times in the Second RC. The obligatory ditransitive condition was significantly slower than the optional ditransitive condition during the second RC ($F(1,43) = 4.187$, $MS_{\text{within}} = 804.2$, $p < 0.05$; $F(1,15) = 7.941$, $MS_{\text{within}} = 125.7$, $p < 0.02$), as predicted by the storage cost hypothesis. No other comparisons were performed with the obligatory transitive condition in this region, because the second RC is sentence final for the transitive condition.

Reading Times in the PP. The PP region consisted of the first three words of the PP except for the sentence final words of items #1, 6, 10, and 15 (see Appendix C). The residual RTs of the obligatory ditransitive condition was slower than the optional ditransitive condition, marginally significant by

participants and significant by items ($F(1,43) = 3.166$, $MS_{within} = 1854$, $p = 0.082$; $F(1,15) = 10.18$, $MS_{within} = 132.4$, $p < 0.007$).

Integration Costs:

Figure 3-4 presents the mean residual RTs per word (msec / word) across the three integration distance conditions for the obligatory ditransitive verbs. Table 3-9 presents the mean residual and raw RTs per word for the six regions defined above.

Figure 3-4

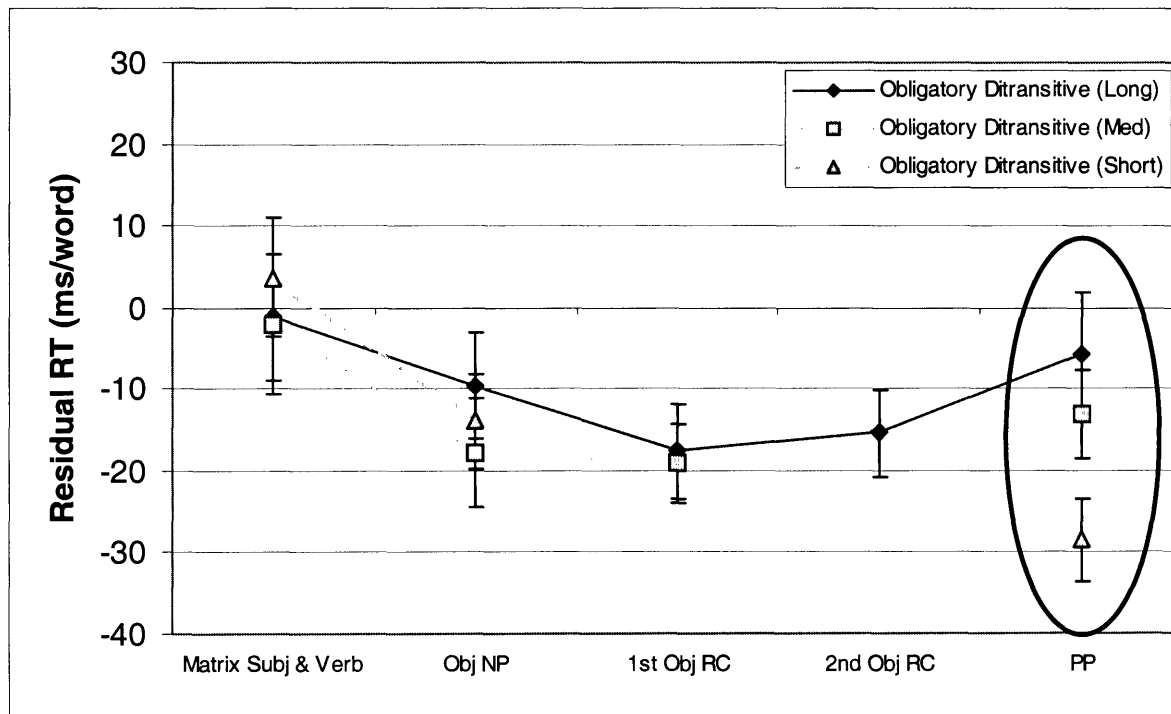


Figure 3-4. Plot of mean (standard error) residual RTs per word for each region of Experiment 1c as a function of integration distance, by participants.

Table 3-9

	Ditransitive PP Integration Distance		
	Long	Medium	Short
Matrix Subj & Verb	-1 (315)	-2 (316)	4 (322)
Object NP	-10 (298)	-18 (297)	-14 (297)
1st Object RC	-18 (298)	-19 (298)	
2nd Object RC	-15 (298)		
PP	-6 (306)	-13 (297)	-29 (287)
Completion	14 (335)	-1 (312)	-18 (299)

Table 3-9. Mean residual RTs (msec / word) for each region of Experiment 1c as a function of integration distance, by participants (raw RTs in parentheses).

Reading Times at the Integration Region. The integration region consisted of the first three words of the PP except for the sentence final words of items #1, 6, 10, and 15 (see Appendix C). In the planned pair-wise comparisons of residual RTs in the integration region, the medium integration distance condition was slower than the short integration distance condition, significant by participants and marginally significant by items ($F(1,43) = 5.889$, $MS_{within} = 880.8$, $p < 0.02$; $F(1,15) = 4.066$, $MS_{within} = 314.9$, $p = 0.062$). Residual RTs for the long integration distance condition was significantly slower than the short integration distance condition at the integration region both by participants and by items ($F(1,43) = 6.267$, $MS_{within} = 1811$, $p < 0.02$; $F(1,15) = 7.574$, $MS_{within} = 446.7$, $p < 0.02$). Finally, residual RTs for the long integration distance condition were numerically slower than for the medium integration distance condition, but insignificantly so ($F_s < 1$).

3.4.5 Discussion

The results of the storage cost comparisons were largely as predicted by the storage cost theories. Most importantly, RTs were longer for the obligatory ditransitive condition compared to either the

obligatory transitive condition or the optional ditransitive condition for the critical region consisting of the direct object of the verb followed by the first RC. Secondly, RTs were longer for the obligatory ditransitive condition compared to the optionally ditransitive condition in the region consisting of the second RC. These results support the hypothesis that there is a storage cost associated with maintaining the second argument of the obligatorily ditransitive verbs during the RCs preceding this argument.

The results of the integration cost comparisons supported components of the integration distance hypothesis. In particular, the integration between the PP and the preceding verb was processed fastest when there was no RC between the PP and direct object NP: both the long and medium distance integrations were processed more slowly. Interestingly, there was no significant difference between the medium and long distance integrations at the PP, although the long-distance integration resulted in numerically longer RTs. The lack of a difference between medium and long distance integrations is suggestive that the cost function associated with integration increases is not linear in the limit, but approaches a maximal cost, following a sigmoidal function (cf. Gibson, 1998). In general, these results confirm that English is a language in which integration costs increase at the rightward end of a head-dependent connection. Such results contrast with results from head-final languages like German (Konieczny, 2000), Hindi (Vasishth, 2002) and Japanese (Nakatani & Gibson, 2003). Additional research is needed to discover the cause for the apparent cross-linguistic differences (see Nakatani & Gibson (2003) for discussion).

3.5 General Discussion

The results from Experiment 1a demonstrated that storing incomplete verbal dependencies is associated with increased on-line processing costs as measured by reading times. The results from Experiment 1b demonstrated that a similar storage cost is associated with keeping track of wh-filler-gap dependencies while the results of Experiment 1c demonstrated a similar RT effect for keeping track of an expected PP argument of a verb. Together, the results of these three experiments are potentially consistent with a variety of on-line syntactic storage cost theories. We evaluate our experimental results together with other results in the literature with respect to four potential theories. These theories differ in terms of what types of syntactic representations incur storage costs at a parser state as follows: (1) incomplete clauses / predicted verbs (Kimball, 1973); (2) incomplete dependencies, thematic role assignments or case assignments (Hakuta, 1981; Gibson, 1991; Lewis, 1996; Stabler, 1994); (3) incomplete phrase structure (PS) rules (Chomsky & Miller, 1963); or (4) predicted syntactic heads / categories (Gibson, 1998).

The first theory, in which on-line syntactic storage cost is indexed by the number of incomplete clauses or predicted verbs at a parser state (Kimball, 1973), is insufficient to account for the results of both Experiments 1b and 1c, because wh-filler-gap dependencies and predicted PP dependencies are associated with storage cost independent of the number of incomplete clauses. Although it is possible that there are multiple different types of syntactic storage that are associated with on-line cost, such a theory is dispreferred to a simpler theory in which the same underlying storage process is responsible for both results.

Under the second class of theories, there is a storage cost for each incomplete syntactic dependency, thematic role assignment or case assignment (Hakuta, 1981; Gibson, 1991; Lewis, 1996; Stabler, 1994). Each of these theories can correctly account for the results of all three experiments presented here. Thus, this class of theories can account for the observed data under a uniform theoretical account. However, such a theory is difficult to reconcile with existing results

from the literature with respect to the processing of head-final constructions. Such a theory predicts an increasing processing load in head-final constructions, across the dependents of a head, which appears in final position. For example, in an SOV language like Japanese, processing is predicted to be slower over the processing of embedded elements depending on the number of dependents of the verb. Thus, processing embedded material (e.g. an adverbial, or an optional prepositional phrase) should be fastest when preceding verbs with one argument (intransitives), slower for verbs with two arguments (transitives) and slowest for verbs with three arguments (ditransitives). Whereas processing does slow down when an additional verb / clause is expected (e.g. in Japanese, Miyamoto, 2002), there is no evidence in head-final languages for slowed processing across verbal dependents that do not require the prediction of an additional clause. In fact, processing appears to speed up over the additional arguments and modifiers of an upcoming head in SOV languages (for results in Japanese, see Miyamoto (2002) and Nakatani & Gibson (2003); for results in Hindi, see Vasishth (2002); for results in German, see Konieczny (2000) and Konieczny & Doring (2003)). As a consequence, such a storage cost theory based on incomplete dependencies seems unlikely.

Under the third theory that we will consider, there is a storage cost for each partially processed phrase structure (PS) rule at a parse state (Chomsky & Miller, 1963), assuming a fully connected structure for the input as each word is processed. Such a theory can account for the results of Experiment 1a, because there is a partially processed PS rule of the form $S \Rightarrow NP VP$ (i.e. the subject-verb rule) for each incomplete verbal dependency. Such a theory can also account for the results of Experiment 1c, because when parsing in such a framework, there is a partially processed PS rule of the form $VP \Rightarrow Verb NP PP$ after an obligatorily ditransitive verb is processed. The comparison between optionally ditransitive verbs and obligatorily ditransitive verbs can be accounted for by having separate rules for the transitive and ditransitive continuations of optionally ditransitive verbs. Such a theory can also extend to account for the results of Experiment 1b by adopting the slash notation from recent phrase-structure grammars. Under recent context-free phrase

structure rule theories – e.g. head-driven phrase structure grammar (Pollard & Sag, 1994) and its predecessor, generalized phrase structure grammar (Gazdar et al., 1985) – wh-fillers are associated with their underlying role-assigning positions by percolating a “slash” feature through intermediate rules until an empty category can be posited in the target position. Because the slash feature indicates a missing constituent, a PS rule storage cost theory can account for the results of Experiment 1b if rules with active slash features also contribute to storage cost in addition to partially processed rules. Thus, a storage cost theory based on incompletely processed PS rules can account for the results of all three experiments naturally. However, this theory has difficulty in accounting for the data from head-final languages. In particular, it is standardly assumed that subject and object NPs are connected to the verb through different phrase structure rules, even in SOV languages, i.e. $S \Rightarrow NP VP$, and $VP \Rightarrow NP Verb$. If this is so, then storage cost should increase through the VP in transitive sentences, but this does not appear to be the case. Thus, the PS-rule based storage cost theory cannot account for the head-final data in much the same way that the incomplete dependency theory cannot.

The final account that we will evaluate is one in which there is storage cost associated with predicted syntactic heads. This theory can account for the results from all three experiments. In Experiment 1a, there is a cost associated with predicting a verb to come. In Experiment 1b, there is a cost associated with predicting an empty category to be associated with the wh-filler. And in Experiment 1c, there is cost associated with predicting the PP argument of an obligatorily ditransitive verb. In addition, this theory does not make the undesirable prediction of increasing complexity across the dependents of a single head in a head-final language: all the dependents support the prediction of the same head, and therefore there is no increasing cost. As a result, this theory can account for all of the existing data.

An interesting consequence of this discussion is that the combination of results supports a theory which includes empty categories mediating wh-filler-gap dependencies. That is, if we accept

that storage costs are indexing predicted categories rather than incomplete dependencies (because of the evidence from head-final languages), then the only way to account for the results of Experiment 1b using predicted categories is to assume the existence of a wh-trace, an empty category. If there were no wh-trace, such that the dependency were represented via a direct link between the wh-filler and the verb to come, then there would be no additional storage cost for this category prediction, because the verbal head would already be predicted by the existence of the embedded clause, and would not induce additional processing cost. Thus the three experiments, in conjunction with existing results from the processing of head-final languages, provide indirect evidence for the existence of empty categories in wh-filler dependencies (cf. Pickering & Barry, 1991; Gibson & Hickok, 1993; Gibson & Warren, 1999).

3.6 Future Work: An Application of a Model of Storage Costs

The evidence from the literature so far seems to suggest that the human sentence processing mechanism is an incremental system (Eberhard et al., 1995; Marslen-Wilson, 1973). In other words, the system attempts to construct and interpret the sentence fragment as each new word comes, without waiting for the completion of the full sentence. The incrementality, plus the potential ambiguity of the lexical items themselves (both in meaning: federal “bank” vs. river “bank”, and syntactic category: “dance”-as-action verb vs. “dance”-as-event noun), often leads to at least temporary confusion as to how the new input word should be syntactically integrated, or parsed, into the sentence. For instance, the prepositional phrase “with the telescope” in sentence (21a) could be associated with either the cop being watched as in (21b) or the action of seeing by the robber as in (21c).

(21)

- a. The robber saw the cop with the telescope.
- b. The robber saw [the cop with the telescope].
- c. The robber saw [the cop] [with the telescope].

In this particular example, the ambiguity is global, such that, even at the end of the sentence, it is not absolutely clear which meaning was intended (though readers and listeners do have a preference for interpreting the sentence as in (21b)). Other ambiguities can be ‘temporary’ in which the sentence eventually contains disambiguating material that makes the correct interpretation clear. Or if the disambiguating material comes before the ambiguity itself, the sentences are considered unambiguous. Some temporary ambiguities include the main verb / reduced relative, the direct object / sentential complement, and the sentential complement / relative clause (for a summary, see Tanenhaus & Trueswell (1995)).

Given that there is so much potential for ambiguity, just how does the human sentence processing mechanism go about building the correct syntactic constructions and making sense of them? How does it deal with these ambiguities? We have begun to explore the anatomical localization of the process of ambiguity resolution in Experiment 4, and together with a model of storage costs developed in Experiments 1a, b, and c, they can play a role in determining the serial vs. parallel nature of the sentence processing mechanism.

3.6.1 Serial & Parallel Accounts of Ambiguity Resolution

Fundamentally, the sentence processing system can deal with ambiguities in either a serial or parallel manner (for a review, see Crocker (1999)). As a serial processor it will consider just one construction at a time, in depth, and make repairs only when it discovers it has made a mistake. A parallel system, on the other hand, considers perhaps several similar constructions simultaneously, in breadth, and chooses the most appropriate one when it arrives at the disambiguating material. For serial models then, the burden of determining the correct interpretation of an ambiguous sentence is placed late on the reanalysis and repair system while parallel models place the burden at the beginning by considering several possible alternatives at once.

3.6.1.1 Serial Models

Among serial accounts, both deterministic and probabilistic models have been proposed (for a review, see Pearlmutter & Mendelsohn (1999)). In deterministic serial models, the processor develops the appropriate preferred structure through some decision criteria or heuristic. This single construction is then pursued until the processing system discovers it has made a mistake, at which point reanalysis and repair must occur. For example, in the Garden Path model the criterion that determines which structure is the first choice is based on the purely syntactic principles of Minimal

Attachment and Late Closure (summarized in Frazier & Clifton (1996)). These parsing principles are implemented in the first stage of the model while the second stage, the Thematic Processor, takes into account various other sources of information including plausible thematic role assignment and verb subcategorization. Reanalysis and repair of potential ambiguities is less defined and can come in either processing stage depending on the syntactic or semantic nature of the disambiguating material.

Probabilistic serial models use similar decision criteria and heuristics to establish the most preferred construction to analyze. However, unlike deterministic models, which will develop the same preferred structure each time it is given a syntactically similar sentence to parse, the probabilistic model will only pursue the preferred structure most of the time. Depending on the feasibility of the alternative construction as being the correct interpretation, the probabilistic model will occasionally attempt to construct the sentence in the less preferred way instead. Probabilistic serial models, as we will see, are the most difficult to experimentally differentiate from parallel approaches.

3.6.1.2 Parallel Models

Parallel accounts include fully and ranked parallel models (for a review, see Pearlmuter & Mendelsohn (1999)). In fully, unconstrained parallel models, the parsing system considers every possible syntactic tree construction simultaneously and in an equally weighted manner. Then, when the ambiguity is clarified, the parser simply chooses the appropriate construction or eliminates the inappropriate ones. Given the clear presence of garden path effects in which the processor encounters measurable processing difficulties, the ability to freely choose the appropriate construction does not seem plausible and most researchers do not seriously consider this instantiation of the parallel approach.

In the case of ranked parallel models, some of the parallel constructions are deemed much more likely to be appropriate than others, based on the available information. But, all of the constructions are still pursued by the processor as possible options. The measurable increase in reading times from garden paths come when the rankings need to be reordered and less preferred structures are reactivated when disambiguating material is encountered. As a result, deliberation at the point of disambiguation of the more preferred construction should be easier and less time-consuming than constructions that are not preferred.

3.6.2 State of the Research

3.6.2.1 The Garden Path Model & Constraint-based Approaches

Much of the evidence for the serial or parallel nature of the parser has come as a consequence of the debate between the Garden Path model and Constraint-based approaches to describing the processor (Ferreira & Clifton, 1986; Trueswell et al., 1994). The Garden Path model is a serial model in which the processing mechanism considers only one structure at a time. The criterion that determines which structure is the first choice is based on the purely syntactic principles of Minimal Attachment and Late Closure (summary in Frazier & Clifton, 1996). These parsing principles are implemented in the first stage of the model while the second stage, the Thematic Processor, takes into account various other sources of information including plausible thematic role assignment and verb subcategorization. Reanalysis and repair of potential ambiguities comes in either processing stage depending on the syntactic or semantic nature of the disambiguating material. Constraint-based approaches, on the other hand, assume a parallel system in which all the possible syntactic constructions are considered at once (e.g. MacDonald et al., 1994). All the constraints, both syntactic and semantic information, apply from the beginning and give certain constructions more weight than others, making them easier and less time-consuming to process.

3.6.2.2 Modularity

The crucial difference examined and reported in the literature between the two approaches then is the modularity of the information that influences the processing system. Does syntactic and semantic information have temporally distinct influences on the human parser? Ultimately, the serial or parallel nature of the processing architecture has been based on arguments about the primacy of syntactic information and the time-course of syntactic and semantic information use. The original works supporting the Garden Path model found that semantic information such as animacy of the noun or verb subcategorization did not measurably affect sentence processing, as evidenced by self-paced reading and eye-tracking (Ferriera & Clifton, 1986). However, subsequent experiments appeared to support Constraint-based proposals by showing that, with the same stimuli that controlled for the appropriate factors, semantic information is indeed used as quickly as measurably possible (Trueswell et al., 1994). Though it seems clearer that multiple information sources are used as soon as measurably possible in determining the preferred sentence construction, other work continues to provide evidence that apparently reflects delay between the application of syntactic and semantic information, and the debate continues (McElree & Griffith, 1995; Binder & Price, 2001).

3.6.2.3 A Recent Study

In the end though, the characteristics (types, strength, and time-course) of information use and issues of modularity do not directly answer the question of the serial or parallel nature of the processor. The results from such studies are entirely compatible with both views. For instance, one can imagine a serial parser pursuing only one construction that is chosen based on all the available information. On the other hand, a parallel parser could begin with multiple alternative constructions based on

syntactic information alone only to be modified or selected later using semantic information, similar to the Referential Theory of Altmann & Steedman (1988) or the Concurrent model of Boland (1997).

Though few studies have attempted a direct approach to the serial vs. parallel question, recent efforts include an experiment by Pearlmuter & Mendelsohn (1999) in which the authors attempted to directly detect the presence of the less preferred alternative construction in both the ambiguous and unambiguous regions using thematic plausibility information. To do this, the sentential complement / relative clause (SC/RC) ambiguity was used. Pearlmuter & Mendelsohn presented sentences to subjects that were all disambiguated as the preferred SC construction. However, the embedded verbs of these sentences were manipulated to make the less preferred alternative RC reading either implausible or as equally likely as the SC interpretation.

(22)

a. RC or SC plausible

[The claim [that the cop *ignored* the informant] [was true]]

b. RC implausible

[The claim [that the cop *shot* the informant] [was true]]

The embedded verb “ignored” in sentence (22a) can be plausibly completed as either SC or RC sentences. However, “shot” in sentence (22b) cannot be plausibly completed as an RC sentence because its direct object refers back to the claim, creating the improbable proposition that “the cop shot the claim”.

A parallel system would show an effect of the plausibility manipulation because the processor would consider the RC alternative as well. Therefore, there should be a measurable effect at the embedded verb when the system realizes the RC interpretation is untenable. This effect is exactly what the authors claim to have found as the embedded verbs that were not plausibly

continued as RCs (as in sentence (22b)) had slower reading times relative to unambiguous controls than those that were equally plausible as RCs or SCs (as in sentence (22a)). However, reanalysis effects measured at the semantic disambiguation point of the embedded verb (“ignored” and “shot”) could be the result of the reanalysis of the simultaneously pursued, less preferred construction in a ranked parallel model, or the result of the occasional pursuit of the less preferred construction in the probabilistic serial model.

To differentiate between probabilistic serial and ranked parallel models, Pearlmuter & Mendelsohn analyzed the correlation between SC-bias of the noun and ambiguity effect in the same sentences as above at both the embedded verb “shot” for the implausible conditions (sentence (22b)) and the direct object “the informant” for the plausible conditions (sentence (22a)). For the former, the embedded verb is the point of semantic plausibility disambiguation, while for the latter the direct object is the point of syntactic disambiguation. They reasoned that in probabilistic serial accounts, the SC and RC interpretations should be in complementary distributions because the parser pursues only one interpretation or the other. The model would then predict that as the SC-bias of the verb increases, the ambiguity effect at the disambiguating direct object decreases for the plausible conditions because the probabilistic model is less likely to pursue the inappropriate RC completion. However, competitive ranked parallel models also make the same prediction because of the complementary nature of competition.

Crucial to differentiating between probabilistic serial and non-competitive ranked parallel models is the authors’ prediction that the same SC-bias correlation should be seen at the embedded verb of the implausible conditions if the parser is probabilistic serial while no such correlation would be evident for non-competitive ranked parallel parser. This is because the same complementary distribution should be evident at this semantic point of disambiguation. The authors, however, did not find identical correlations to SC-bias at both the direct object and the embedded verb. In fact, there was no reliable correlation of SC-bias at the embedded verb of the implausible conditions at all.

They arrived at the conclusion that the human sentence processor is a non-competitive ranked parallel system.

The conclusion ultimately rests upon a null result, namely that there is no SC-bias correlation at the embedded verb of the implausible conditions. Otherwise, there is still no distinction between probabilistic serial and non-competitive ranked parallel models, let alone competitive ranked parallel models. Furthermore, the parallel model proposed here is very different from most, both in terms of the reading time cost to eliminate a less preferred alternative construction and the model's non-competitive characteristics. As a result, the debate between serial and parallel nature of sentence processing continues.

3.6.3 Alternative Constructions in the Ambiguous Region

One promising way to differentiate between ranked parallel and probabilistic serial models is to search for the presence of alternative structures in the region of ambiguity. Lewis (2000) proposed that in a parallel system, the existence of additional constructions should increase the processing load and have a measurable effect on reading times in the ambiguous region. Embedding multiple ambiguities within each other can take full advantage of this concept, expanding the number of alternative constructions a parallel system would need to consider and allocate resources to. It can then be predicted that a parallel system would show increased reading times at the most embedded region of ambiguity with greater numbers of embeddings, whereas a serial system, which still considers only the most preferred construction based on whatever criteria is most appropriate, would show no effect.

Such an examination of the sentence processing system requires a clear theory of online resource allocation and use. In particular, the theory must provide a prediction regarding which aspect of processing an additional construction increases reading times and at what magnitude. The

Dependency Locality Theory (DLT) provides such a specific description of the predicted resource costs in the ambiguous region of a sentence through the storage cost portion of the theory.

Utilizing the results of storage costs described in Experiment 1, we predict that the presence of storage costs associated with the syntactic predictions required by the alternative construction should result in increased reading times. Based on preliminary experiments, instead of multiply nested ambiguities we plan to use only a single ambiguity (Grodner et al., 2001). This is because one extra verb prediction by an alternative construction should be enough to produce a significant storage cost effect.

Stimuli

(23)

a. Ambiguous, Verb (SC-bias) Interpretation:

[The student [_{VP} *wishes* **at the meeting** [_{SC} that the administration would listen to him]]].

b. Unambiguous, Verb (SC-bias) Interpretation, Low Storage Cost:

[The student [_{VP} had wished **at the meeting** [_{SC} that the administration would listen to him]]].

c. Ambiguous, Noun Interpretation:

[The student *wishes* **at the meeting** [_{VP} were ignored by the administration]].

d. Unambiguous, Noun Interpretation, Mid Storage Cost:

[The student's wishes **at the meeting** [_{VP} were ignored by the administration]].

e. Unambiguous, Embedded Verb (SC-bias) Interpretation, Mid Storage Cost:

[The reporter's claim [_{SC} that the student [_{VP} had wished **at the meeting** [_{SC} that the administration would listen to him]]] [_{VP} was unfounded]].

f. Unambiguous, Embedded Noun Interpretation, High Storage Cost:

[The reporter's claim [_{SC} that the student's wishes **at the meeting** [_{VP} were ignored by the administration]] [_{VP} was unfounded]].

Predictions. The stimuli in this experiment rely on verbs like “wishes” that take a sentential complement as an argument and the lexical ambiguity with their nominalizations. The initial sentence fragment, “The student wishes at the meeting”, is ambiguous between the verb “wishes” that takes a sentential complement as in sentence (23a) and the nominalization of “wishes” in a simple active sentence as in (23c), both included for completeness though they are identical through the critical region. Sentence (23b) is the unambiguous equivalent of sentence (23a), the subcategorization of the verb “wishes” predicts at least another noun and verb in the sentential complement to complete a grammatical sentence. Sentence (23d) is the unambiguous version of sentence (23c) and the nominalization of “wishes” predicts only a matrix verb to complete the fragment grammatically.

If we hypothesize a parallel processing system, the presentation of the ambiguous sentence fragment “The student wishes at the meeting” may lead the system to pursue both sentence (23a) and sentence (23c) simultaneously as possible completions. If this is the case, then the system must maintain 1 noun and, critically, 2 verb predictions for both alternative constructions combined during the length of the critical ambiguous region (bold type). One verb will be required by the subcategorization frame of the verb form for the “wishes” while one verb will syntactically required by the clause initial noun form of “wishes”. These two different verbs being predicted could then conceivable slow the processor down. For sentences (23b) and (23d), the two unambiguous controls, only one verb prediction needs to be maintained. Just as in Experiment 1a, we would expect an increase in reading times for the region of interest when there are two verb predictions being maintained for the ambiguous fragment versus one for either of the unambiguous fragments. The predictions here also crucially depend on the results of Experiment 1c in which we examine if storage costs are sensitive to the predictions of verb subcategorization frames.

If we hypothesize a serial system, on the other hand, we expect no significant differences in reading times between ambiguous and unambiguous sentences. We expect no differences because, with the presentation of the ambiguous sentence fragment, only one of the two alternatives is pursued by the serial system and both the verb and nominalization interpretations of “wishes” predict just one critical verb over the ambiguous region. This storage cost prediction is the same as the unambiguous controls and as a result we predict no measurable differences in the ambiguous region between ambiguous and unambiguous sentence fragments.

Because of the possibility of a null result for a serial parser, embedded controls are included in the stimulus set (sentences (23e) and (23f)). For sentences (23e) and (23f) respectively, the unambiguous sentences (23b) and (23d) are embedded as SCs to the nominalized form of “claim”. As described in Experiment 1a, the DLT predicts that the nominalization of “claim” requires the processor to maintain an additional matrix verb prediction throughout the embedded SC, including the region of interest, “at the meeting”. Thus, in terms of storage cost load, these unambiguous, embedded control sentences should behave like ambiguous sentences in a parallel processor in that they both should have two VP predictions in the region of interest. Therefore, if the sentence processor is parallel in nature, the DLT predicts reading times will be equivalent for ambiguous sentences (23a) and (23c) as well as embedded control sentences (23e) and (23f). For unambiguous control sentences (23b) and (23d), however, these reaction times should be lower in magnitude as only one VP prediction is maintained. If the sentence processor is serial in nature, on the other hand, the DLT predicts that ambiguous sentences (23a) and (23c) should match the lower reaction times of the single VP prediction for sentences (23b) and (23d) instead. Meanwhile, unambiguous embedded control sentences (23e) and (23f) will still have higher reaction times for maintaining two VP predictions through the region of interest, “at the meeting”.

Chapter 4

Neuroimaging Studies on Syntactic Storage Costs

4.1 Introduction

Having behaviorally established the DLT model of syntactic storage costs, the next step is to begin to identify the neurological basis of processes associated with syntactic storage and integration costs. Some previous work has been done on the localization of these processes. Deficit-lesion correlations have identified the dominant (left) perisylvian cortex as the general locus for the operations of syntactic processing. Some researchers have argued for further localization of specific syntactic operations within the dominant perisylvian cortex. For instance, it has been claimed that the co-indexation of traces is selectively disrupted by lesions in and around the left inferior frontal cortex (Broca's area) (Grodzinsky, 2000; Swinney & Zurif, 1995; Swinney et al., 1996). The “co-indexation of traces” using terms from Chomsky's (1981) theory of syntax can be informally described as the linking of noun phrases to distant positions in sentences. In sentence (24) we have a relative clause, marked by the square brackets, embedded within the main clause. The co-indexation of traces, indicated by the subscripted “i”, connects the noun phrase “the reporter” to the gap in the embedded relative clause where it intuitively belongs as the object of the senator's attack.

(24)

The reporter, [that the senator attacked ____,] admitted the error.

This strong localization view of the co-indexation of traces is heavily contested (Caplan, 1995; Caramazza et al., 2001). Deficits in syntactic processing, including in the co-indexation of traces, have been associated with lesions throughout the left perisylvian region and with all aphasic syndromes (Berndt et al., 1996; Caplan et al., 1985, 1996, 1997), suggesting either distributed models (Bates & Goodman, 1997; Damasio, 1992; Mesulam, 1990; Dick et al., 2001) or models that postulate individual variability in the functional neuroanatomy of this and other aspects of syntactic processing (Caplan et al., 1985, 1996).

A different methodology, functional neuroimaging, has begun to provide data relevant to the localization of syntactic processing, and the results also point to the importance of Broca's area for syntactic processing. Many of the initial neuroimaging studies compared stimulus sentences to simple baseline tasks (Mazoyer et al., 1993; Bookheimer et al., 1993; Stowe et al., 1994; Bavelier et al., 1997; Carpenter et al., 1999). Because of the simplicity of the baseline tasks, the experimental designs can reveal activity attributable to sentence processing in general but precise syntactic processing localization conclusions cannot be drawn reliably. A small number of studies have focused more specifically on the neurological substrates of different aspects of syntactic processing, and this literature has confirmed a role for Broca's area in several syntactic operations, including the detection of agreement violations (Ni et al., 2000) and transitivity violations (Noguchi et al., 2002), ambiguity resolution (Stowe et al., 1994), and computing word order canonicity (Röder et al., 2002) and predicate relations and pronominal co-reference (Hashimoto & Sakai, 2002).

The most sustained series of functional neuroimaging studies has focused on the process associated with the co-indexation of traces. Ben-Shachar et al. (in press) showed activation in

Broca's area, as well as Wernicke's area and its right hemisphere homologue, when contrasting sentences in Hebrew with relative clauses (which contain traces) to sentential complements (which do not). Dapretto & Bookheimer (1999) found activation in Broca's area for synonymity judgments about the meaning of active and passive sentences compared to a baseline resting condition and a word synonymity judgment task. They concluded specifically that BA 44 was responsible for processing sentence-level syntactic structures, including passives, which in Chomsky's (1981) model contain traces that have to be co-indexed with the subject noun phrase.

Broca's area has been linked specifically to the underlying memory demands of the co-indexation of traces as opposed to the process itself (Stowe et al., 1998; see Stowe, 2000, for brief review). Just et al. (1996) reported increased Blood Oxygenation Level Dependant (BOLD) signal in the area when participants answered questions about sentences with relative clauses in which the co-indexation of traces impose different memory demands. The authors compared so-called subject-object (SO) structures (sentence (25a)), in which the relative clause is attached to the **subject** of the main clause and the gap is in the **object** position within the relative clause, against subject-subject (SS) structures (sentence (25b)):

(25)

a. Subject-Object (SO) Structure:

The reporter_i [that the senator attacked _____i] admitted the error.

b. Subject-Subject (SS) Structure:

The reporter_i [that _____i attacked the senator] admitted the error.

The SO structure is theorized to impose greater memory demands than the SS structure because of the greater distance between the gap and the noun phrase it refers to. Stromswold et al. (1996) and Caplan et al. (1998, 1999, 2000) reported a similar activation locus in Broca's area using PET when

participants made plausibility judgments about SO sentences compared to object-subject (OS) sentences, shown in (26).

(26)

a. Subject-Object (SO) Structure:

The juice_i [that the child spilled _____i] stained the rug.

b. Object-Subject (OS) Structure:

The child spilled the juice_i [that _____i stained the rug].

In addition to the greater integration costs associated with the distance between the noun phrase and its gap, the SO structure also has the relative clause embedded within the main clause while the OS structure has the relative clause attached to the right. Embedded clauses are believed to impose their own storage cost demands because the interruption in the processing of the main clause requires the parser to maintain the prediction of a verb in memory (for discussion, see Chapter 3).

In addition to Broca's area in the dominant (left) inferior frontal cortex, regions in the posterior perisylvian cortex, including Wernicke's area in the posterior portion of the superior temporal gyrus, and the inferior parietal cortex, have also been activated in some of these studies but not all. For the less consistent posterior perisylvian activation, a variety of results have been reported, including: (1) selective activation of Broca's area only (Stromswold et al., 1996; Caplan et al., 1998, 1999, 2000; Hashimoto & Sakai, 2002; Homae et al., 2002); (2) activation of Broca's area and its right hemisphere homologue (Waters et al., 2003, in fast responders); (3) activation of Broca's and Wernicke's area as well as their right hemisphere homologues (Just et al., 1996); and (4) activation of parietal structures only (Caplan et al., 2002; Waters et al., 2003; Caplan et al., 2003, for slow responding subjects).

Caplan et al. (2002) reported exclusively posterior perisylvian activation (in inferior parietal cortex and, to a lesser degree, Wernicke's area) in a plausibility judgment task on SO and SS structures (seen in (27)) with long filler phrases in italics, using an externally paced word-by-word presentation in an event-related design with fMRI.

(27)

a. Subject-Object (SO) Structure:

The reporter _i *covering the big story carefully* [who the photographer admired _____i] appreciated the award.

b. Subject-Subject (SS) Structure:

The reporter _i *covering the big story carefully* [who _____i admired the photographer] appreciated the award.

Caplan et al. (2002) noted that task demands could have loaded a non-syntactic short-term verbal memory system in those studies that showed posterior perisylvian activation, lumping together activation in the inferior parietal cortex and Wernicke's area. The task demand load was especially high in these experiments because of the slow presentation of the sentence, one word at a time, as well as the sizable word length of the sentences. In Just et al. (1996), the answering of a comprehension question after each sentence also has a higher task memory load because of the need to read a second sentence and decide whether it accurately represents part of the propositional content in the first. In contrast, in Stromswold et al. (1996) and similar studies, participants were asked to make plausibility judgments while the sentences were still displayed before them. Caplan et al. (2002) proposed two working memory systems, each associated with a different location in the perisylvian cortex. Activation in Broca's area represented syntactic processing load alone, while

activation in Wernicke's area and the inferior parietal cortex represented task-related short term memory load in addition to the syntactic processing load.

This explanation does not account for the lack of Broca's area activation in the Caplan et al. (2002) study or the presence of activation in Wernicke's area and its non-dominant (right) hemisphere homologue in the Ben-Shachar et al. (in press) study. Furthermore, in a study relevant to the issues discussed here, Cooke et al. (2001) attempted to vary the memory demands of syntactic processing and short-term memory independently, contrasting SO and SS sentences and inserting additional material in the sentences to manipulate the distance between the subject noun phrase (NP) and the relative clause (RC) gap, as shown in (28).

(28)

a. Subject-Object (SO) Structure, Short Distance to Gap:

The flower girl_i [who Andy punched ____i] in the arm was five years old.

b. Subject-Object (SO) Structure, Long Distance to Gap:

The messy boy_i [who Janet *the very popular hairdresser* grabbed ____i] was extremely hairy.

c. Subject-Subject (SS) Structure, Short Distance to Gap:

The strange man_i *in black* [who ____i adored Sue] was rather sinister in appearance.

d. Subject-Subject (SS) Structure, Long Distance to Gap:

The cowboy_i *with the bright gold front tooth* [who ____i rescued Julia] was adventurous.

In an event-related fMRI experiment with word-by-word presentation at two separate rates, participants had to determine the gender of the agent of the action in the sentence. The only contrasts in which differential Broca's area activation (BA 47) was seen were the long SO vs. short SS and the long SO vs. short SO comparisons ((28b) – (28c) and (28b) – (28a), respectively). Cooke et al. (2001) anticipated that the additional words in the long distance-to-gap sentences would impose

demands on non-syntactic short term memory; the system Caplan et al. (2002) suggested was activated in their study and in Just et al. (1996). From their results, Cooke et al. (2001) concluded that Broca's area activation reflects a super-additive interaction of cognitive resources necessary for the maintenance of items in short term memory and the processing of syntactically complex sentences, counter to the hypothesis of Caplan et al. (2002) that Broca's area activation reflects sentence-level syntactic processing load alone while the interaction of high task-related short term memory demands and syntactic processing involves the posterior perisylvian cortex. However, the Cooke et al. (2001) study confounded syntactic complexity and short term memory load. The long SO condition such as sentence (28b) is the only sentence type in which the additional material used to lengthen the distance to the gap follows a proper noun. This material is an appositive construction that adds to the syntactic complexity of the SO structure. Thus, the manipulation of the distance to the gap did not simply increase short term memory load in the SO structure, but also increased the syntactic complexity of that structure. Accordingly, it is possible that the Broca's area activation seen by Cooke et al. (2001) reflects the syntactic processing necessary to comprehend the appositive construction and is not the result of interactions with short term memory demands.

Another account of the inconsistent presence of posterior perisylvian cortex activation found in these studies suggests that individual differences in syntactic processing ability may be related to patterns of vascular responses to syntactic processing. Using the same paradigm as Stromswold et al. (1996), Waters et al. (2003) found that participants who responded more slowly in making judgments about plausibility of sentences activated Wernicke's area, whereas participants who responded more quickly activated Broca's area and its right hemisphere homologue. Caplan et al. (2003) reported similar effects of individual differences in both young and old participants: slow responding participants activated left parietal cortex and fast responding participants activated Broca's area, regardless of age. It is not possible to compare participants who have been tested on different paradigms with respect to their processing efficiency, so whether this hypothesis accounts

for the different patterns of vascular responses to syntactic operations related to the co-indexation of traces seen in all the functional neuroimaging studies reviewed here cannot be determined.

In summary, the syntactic operation that has been most intensively studied using functional neuroimaging techniques is the co-indexation of traces. In a study that contrasted sentences with and without traces, Ben-Shachar et al. (in press) found increased vascular response in Broca's area and in Wernicke's area and its right hemisphere homologue. For contrasts that compared sentences in which the co-indexation of traces was more difficult (because of a sentence-internal memory load), a variety of results have been reported: (1) selective activation of Broca's area; (2) activation of Broca's area and its right hemisphere homologue; (3) activation of Broca's and Wernicke's area as well as their right hemisphere homologues; and (4) activation of parietal structures only. The reasons for these different patterns of activation are not clear and it is the purpose of this experiment to take another look at activation in Broca's area and the posterior perisylvian region for the SO/OS contrast using event-related fMRI.

4.2 Experiment 2: The SO/OS Contrast

The present study provides additional data regarding vascular responses to syntactic processing. We replicated the experiment of Stromswold et al. (1996) with event-related fMRI. This replication serves several purposes.

First, it allows us to see whether differences in results obtained with this plausibility judgment task and other tasks could be the result of differences in imaging technique. All the results reported with whole sentence presentation, using this plausibility task and these materials, have shown exclusively inferior frontal activation while utilizing PET, whereas studies testing slightly different, but similar, structures using fMRI have produced more posterior perisylvian activation. Second, the event-related design (for a brief description, see Chapter 2) gives us the capability to examine BOLD signal response to the plausible sentences alone. This makes the results more directly comparable both to studies such as Just et al. (1996) and Cooke et al. (2001), which only used plausible sentences, and to studies by Caplan et al. (2002) and Ben-Shachar et al. (in press), which isolated vascular responses to plausible sentences in paradigms that contained both plausible and implausible materials. Examining plausible and implausible sentences separately eliminates the concern that previous PET results with this paradigm were due to processes associated only with implausible sentences. Third, the event-related fMRI design eliminates any potential strategy effects induced by a block design. Though there is no evidence of significant use of strategies to accomplish the judgment task in previous PET studies with this paradigm (experimental controls eliminated the possibility that any identified strategy could result in above chance performance on the judgment task, and previous PET studies did not find effects of block, indicating that participants did not develop unknown strategies over the course of the experiments), this is an additional reason to undertake this study. Finally, event-related fMRI allows us to examine the hemodynamic response function. This may document differences in the time course of any vascular responses that occur in frontal, temporal, and parietal regions, and in the left and right hemispheres, which may be

useful in interpreting regional vascular responses in relation to psychological processes such as initial parsing and retention of the verbatim content of sentences in short term memory.

4.2.1 Methods

Participants. Twelve participants from the Massachusetts General Hospital community (10 female, 2 male; mean age 22.7 years, range 18-25; mean years of education 15.5, range 13-16) were paid for their involvement (see Appendix L). All were right-handed, native speakers of English and naïve as to the purposes of the study. Informed consent was obtained from all participants. Gender of the participants was not closely controlled because earlier results in PET showed no differences in activation localization for the SO vs. OS contrast (Stromswold et al., 1996; Caplan et al., 1998).

Materials & Design. The experimental items consisted of 144 pairs of SO and OS sentences from Stromswold et al. (1996), illustrated here in (29), for a total of 288 items (see Appendix D).

(29)

a. Plausible Subject-Object (SO) Structure:

The juice_i [that the child spilled _____i] stained the rug.

b. Plausible Object-Subject (OS) Structure:

The child spilled the juice_i [that _____i stained the rug].

(30)

a. Implausible Subject-Object (SO) Structure:

The storm_i [that the man drenched _____i] brought the umbrella.

b. Implausible Object-Subject (OS) Structure:

The man brought the umbrella_i [that _____i drenched the storm].

Each matching pair of sentences had the same propositional scenario and identical lexical items, advantages provided by using the SO vs. OS contrast. In one quarter of the SO and OS sentence pairs: (1) both items were plausible; (2) both items were implausible; (3) the SO sentence was plausible while the OS sentence was implausible; and (4) the SO sentence was implausible while the OS sentence was plausible. In total, half of the items in each condition was plausible and the other half was implausible. The violations in plausibility that are illustrated in (30) were the result of mismatches in the animacy required by the matrix verb and the animacy of the subject or object noun of the matrix clause. The items from Stromswold et al. (1996) also controlled for several other factors to minimize possible strategies in the block design of their PET study, including the animacy of the subject and object noun phrases and the point at which the sentence became implausible.

Procedures:

Stimulus Presentation. Each stimulus sentence item was visually displayed in its entirety on a single line in the center of the screen. A given experimental trial consisted of a brief 300 ms fixation cross (a centered “+”), followed by the sentence item presented for 5.5 sec and a final 200 ms inter-trial interval for a total trial length of 6 sec. The task for the participants during the experimental trial was to read the sentence and judge the plausibility (plausible or implausible) of the presented item as quickly and accurately as possible. A plausible sentence was described as a sentence that had a meaning the participant could imagine happening in the real world. The pseudo-randomized item presentation in the event-related design was according to a computer program developed to randomize trial types and vary the duration of the fixation trials for optimum efficiency in the deconvolution and estimation of the hemodynamic response (Burock et al., 1998; Dale, 1999). Thus, randomly interspersed between each 6 sec sentence trial was a 2, 4, 6, or 8 sec fixation trial.

The 288 stimulus items interspersed with the fixation trials were divided into 8 runs. No pair of matched SO and OS sentences were presented in the same run. Participants were allowed a short break between each run. The sentences were projected to the back of the scanner using a Sharp LCD projector and viewed by the participants as a reflection in a mirror attached to the head coil. Responses of “plausible” and “implausible” were recorded via a custom-designed, magnet compatible button box. A Dell Inspiron 4000 computer running the E-Prime v1.0 software package (Psychology Software Tools, Inc., Pittsburgh, PA) was used to both present the stimuli and record the response and reaction times.

MR Imaging Parameters. Participants were scanned in two separate sessions. In both scanning sessions, participants were placed in the standard Siemens quadrature head coil. In the structural session, two sets of high-resolution anatomical images were acquired in a 1.5T whole-body Siemens Sonata scanner (Siemens Medical Systems, Iselin, NJ) using a T1-weighted MP-RAGE sequence (Time to Repetition (TR) = 7.25 ms, Time to Echo (TE) = 3.0 ms, and flip angle = 7°). Volumes consisted of 128 sagittal slices with an effective thickness of 1.33mm. The in-plane resolution was 1.33 x 1.0 mm (192 x 256 matrix, 256 mm Field of View (FOV)).

The functional session utilized a 3.0T head-only Siemens Allegra scanner (Siemens Medical Systems, Iselin, NJ). Scans included a lower-resolution T1-weighted MP-RAGE sequence, (TR = 11.08 ms, TE = 4.3 ms, flip angle = 8°) which consisted of 128 sagittal slices (effective slice thickness = 1.33 mm, matrix size = 128 x 256, FOV = 256 mm, in-plane resolution = 2.0 mm x 1.0mm). A T1-weighted inversion-recovery echo-planar sequence (TR = 6s, TE = 29 ms, flip angle = 90°) with 33 slices aligned parallel to the line defined by the anterior- and posterior-commissures (AC-PC) was also acquired to aid registration of the functional images with the high resolution anatomical images. The slices were effectively 3 mm thick and had a distance of 0.9 mm between slices. The in-plane resolution was 3.13 x 3.13 mm (64 x 64 matrix, 200 mm FOV).

The functional volume acquisitions utilized a T2*-weighted gradient-echo pulse sequence (TR = 2 s, TE = 25 ms, and flip angle = 90°). The volume was comprised of 33 transverse slices aligned along the same AC-PC plane as the registration volume. The interleaved slices were effectively 3mm thick with a distance of 0.9mm between slices. The in-plane resolution was 3.13 x 3.13 mm (64 x 64 matrix, 200 mm FOV). Each run consisted of 120 such volume acquisitions for a total of 3960 images. The 33 slices of a single volume took the entire TR (2s) to be fully acquired and a new volume was initiated every TR by definition. An initial 8 s (4 TR equivalent) buffer of radio frequency (RF) pulse activations, during which no stimulus items were presented and no functional volumes were acquired, was employed to ensure maximal signal during the length of the functional run.

Cortical Surface Reconstruction. High-resolution anatomical MP-RAGE scans were used to construct a model of each participant's cortical surface. An average of the two structural scans was used to maximize the signal to noise ratio. The cortical reconstruction procedure involved: (1) segmentation of the cortical white matter; (2) tessellation of the estimated border between gray and white matter, providing a geometrical representation for the cortical surface of each participant; and (3) inflation of the folded surface tessellation to unfold cortical sulci, allowing visualization of cortical activation in both the gyri and sulci simultaneously (Dale et al., 1999; Fischl et al., 1999a, 1999b).

For purposes of inter-subject averaging, the reconstructed surface for each participant was morphed onto an average spherical representation. This procedure optimally aligns sulcal and gyral features across participants, while minimizing metric distortion, and establishes a spherical-based coordinate system onto which the selective averages and variances of each participant's functional data can be resampled (Fischl et al., 1999a, 1999b). This non-rigid surface-based deformation procedure results in a substantial reduction in anatomical and functional variability across participants relative

to the more commonly used normalization approach of Talairach (Talairach & Tournoux, 1988), thereby improving the anatomical precision of the inter-participant averages.

Functional Pre-processing. Pre-processing and statistical analysis of the functional MRI data was performed using the FreeSurfer Functional Analysis Stream (FS-FAST) developed at the Martinos Center (Burock & Dale, 2000). (See Appendix G for an overview of the FS-FAST analysis stream and Appendix J for a summary of the FS-FAST commands). For each participant, the acquired native functional volumes were first corrected for potential motion of the participant using the AFNI algorithm (Cox, 1996). Next, the functional volumes were spatially smoothed using a 3-D Gaussian filter with a full-width half-max (FWHM) of 6mm. Global intensity variations across runs and participants were removed by rescaling all voxels and time points of each run such that the mean in-brain intensity was fixed at an arbitrary value of 1000.

The functional images for each participant were analyzed with a General Linear Model (GLM) using a finite impulse response (FIR) model of the event-related hemodynamic response (Burock & Dale, 2000). The FIR gives an estimate of the hemodynamic response average at each TR within a peristimulus window. The FIR does not make any assumption about the shape of the hemodynamic response. Mean offset and linear trend regressors were included to remove low-frequency drift. The autocorrelation function of the residual error, averaged across all brain voxels, was used to compute a global whitening filter in order to account for the intrinsic serial autocorrelation in fMRI noise. The GLM parameter estimates and residual error variances of each participant's functional data were resampled onto his or her inflated cortical surface and into the spherical coordinate system using the surface transforms described above. Each participant's data were then smoothed on the surface tessellation using an iterative nearest-neighbor averaging procedure equivalent to applying a two-dimensional Gaussian smoothing kernel with a FWHM of

approximately 8.5 mm. Because this smoothing procedure was restricted to the cortical surface, averaging data across sulci or outside gray matter was avoided.

Functional Statistical Analysis. Contrasts of interest were tested at each voxel on the spherical surface across the group using a random effects model of the cross-participant variance of the FIR parameter estimates. Contrasts were constructed over a range of post-stimulus delays in the FIR model corresponding to the delays at which vascular responses were expected to be peaking. BOLD signal changes follow electrophysiological events associated with elementary sensory stimuli and simple motor functions by as little as 2 seconds, with an established response by 4-6 seconds (Bandettini et al., 1993; Turner, 1997). Thus, the hemodynamic response was collapsed across 4 post stimulus delay intervals from 4 to 12 sec, long compared to the reading times of Experiment 1, but a typical analysis time window for fMRI.

Group statistical activation maps were constructed for contrasts of interest using a t statistic. To correct for multiple comparisons, a clustering program was run on these maps to extract clusters whose members each exceeded a voxel-threshold of $p = 0.012589$ and whose area was 200 mm^2 or more (see Appendix I). The significance of the resulting clusters was computed empirically using a Monte Carlo simulation in the following way. In this simulation, a volume of Gaussian distributed numbers was generated for each participant. Most of the processing steps applied to the real data were then applied to these simulated data sets, including volumetric smoothing, resampling onto the sphere, smoothing on the spherical surface, random effects analysis, and significance map generation. The result was then subjected to the same clustering as applied to the real data, and the number of clusters reported. This process was repeated 3500 times to compute the likelihood of getting one or more clusters under the null hypothesis. We used this likelihood as the cluster-level significance. The resulting statistical activation maps using a cluster-level threshold of $p < 0.05$ are shown in Figure 4-3. The accompanying Talairach coordinates (Table 4-1) correspond to the vertex

with the minimum local p-value in the cluster. The functional activations are displayed on a map of the average folding patterns of the cortical surface, derived using the surface-based morphing procedure (Fischl et al., 1999a, 1999b).

Region of Interest (ROI) Analysis. Regions of interest (ROI) were anatomically defined on the average cortical surface in accordance with the MGH Center for Morphometric Analysis (CMA) parcellation system (Caviness et al., 1996; Rademacher et al., 1992). (See Appendix H for an overview of anatomically defined regions). The defined ROIs were then transformed to the cortical surface of each individual participant in this experiment using the spherical transformations described previously. ROIs in both hemispheres were selected a priori based on the results of the experiments described in the introduction: (1) the inferior frontal cortex (Broca's area); (2) inferior parietal cortex; and (3) superior temporal cortex (Wernicke's area). The inferior frontal ROI includes the inferior frontal gyrus pars opercularis and pars triangularis as well as the inferior frontal sulcus. The inferior parietal ROI includes the supramarginal gyrus, the angular gyrus, and the intra-parietal sulcus. The superior temporal ROI includes the superior temporal gyrus and sulcus. In addition to these anatomical masks for the ROI analysis, an additional second mask described by the active vertices (vertex-level p-value < 0.01) of the functional analysis for all target items vs. fixation was used in order to focus on the BOLD signal of areas specifically relevant to sentence processing.

The mean % signal change relative to a pre-stimulus baseline of the first three volume acquisitions (-4 to 2 sec relative to sentence onset) was calculated for each condition and across all the voxels of the ROI anatomical plus functional mask. The mean values by participants were then plotted for each 2 sec TR interval of a trial, giving the hemodynamic response function. For each ROI, both for plausible sentences only and then for implausible sentences only, one-tailed, pair-wise t-tests were performed on the mean % signal change of the two SO and OS structure types for the time interval between 4 and 12 second post stimulus-onset.

Figure 4-1

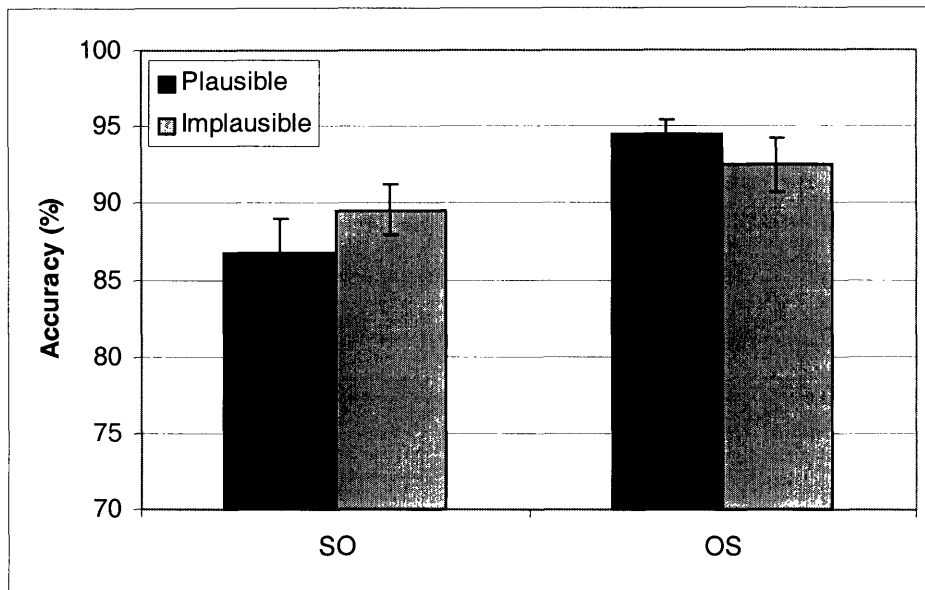


Figure 4-1. Bar graph of mean (standard error) response accuracy for each condition to the comprehension question following every item in Experiment 2, by participants.

4.2.2 Results

Behavioral Results:

Plausibility Judgment Accuracy. Mean percent correct and the standard error for the plausibility judgment are shown in Figure 4-1.

In the 2x2 ANOVA (SO/OS Structure x plausible / implausible), there was a significant main effect of structure type both by participants and by items ($F(1,11) = 30.994$, $MS_{\text{within}} = 0.0011$, $p < 0.001$; $F(1,284) = 9.799$, $MS_{\text{within}} = 0.0204$, $p < 0.01$) in which the less complex OS structure had a higher average percent correct score than the more complex SO structure. No significant effect of plausibility was noted and although the interaction showed a numerical trend in the analysis by participants, it was non-significant ($F(1,11) = 4.154$, $MS_{\text{within}} = 0.0016$, $p = 0.066$; $F(1,284) =$

Figure 4-2

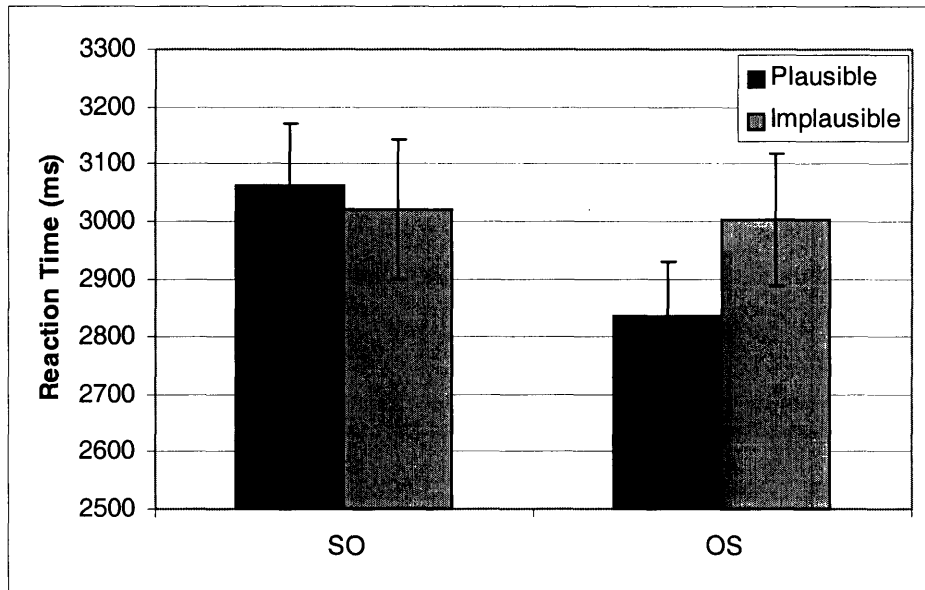


Figure 4-2. Bar graph of mean (standard error) reaction times for each condition to the comprehension question following every item in Experiment 2, by participants.

1.989, $MS_{\text{within}} = 0.0204$, $p = 0.160$). In planned pair-wise comparisons, there was an effect of sentence structure type for plausible sentences alone, both by participants and by items ($F(1,11) = 15.908$, $MS_{\text{within}} = 0.002$, $p < 0.01$; $F(1,142) = 11.085$, $MS_{\text{within}} = 0.019$, $p < 0.01$) in which accuracy was higher for the less complex OS structures than SO structures. For implausible sentences a similar pattern was seen though the effect was significant only by participants ($F(1,11) = 10.066$, $MS_{\text{within}} = 0.0005$, $p < 0.01$; $F(1,142) = 1.382$, $MS_{\text{within}} = 0.0218$, $p = 0.242$).

Plausibility Judgment Reaction Times. Mean reaction times and the standard error for the plausibility judgment are shown in Figure 4-2. Only the reaction times of correctly judged items are included in the analysis and no other trimming was employed.

In the 2x2 ANOVA (SO/OS Structure x plausible / implausible), there was once again a significant main effect of structure type both by participants and by items ($F(1,11) = 10.421$, MS_{within}

= 17323.674, $p < 0.01$; $F(1,284) = 8.831$, $MS_{\text{within}} = 210714.062$, $p < 0.01$) in which the less complex OS structure had faster average reaction times than the more complex SO structure. No significant effect of plausibility was noted. However, the interaction was significant by participants and by items ($F(1,11) = 8.181$, $MS_{\text{within}} = 15957.112$, $p < 0.05$; $F(1,284) = 4.211$, $MS_{\text{within}} = 210714.062$, $p < 0.05$). In planned pair-wise comparisons, there was an effect of sentence structure type for plausible sentences alone, both by participants and by items ($F(1,11) = 27.620$, $MS_{\text{within}} = 11189.53$, $p < 0.001$; $F(1,142) = 11.735$, $MS_{\text{within}} = 226577.01$, $p < 0.001$) in which reaction times were faster for the less complex OS structures than SO structures. For implausible sentences, the effect of structure type was non-significant both by participants and by items. It appears, therefore, that the behavioral results reflect the greater difficulty of processing SO compared to OS structures. Furthermore, this structural effect is carried primarily by the plausible sentences, giving confidence that the effect is the result of syntactic processing rather than undefined processes employed in comprehending implausible sentences.

fMRI Results:

Activation Maps. Figure 4-3 shows the statistical maps of the contrasts of interest and Table 4-1 is a listing of the Talairach coordinate locations of minimum p-values within each cluster.

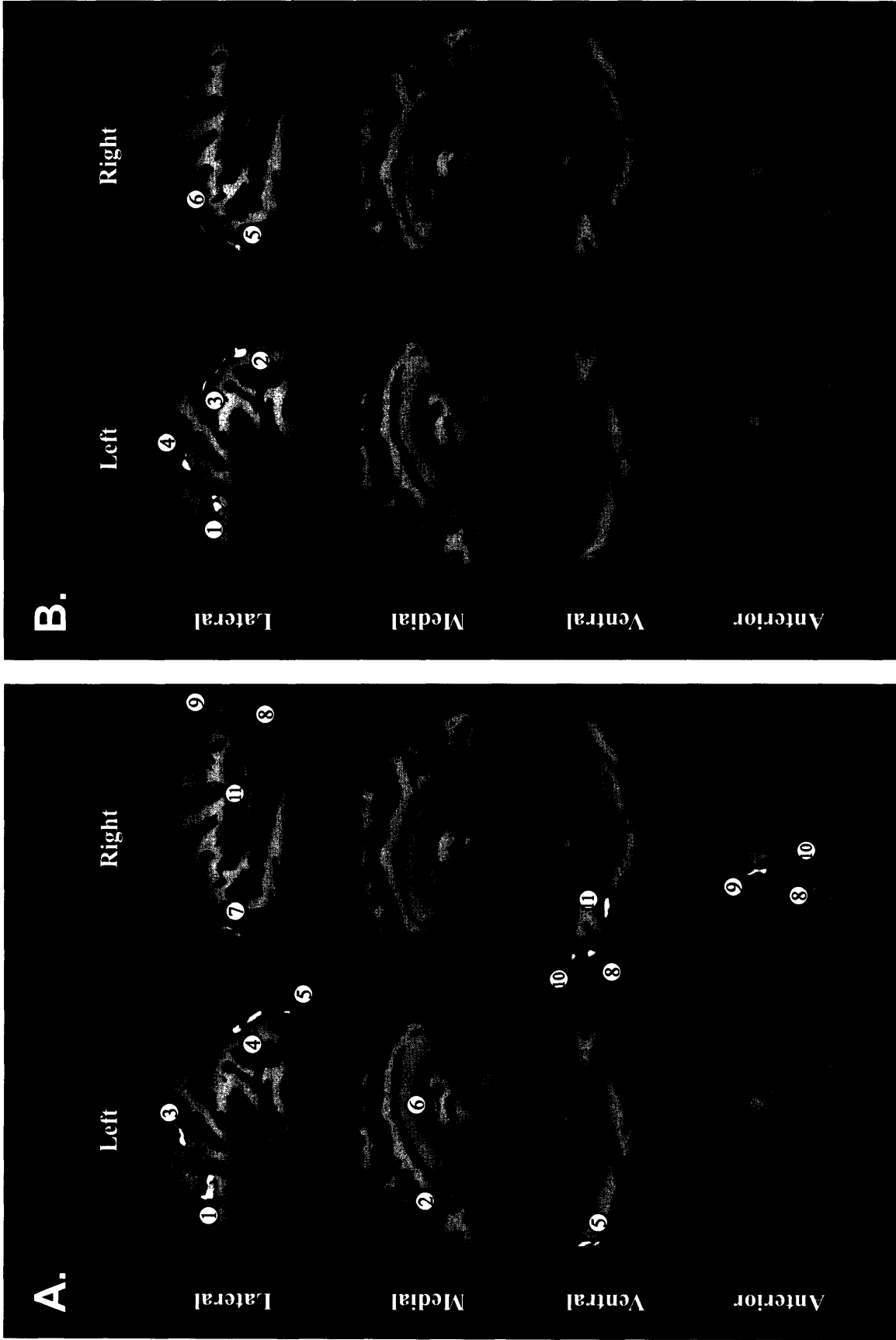
For the contrast of syntactic structure averaged across plausibility (Figure 4-3a), multi-focal areas of significant activation were seen in the left hemisphere. Of particular interest is Cluster #1 in the inferior frontal cortex, in which the vertex with the minimum p-value is located in the pars triangularis (Brodmann Area (BA) 44), and Cluster #4 straddling the inferior parietal and occipital cortex with the minimum p-value vertex located in BA 19. The latter cluster is bi-modal and when smaller clusters of 100 mm² with the appropriate, higher vertex-level p-value thresholds are examined, it separates into two distinct clusters of activation, one peaking in the occipital cortex (BA 19) and one in the angular gyrus (BA 39). Additional clusters of activation in the left hemisphere

can also be seen in the superior parietal cortex (#2; BA 7), dorsal prefrontal cortex (#3; BA 9/6), and occipital lobe (#5; BA 18). In the right hemisphere, a significant cluster of activation can be seen in the angular gyrus of the inferior parietal cortex (#7; BA 39). Significant activation is also evident in the insular region (Cluster #11) as well as several frontal cortex locations (#8-10; BA 10 & 11). Thus, there is significant activation in the language-related areas of the left inferior frontal cortex and inferior parietal cortex of both hemispheres, but none in the superior temporal cortex. There is however, also significant activation in several areas not traditionally associated with language processing. The occipital lobe activation could be due to re-reading of the more complex SO structures while dorsal frontal activation has been seen in some neuroimaging studies of syntax.

In the SO vs. OS contrast averaged across plausible sentences only (Figure 4-3b), the pattern of activation is similar, especially the activation of the left inferior frontal cortex (#1; BA 44), left dorsal prefrontal cortex (#4; BA 9), and the bilateral activation of the inferior parietal cortex (#2, 3 & 5; BA 39, 40). A cluster of activation in the right superior parietal cortex (#6; BA 7) is also seen for this comparison.

Finally, when averaged across only implausible sentences (Figure 4-3c), the SO vs. OS structural contrast revealed two significant clusters of activation in the right hemisphere. One cluster is located in the frontal cortex (#1; BA 10) and the second is located in the inferior parietal cortex (#2; BA 40).

Figure 4-3





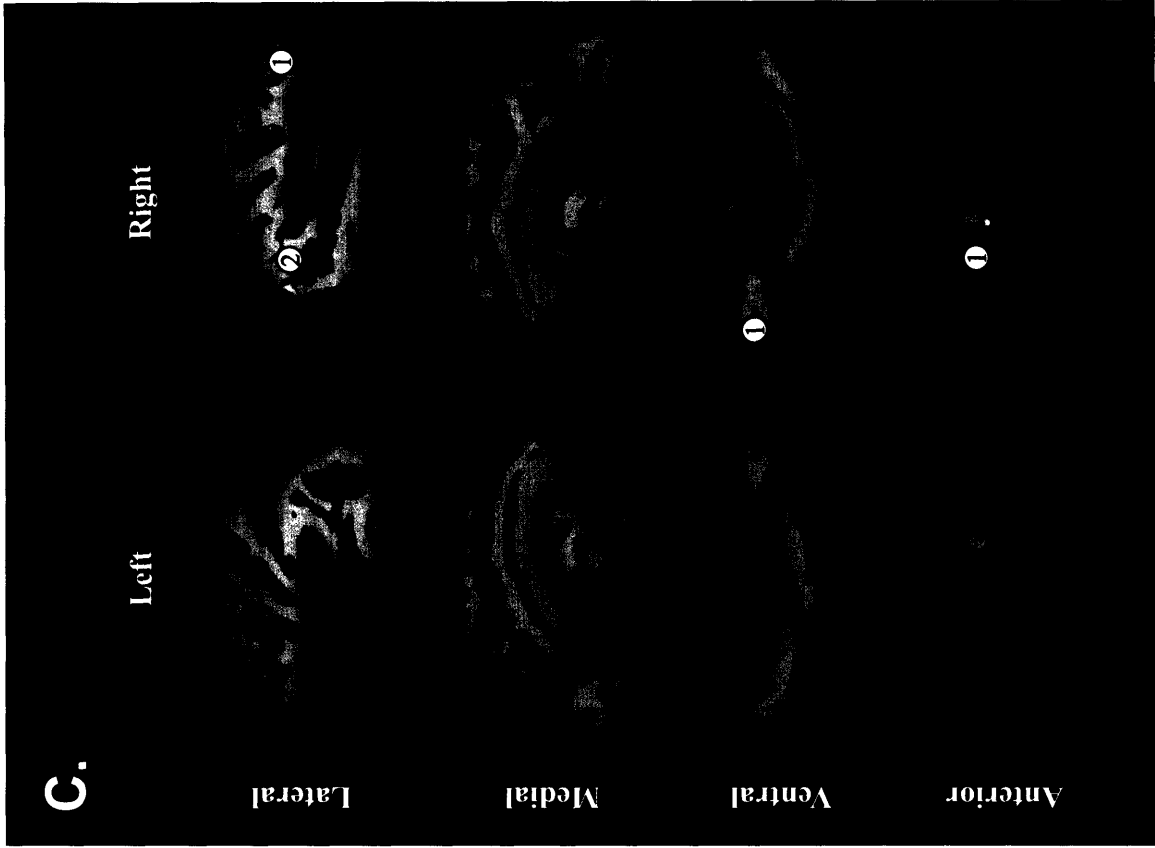


Figure 4-3. (A) Color overlays represent p-values of the SO vs. OS contrast averaged across plausible and implausible items. For this and subsequent figures, color threshold (red) corresponds to $p = 0.012589$ and ceilings (yellow) at $p = 0.001259$. Each cluster has a minimum area of 200 mm^2 and a false-positive $p < 0.05$. The number label for each cluster corresponds to the cluster # of the same contrast in Table 4-1. (B) Color overlays represent p-values of the SO vs. OS contrast averaged across plausible items only. (C) Color overlays represent p-values of the SO vs. OS contrast averaged across implausible items only.

Table 4-1

Contrast	Left Hemisphere				Right Hemisphere								
	Cluster #	Region	BA	Talairach (x,y,z)	Size (mm2)	p-value	Cluster #	Region	BA	Talairach (x,y,z)	Size (mm2)	p-value	
A. SO vs. OS	1	Inferior Frontal	44/6	(-53, 11, 15)	495	0.000050	7	Inferior Parietal	39/19	(35, -56, 36)	1027	0.000251	
	2	Superior Parietal	7	(-17, -67, 45)	267	0.005012	8	Frontal	10	(21, 47, -8)	236	0.000631	
	3	Dorsal Frontal	9/6	(-54, 3, 39)	260	0.000316	9	Frontal	10	(21, 51, 0)	451	0.000794	
	4	Occipital	19	(-38, -82, 19)	877	0.000013	10	Frontal	11	(12, 54, -18)	274	0.000794	
	5	Occipital	18	(-39, -94, -0)	852	0.000316	11	Insula		(27, 23, -1)	387	0.000032	
	6	Callosum		(-12, -9, 12)	216	0.003162							
B. SO vs. OS (Plausible Only)	1	Inferior Frontal	44	(-60, 16, 14)	285	0.000398	5	Inferior Parietal	40/39/19	(36, -61, 38)	474	0.000079	
	2	Inferior Parietal	39	(-46, -68, 30)	551	0.000003	6	Superior Parietal	7	(25, -56, 48)	401	0.000251	
	3	Inferior Parietal	40	(-41, -43, 40)	394	0.000631							
	4	Dorsal Frontal	9	(-53, 9, 40)	293	0.000050							
C. SO vs. OS (Implausible Only)				None			1	Frontal	10	(9, 58, -6)	469	0.001259	
							2	Inferior Parietal	40	(50, -48, 41)	254	0.001585	

Table 4-1. Talairach coordinates of peak activation corresponding to the local minimum p-values for each cluster of activated vertices. Cluster # for each contrast (SO vs. OS, SO vs. OS plausible only, and SO vs. OS implausible only) corresponds directly to cluster labels on the corresponding contrasts in Figure 4-3.



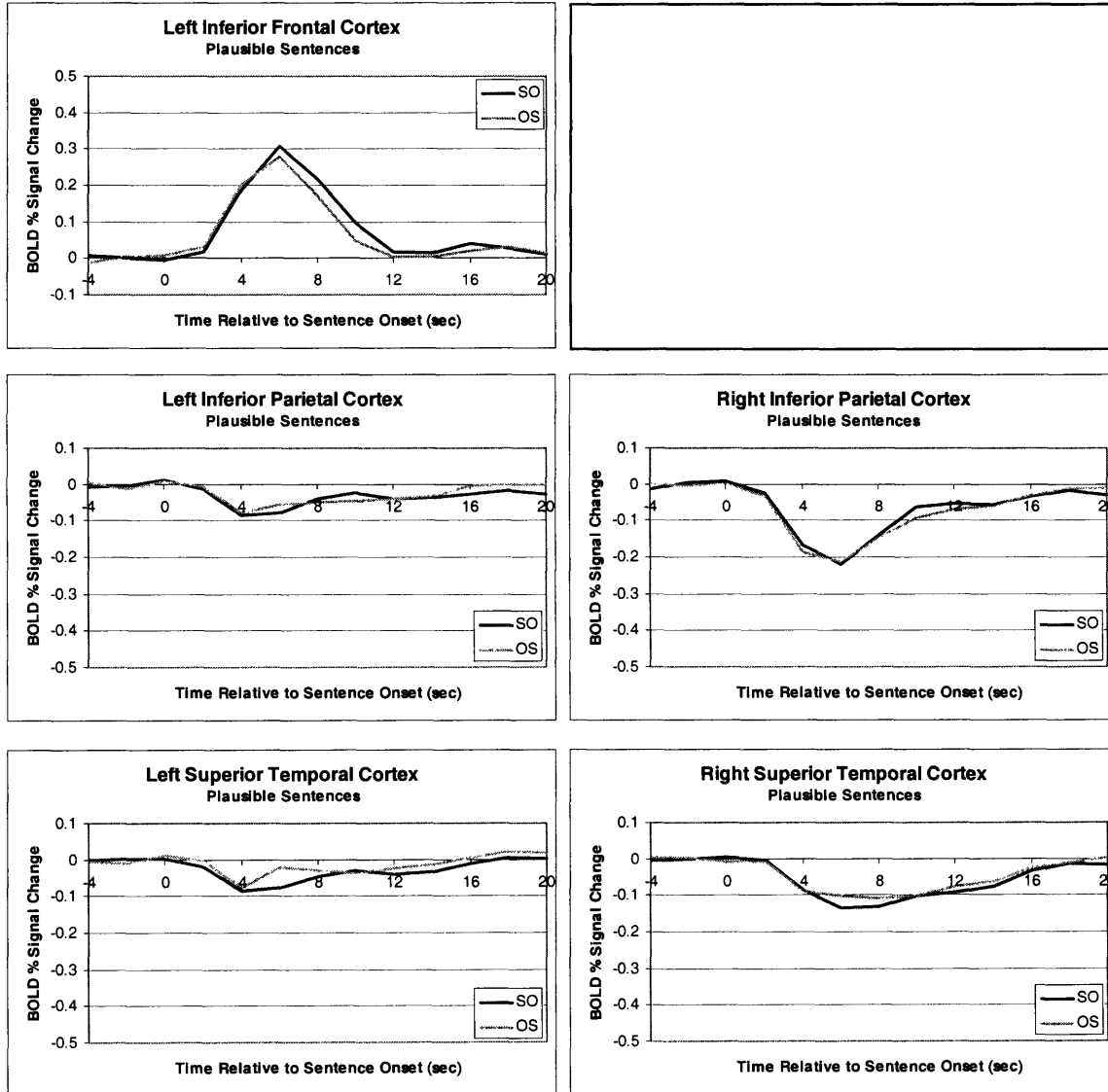
ROI Analysis. Figure 4-4 shows the hemodynamic response function of the inferior frontal, inferior parietal, and superior temporal ROIs in the left and right hemispheres for plausible sentences only (Figure 4-4a) and implausible sentences only (Figure 4-4b).

Overall, the left inferior frontal region exhibits the classic hemodynamic response function peaking at approximately 6 sec post sentence onset. On the other hand, the right inferior frontal region showed no significant functional activation and thus no hemodynamic response function could be generated in the ROI analysis. For the posterior regions, both the inferior parietal cortex and the superior temporal cortex, a negative-going response function was evident as was a largely greater magnitude in the right hemisphere.

For only plausible sentences in Figure 4-4a, the pair-wise t-tests in the left inferior frontal cortex were based on data from 207 voxels on average (range: 178-226) in the anatomical plus functional masks. Significant effects of structure type ($p < 0.003$) were seen for the last three TR Intervals (6-12 sec), in which SO structures had significantly greater % signal change than OS structures. In the right inferior frontal cortex, there were no overlapping voxels between the anatomical and functional masks. In the left inferior parietal region, pair-wise t-tests were based on data from 210 voxels on average (range: 191-235). Significant structure type effects were seen only for the final TR Interval (10-12 sec; $p < 0.007$). In the right inferior parietal region, the t-tests were based on data from 314 voxels on average (range: 253-349). Significant effects were seen for structure type in the first TR Interval (4-6 sec; $p < 0.03$) and the last TR Interval (10-12 sec; $p < 0.008$). For the left superior temporal cortex analysis, the t-tests were based on data from 86 voxels (range: 77-101) on average. Significant main effects of structure type were seen only for the TR Interval from 6 to 8 sec ($p < 0.01$). For the right superior temporal cortex, the anatomical and functional mask overlapped for an average of 257 voxels (range: 220-313). No significant effects of structure type were seen.

Figure 4-4

A.



B.

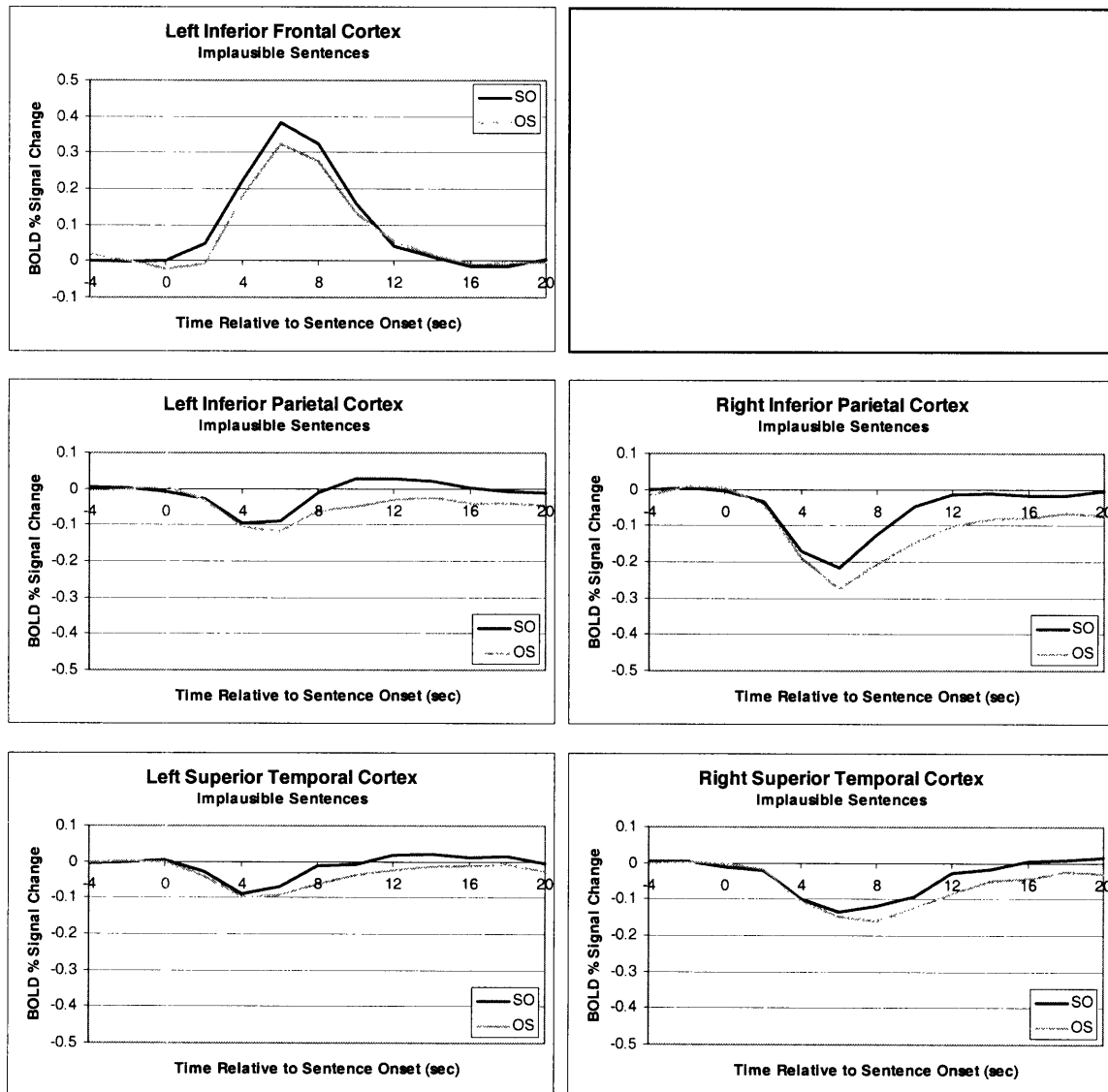


Figure 4-4. (A) Hemodynamic response functions for SO and OS structures, averaged across plausible sentences only, by participants. (B) Hemodynamic response functions for SO and OS structures, averaged across implausible sentences only, by participants.

Table 4-2

		TR Interval of significant SO minus OS	
		Left Hemisphere	Right Hemisphere
Frontal	Plausible	6-12 sec	-
	Implausible	6-8 sec	-
Parietal	Plausible	10-12 sec	4-6 & 10-12 sec
	Implausible	10-12 sec	4-6 & 10-12 sec
Temporal	Plausible	6-8 sec	-
	Implausible	-	-

Table 4-2. Summary of significant ($p < 0.05$) pair-wise differences for SO vs. OS structures at each TR interval (4-12sec) for the ROIs examined in both hemispheres (see Figure 4-4).

For only implausible sentences in Figure 4-4b, the pair-wise t-tests in the left inferior frontal cortex revealed significant effects for structure type in only the TR Interval from 6 to 8 sec ($p < 0.03$). For the left inferior parietal region, the t-tests revealed significant structure type effects only for the last TR Interval (10-12 sec; $p < 0.03$). In the right inferior parietal region, the t-tests revealed significant effects of structure type for the first (4-6 sec; $p < 0.05$) and last (10-12 sec; $p < 0.04$) TR Intervals. In the left superior temporal cortex, the t-tests did not reveal any significant structure type effects, nor did t-tests in the right superior temporal cortex.

Table 4-2 summarizes the significant pair-wise differences of BOLD signal changes for SO vs. OS structures at each TR interval for the ROIs in both hemispheres. For plausible sentences only, SO structures had a significantly greater percent signal change than OS structures ($p < 0.05$) in the left inferior frontal cortex from 6 to 12 sec post stimulus onset. Inferior parietal cortex also had moments of significant effects in both hemispheres, mostly from 10 to 12 sec post onset. For implausible sentences only, SO structures had significantly greater % signal change than OS structures in an almost identical pattern as for plausible sentences only, except for a briefer period of

significant activity in the left inferior frontal cortex. Overall, the superior temporal cortex had little significant effect of structure type.

4.2.3 Discussion

The behavioral results revealed that, as expected, the syntactically more complex SO structures were more difficult to process than the less complex OS structures. SO structures had lower average percent correct scores and slower reaction times than OS structures. The imaging results for the SO vs. OS contrast show a pattern of activation that is comparable to that seen in the Just et al. (1996) and Ben-Shachar et al. (in press) studies. When BOLD signal responses were averaged across plausibility, there was significant activation in Broca's area and in the inferior parietal cortex bilaterally. A similar pattern of activation was also seen for the SO vs. OS comparison of plausible sentences only. Because the activation patterns are very similar, it reinforces the conclusion that the activations found in previous PET studies with this paradigm reflect differences in syntactic processing and are not the result of undefined processes that occur as the sentence comprehension mechanism attempts to deal with implausible sentences. Averaged across implausible sentences only, the SO vs. OS contrast revealed no significant activation. The lack of activation for the structural contrast could be because the processes used to deal with implausible sentences overshadow any subtle differences in syntax. However, the exact processes involved in processing implausible sentences and their relationship with syntactic processing are not clear and will of course require further study.

Overall, the activation maps appear to indicate that the neural substrates of syntactic processing for SO vs. OS sentences simultaneously involve Broca's area and the inferior parietal cortex, bilaterally. The hemodynamic time course data, on the other hand, do not support the conclusion that these regions all show identically greater activation for SO structures as compared to OS structures. Effects of structure can be seen in Broca's area just after the peak of the

hemodynamic response. SO structures had a significantly greater % signal change than OS structure from 6 to 12 seconds after the onset of the stimulus sentences. For pair-wise comparisons in the inferior parietal cortex, SO structures had greater % BOLD signal change than OS structures over only some TR intervals: 10-12 sec in the left hemisphere but 4-6 and 10-12 sec in the right. It would therefore appear that the effect of structure type in Broca's area is robust and remains in both the activation map analysis and the ROI hemodynamic response analysis. Bilateral inferior parietal effects of structure type are less robust and, though apparent in the activation map analysis, do not have hemodynamic response functions that behave in an identical fashion as activation in Broca's area. In fact, the hemodynamic response functions in bilateral inferior parietal regions do not consistently show significant effects of structure at all.

Another difference between the hemodynamic response functions of Broca's area and the inferior parietal region is the direction of the peak BOLD signal. Activation in the left inferior frontal ROI showed a positive hemodynamic response function relative to the pre-sentence baseline, peaking at 6 seconds after the onset of the stimulus sentence, while in the bilateral inferior parietal ROI the hemodynamic response was in the negative direction. What fundamental differences in processing are represented by the negative- and positive-going response functions is not clear, other than that the processing of SO and OS syntactic structures depress the BOLD signal in the inferior parietal region as compared to the processes occurring in that region while the participant fixates on a "+", indicating a relative decrease in processing load. Meanwhile, the increase of the BOLD signal in Broca's area could indicate recruitment of the area in the processing of SO and OS syntactic structures.

The results add to the documentation of vascular responses occurring in both anterior and posterior perisylvian cortex in association with one aspect of syntactic processing -- an increase in processing load corresponding to object relativization that can be associated with operations related to the co-indexation of traces. While both the deficit-lesion correlational literature and the functional

neuroimaging literature focus to a large extent on a possible special role for Broca's area in this operation, perhaps by supporting memory requirements, these data as well as other results in the literature provide evidence that posterior perisylvian cortex is also somehow involved in these operations though possibly in a different and less robust way (see Fiebach et al., 2001a, 2001b, for related results). This conclusion raises the question of whether these anatomical regions play the same role as a connected network of cortex contributing to the co-indexation of traces or, if a more modular view is taken, what differentiates their contributions to the processing of SO and OS syntactic structures.

Most researchers have argued either that the roles of the Broca's area and posterior perisylvian cortex in processing these constructions differ (Grodzinsky, 2000; Ben-Shachar et al., in press; Kaan & Swaab, 2002; Caplan et al., 2002), or that these regions accomplish similar functions in different groups of participants (Waters et al., 2003; Caplan et al., 2003). The results of this study do not fully support the former view. The hypothesis that Broca's area and posterior perisylvian cortex areas differ in their role in syntactic processing would receive support from finding a complete difference in the vascular response associated with processing object-relativized sentences. This was not the pattern of results found here as Broca's area and bilateral inferior parietal regions all showed significant effects of structures type in the activation map analysis. .

Examining in more detail the specific hypotheses that have been proposed about the roles of Broca's area and the posterior perisylvian cortex in syntactic processing also leads us to view the results of this study as failing to fully support any of the articulated models. Kaan & Swaab (2002) suggested that Wernicke's area is involved in activating syntactic information associated with lexical items. The SO and OS sentences in this study contained identical lexical items, and therefore should not have lead to BOLD differences in posterior perisylvian cortex, according to this hypothesis. A similar problem arises for the suggestion of Ben-Shachar et al. (in press) that differences in the number or content of propositions might account for posterior perisylvian activation in some

sentence comprehension studies, since the SO and OS had the same propositional content. The present data are consistent with the proposal that Broca's area activation results from intrinsic memory demands of syntactic processing (Stromswold et al., 1996, Caplan et al., 1998, 1999, 2000, 2002) and with the closely related view that Broca's area is responsible for storing information that has not yet been integrated into an accruing syntactic and propositional representation when processing load increases (Kaan & Swaab, 2002). However, there is nothing in the present results that supports the view that the similar processes do not underlie BOLD signal responses in posterior perisylvian cortex. Specifically, the proposal of Caplan et al. (2002) that posterior perisylvian activation reflected task-related short-term memory demands would receive support by a significant delay in the time course of the BOLD signal activity in the posterior perisylvian cortex, since the occurrence of that response is hypothesized to be related to re-analysis of complex sentences rather than their initial parsing. However, no such delay in the time course of BOLD signal responses were found here, perhaps because of insufficient temporal resolution in the hemodynamic response function. Finally, the proposal that the posterior perisylvian cortex may support parsing in less efficient language processors (Waters et al., 2003; Caplan et al., 2003) cannot be tested in these data because the reaction times and accuracy of the participants in this study were comparable to those of "fast responders" in previous studies, who showed Broca's area activation.

While the data reported here fail to support any of the models that posit different roles for Broca's area and the posterior perisylvian cortex in this one aspect of syntactic processing, it is premature to conclude that they provide strong evidence for the view that these different regions are all responsible for the same operations as a connected network of cortical areas. A possible role for the posterior perisylvian cortex in processing lexical syntactic information or sentential propositional representations was not examined in this study because these factors were purposefully controlled, not varied, across experimental conditions. The temporal resolution of the BOLD signal measure as well as the variable and extended relation of BOLD signal response to physiological (and

psychological) events may well have made it impossible to test Caplan et al.'s (2002) theory in this experiment. Paradigms such as question-answering (Just et al., 1996), which require a longer period of maintaining material in short term memory than does end-of-sentence plausibility judgment, are better suited to exploring possible differences in the time course of BOLD signal activity that might be associated with initial parsing (putatively carried out in Broca's area) and maintaining a representation of a sentence in short-term memory (putatively carried out in the posterior perisylvian cortex). Alternatively, interfering with short term memory through concurrent tasks may affect Broca's area and posterior perisylvian response patterns differently.

In summary, using the same stimuli as Stromswold et al. (1996), we found activation in both Broca's area and bilateral inferior parietal cortices, comparable to the findings by Just et al. (1996) and Ben-Shachar et al. (in press). The magnitude and time course of the hemodynamic response and the different responses to plausible and implausible sentences were not identical between Broca's area and the posterior perisylvian cortex. These results suggest that both regions of activation are due to similar but not identical aspects of processing the more complex sentences used in this study and in keeping with the aphasia data. However, a better understanding of the differences in the functions accomplished by these different brain regions may yet be uncovered through the use of technologies with better temporal resolution than event related fMRI and of different tasks that place greater loads on short term memory.

4.3 Experiment 3: Wh-trace Predictions

Having established the DLT theory of storage costs for wh-trace predictions behaviorally in Experiment 1b (Chapter 3), we developed further stimulus items for a comparable neuroimaging experiment. First, because many items were needed for each condition and participants could only remain in the scanner for a limited amount of time, only two conditions, the ambiguous sentential complement (SC) and the unambiguous relative clause (RC) illustrated in (31), were chosen for the neuroimaging experiment.

(31)

a. SC, Ambiguous

The announcement [that **the baker from a small bakery in New York City** received the award] helped the business of the owner.

b. RC, Unambiguous

The announcement [which **the baker from a small bakery in New York City** received ____] helped the business of the owner.

The two conditions were chosen because they had the fewest differences in the initial three words prior to the critical clause “the baker from a small bakery in New York City”. Thus, any differences in the measured Blood Oxygenation Level Dependant (BOLD) signal are unlikely to be the result of a trickle-down-effect from prior words. Furthermore, the unambiguous SC condition was not utilized because it was decided that the disambiguating verb (e.g. “alleging” in sentence (16b) on pg. 46) created a less than natural sounding sentence. In order to encourage participants to pursue the SC reading of the condition, the matrix subject nouns were chosen to be SC-biased with greater than 80% SC completions based on norming data from Kennison (2000) (see Appendix E). Finally, two prepositional phrases (PP) were used to lengthen the region rather than the more complex wh-

clauses, simplifying the critical clause. Otherwise, the predictions are the same: (1) the unambiguous RC condition should have greater activation than the ambiguous SC condition, based on the behavioral data of Experiment 1b; and (2) localization of the activity should be focused in the left inferior frontal cortex and possibly the inferior parietal cortex bilaterally, based on the neuroimaging data of Experiment 2. First, the items were tested behaviorally before committing to the substantially more expensive fMRI technique.

4.3.1 Behavioral Methods

Participants. Twenty-nine participants from the MIT community who did not take part in the previous experiments were paid for their involvement. All were native speakers of English and were naïve as to the purposes of the study.

Materials & Design. Sixty sets of sentences were constructed with 2 conditions each, following the form of (31). See Appendix E for a full list of the items.

The target sentences were split into two lists of 60 items, each balancing all factors in a Latin-Square design. Each list was combined with 80 fillers of various types, including 60 sentences from an expanded version of Experiment 1c with similar neuroimaging goals. The stimuli were pseudo-randomized independently for each participant such that at least one filler item separated any two targets.

Procedure. The procedure was the same self-paced word-by-word paradigm as in Experiments 1. The comprehension task that followed each item consisted of simple yes / no questions.

Figure 4-5

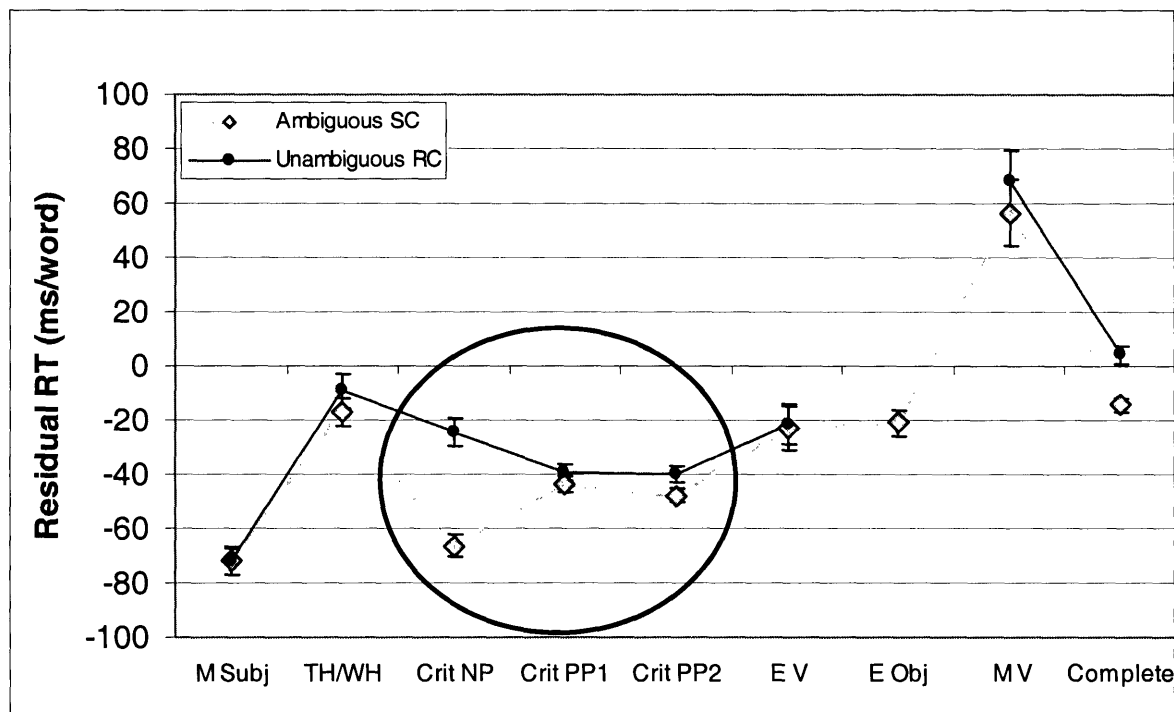


Figure 4-5. Plot of mean (standard error) residual RTs per word for each region of Experiment 3 as a function of condition, by participants (circled data indicates the critical region).

4.3.2 Behavioral Results

Three participants' data were omitted from all analyses: one because of poor comprehension question performance (< 67% accuracy overall), and the other two because of unclear native English status.

Comprehension Question Performance. Overall, the comprehension questions for the experimental items in Experiment 3 were answered correctly in 85.4% of the trials. For the ambiguous SC condition the mean was 85.26% (standard error = 1.72%) while for the unambiguous

Table 4-3

	Ambiguous SC	Unambiguous RC
Matrix Subj	-72 (357)	-72 (356)
That/Which	-17 (386)	-9 (409)
Critical NP	-66 (347)	-24 (392)
Critical PP1	-44 (375)	-39 (375)
Critical PP2	-48 (361)	-40 (375)
Embedded Verb	-23 (415)	-21 (417)
Embedded Obj	-21 (399)	
Matrix Verb	57 (502)	68 (514)
Completion	-14 (398)	4 (421)

Table 4-3. Mean residual RTs (msec / word) for each region of Experiment 3 as a function of condition, by participants (raw RTs in parentheses).

RC condition the mean was 85.52% (standard error = 1.84%). A paired t-test revealed no main effect of sentence structure (all $F_s < 1$).

Reading Times in the Critical Region. Only items with correctly answered comprehension questions were analyzed. Reading time (RT) data points that were greater than 3 standard deviations from the mean were excluded from the analysis, affecting less than 1.7% of the data points in Experiment 3 overall. Figure 4-5 presents the mean residual RTs (msec/word) across the two conditions in this experiment. The critical material in the storage cost analysis consisted of the embedded subject noun phrase and prepositional phrases, “the baker from a small bakery in New York City”. Table 4-3 presents the mean RTs per word for all regions, in both residual and raw forms.

Averaged across the critical region, residual RTs were slower for the unambiguous RC condition (mean = -38; standard error = 8) than for the ambiguous SC condition (mean = -51;

standard error = 9), significant by participants but not by items ($t_1(25) = 6.458$, $MS_{\text{within}} = 334.6$, $p < 0.02$; $t_2(59) = 2.671$, $MS_{\text{within}} = 1076$, $p = 0.108$).

4.3.3 Discussion

Unfortunately, despite the significant effect by participants of wh-trace storage costs across the critical region, as predicted by the DLT, it was decided not to pursue the associated neuroimaging experiment. The reasons for this decision include the fact that the effect was significant only by participants and not by items. Furthermore, from the plot of the RTs in Figure 7-1, it appears that most of the effect is sustained by the embedded subject NP while the additional PPs have virtually identical RTs for the two conditions. The ramifications of this effect are not clear and could represent some ceiling effect in which PPs are extremely easy to process by the sentence comprehension mechanism, or could be from the lack of large enough numbers of participants to detect any significant effects. In the end, it was not apparent that the item set of Experiment 3 would generate enough of a difference in BOLD signal between the two conditions to justify a relatively expensive neuroimaging experiment.

Chapter 5

Neuroimaging Study on Ambiguity Resolution

5.1 Introduction

Sentences that are temporarily ambiguous between two different syntactic structures have been used extensively in psycholinguistic research to examine the human sentence processing mechanism. An important example is the main verb / reduced relative (MV/RR) ambiguity, a robust behavioral contrast in which the sentence processing mechanism often has greater difficulty with the RR structure and prefers the MV structure interpretation (Bever, 1970). Studies of this contrast have been particularly fruitful in understanding the influence of verb subcategorization information and plausibility on the syntactic parser. A neuroimaging examination of the resolution of this ambiguity will complement our examination of storage costs in our effort to understand the sentence comprehension mechanism. Gibson & Pearlmutter (1998) believe that the reason that such factors as subcategorization and plausibility play a prominent role in ambiguity resolution is because the MV and RR alternatives have such similar storage and integration costs.

Example (32) below illustrates the two possible syntactic structures for the ambiguous initial words “The defendant examined”. After the verb “examined”, the MV structure of sentence (32a) is

disambiguated by the determiner “the” of the direct object noun phrase (NP) “the documents” while the RR structure of sentence (32b) is disambiguated by the preposition “by” of the prepositional phrase (PP) “by the lawyer”. By disambiguated, we mean that the processing mechanism now has enough information to construct the proper syntactic structure.

(32)

a. Main Verb (MV)

The defendant examined the documents.

b. Reduced Relative (RR)

The defendant examined by the lawyer was unreliable.

In reading time experiments involving the MV/RR ambiguity, participants are presented visually with the less preferred RR structure (repeated as sentence (33a)), which is initially ambiguous with the MV structure. The baseline condition is the unambiguous relative clause (RC) in which the RR structure is signaled overtly and before the verb by the words “that was” (sentence (33b)). The ambiguous RR structure has the exact meaning of the RC structure except that it is “reduced” because of the option in English to omit instances of “that was”. The critical region to be examined is the PP “by the lawyer”, and reading times for this region in the unambiguous RC condition are subtracted from the same region of the ambiguous RR condition. The difference in reading times is called the ambiguity effect and is thought to reflect the increased processing engaged in resolving the ambiguity. A positive ambiguity effect has been found in the literature (Ferreira & Clifton, 1986; Trueswell et al., 1994) and is thought to reflect the greater difficulty that the sentence processing mechanism has with the ambiguous RR structure over the unambiguous RC structure because of the confusion associated with disambiguating to the RR structure when the MV structure is preferred and initially expected.

(33)

a. Ambiguous RR

The defendant examined **by the lawyer** was unreliable.

b. Unambiguous Relative Clause (RC)

The defendant *that was* examined **by the lawyer** was unreliable.

There are at least two fundamentally different approaches to explain the observed difference in preference for MV over RR structures and the process of syntactic ambiguity resolution. One approach is centered on rules-based syntactic processing. An example of this approach is the Garden Path model in which the sentence processing mechanism is based on two modules (Ferriera & Clifton, 1986; Frazier & Fodor, 1978). As described previously, the first module is the syntactic processor, or parser, which is privy only to syntactic information such as the category of the words (e.g. noun or verb) and utilizes syntactic rules such as Minimal Attachment and Late Closure. The principle of Minimal Attachment, relevant to our discussion of the MV/RR ambiguity, states that new material is attached to the existing syntactic structure using the least number of nodes possible and that a parse with fewer nodes is preferred over one that has more nodes. The second module is the thematic processor, which detects errors in semantics (e.g. the plausibility of the propositions) described by the syntactic structure produced by the first module and attempts to correct errors as well as integrate the sentence into the rest of the discourse. With this framework, the Garden Path model processes a sentence by considering a single construction at a time, first parsing it using only syntactic information and rules, and then detecting any errors and correcting them. In the model, the MV structure in (32a) is preferred because it is syntactically less complex as it needs fewer new nodes to attach a direct object NP to the verb in the existing syntactic structure and thus satisfies the principle of Minimal Attachment, and not as a result of differences in storage and integration costs

(Gibson & Pearlmutter, 1998). Meanwhile, the RR structure is more complex, requiring more intervening nodes to create the relative clause structure when attaching the PP “by the lawyer”, and is therefore less preferred. When presented with an ambiguous RR sentence as in (33a), the parser pursues the MV structure initially and when the disambiguating PP material is presented that signals the RR structure instead, the parser must backtrack to repair or re-parse the sentence to resolve the ambiguity, resulting in longer reading times and an ambiguity effect.

The second approach to explain the process of syntactic ambiguity resolution focuses on the fact that the MV/RR structural ambiguity can also be interpreted as an ambiguity of the verb (in our examples, “examined”). In particular, for the MV structure, the verb “examined” must ultimately be recognized to be in the past tense active form. For the RR structure on the other hand, the verb is actually in the past participle form, which in sentence (33a) overtly looks like the simple past tense. On this view, the structural ambiguity could be resolved as a lexical ambiguity and in fact involve the same representations and processing. MacDonald et al. (1994) have proposed just such a mechanism for ambiguity resolution. In their framework, the lexical representation of words such as the verb “examined” carry with them syntactic information such as their argument structures and the representation of the immediate syntactic tree. Thus, the past tense form of “examined” carries key components of the MV structure while the past participle form carries the RR structure. Since the simple past tense forms of verbs are typically more frequent (both in terms of corpus counts and sentence completion studies), the MV structure (32a) is the preferred interpretation of the MV/RR ambiguity. Additionally, resolution of the ambiguity in (33a) brings with it the other properties of lexical ambiguity resolution. That is, it is parallel, considering all of the possible forms initially before selecting the correct one based on the contextual information, and is constrained by such factors as the frequencies of occurrence and co-occurrence of specific words. The ambiguity effect on reading times comes as a result of the sentence processor initially pursuing the MV interpretation

because of the frequency bias, only to need to reactivate the RR interpretation at the presentation of the disambiguating PP material.

The lexically mediated model of syntactic ambiguity resolution is attractive because it offers a simple explanation for the apparent immediate influence of semantic information on syntactic processing. In the experiments of Ferriera & Clifton (1986), the authors took items like (33) and manipulated the semantic information of animacy in the sentence initial noun as seen in (34).

(34)

a. Ambiguous RR

The evidence examined **by the lawyer** was unreliable.

b. Unambiguous RC

The evidence *that was* examined **by the lawyer** was unreliable.

Whereas example (33) has an animate initial noun “the defendant”, example (34) has an inanimate initial noun “the evidence”. An animate initial noun such as “defendant”, because it often plays a good AGENT thematic role for the initial verb “examined” and is in the putative subject position, reinforces the preference for the MV interpretation. An inanimate initial noun such as “evidence”, on the other hand, makes a poor AGENT and reinforces the RR interpretation. The argument is that, if the Garden Path model is true and syntactic information does take precedence over semantic information like animacy at first, then there should be no difference in the ambiguity effect between the two types of animacy during the initial stages of processing. Using eye-tracking as a window into the initial parsing, the authors found that there was indeed no difference in ambiguity effect and concluded that the data supported the Garden Path model. However, in a subsequent experiment, Trueswell et al. (1994) noticed that some of the inanimate initial nouns coupled with the initial verb

made very plausible MV sentences in which the initial noun plays a INSTRUMENT thematic role rather than an AGENT role (see (35)).

(35)

a. RR Completion

The car towed **by the truck** had broken down.

b. MV Completion

The car towed the trailer.

So animacy, though highly correlated with plausibility, may not quite be the appropriate factor to manipulate. Consequently, Trueswell et al. modified the items to reflect a division in the plausibility of the initial noun-verb pair as a MV completion rather than just the animacy of the initial noun alone. Using eye-tracking also, Trueswell et al. found that with these manipulations, there was indeed a difference in the ambiguity effect such that noun-verb pairs that were plausible as MV completions had significant ambiguity effects while implausible pairs had no significant ambiguity effects. The authors concluded that the data did not support the Garden Path model but rather supported constraint-based approaches in which semantic information had an immediate influence on syntactic parsing. The lexically-mediated model proposed by MacDonald et al. offers a simple instantiation of constraint-based approaches since semantic information such as animacy and plausibility along with the local syntactic structure are all carried by the lexical representations and immediately available to each other.

These two fundamentally different approaches of syntactically-mediated and lexically-mediated ambiguity resolution could have very different neurological substrates as well. In fact, Ullman (2001) has proposed the declarative/procedural model of language in which the lexicon is a form of declarative memory and anatomically located in the temporal lobe while syntactic processing

utilizes procedural memory and is located in the frontal cortex and basal ganglia. Indeed, as we discussed in Experiment 2 (Chapter 4), significant evidence seems to indicate that Broca's area (the inferior frontal gyrus of the dominant (left) hemisphere) is crucial for many different aspects of syntactic processing but that posterior perisylvian cortex is also involved.

Several aspects of lexical processing seem to be predominantly located in the posterior portion of the dominant perisylvian cortex. One important point to consider in our discussion of lexically-mediated ambiguity resolution is which aspect of lexical processing is the local syntactic tree/argument structure associated with. It could be associated with phonological processing, as access needs to be immediate and jabberwocky verbs such as "glicked" seem intuitively to have a transitive, or possibly intransitive, argument structure despite having no real semantic meaning. One could hypothesize that the argument structure is linked, in the English past tense, to the morphological "-ed" regularized ending. However, verbs such as "hoped", even with a regular past tense ending, have a more complex argument structure that requires a sentential complement, with or without the overt "that", as in sentence (36a) and cannot take a simple transitive direct object as in sentence (36b).

(36)

a. The man hoped that the war would end.

The man hoped the war would end.

b. *The man hoped the war.

The hypothesis could then be that argument structures like the RR structure of the past participle form may be more deeply associated with the full meaning of the verb and available only through lexico-semantic processing.

In a review of the lexical processing neuroimaging literature, Binder & Price (2001) discuss many of the issues and brain regions involved in lexical processing. For phonological processing, most important regions appear to be in the anterior half of the perisylvian region and include the posterior and dorsal portions of the inferior frontal gyrus (BA 44/6) for articulatory rehearsal, the frontal operculum (BA 45) for phonemic perception, and BA 40 for verbal short-term memory demands in orthographic to phonographic translation. For lexico-semantic processing, the relevant regions seem to be primarily in the posterior half of the perisylvian region. Taking task and baseline stimuli into account, Binder & Price conclude that the angular gyrus (BA 39) is most consistently involved in lexico-semantic processing. Another pertinent brain area for lexico-semantic processing is the dorsal prefrontal cortex, which they propose to be the motivator for semantic information retrieval. The ventral and lateral left temporal lobe (BA 20/36), possibly related to storing or processing sound-based and semantic representations of words according to Binder & Price, are also areas of interest.

Thus, while it seems that many regions of importance for lexical processing are located in the posterior portion of the perisylvian region, especially for lexico-semantic processing, anterior portions of the perisylvian region still play significant roles for both lexico-semantic and phonological processing. In particular, the inferior frontal gyrus or Broca's area appears to subsume many processes: phonological processing, semantic processing, syntactic processing and working memory (Stowe et al., 1998; Dapretto & Bookheimer, 1999; Poldrack et al., 1999; Caplan et al., 2000; Muller et al., 2003; Chen et al., in preparation). Of course, a modular division of functional processes and anatomy is not the only possible view. Some researchers have proposed that lexical as well as syntactic processing levels both involve a network of frontal and posterior perisylvian regions and that both levels in fact interact and may share common resources (Keller et al., 2001).

5.2 Experiment 4: The MV/RR Ambiguity

In summary, the MV/RR ambiguity is an important contrast in the psycholinguistic literature. A neuroimaging examination of this very robust and fundamental behavioral phenomenon, to establish possible neural substrates, is essential to a greater understanding of the sentence processing mechanism and this study could provide further insight to ambiguity resolution and information use. We present items like those used by Trueswell et al. (1994), ambiguous RR and unambiguous RC sentences with initial nouns that are plausible and implausible subjects for the initial verb, and hypothesize that if resolution of the MV/RR ambiguity is syntactically mediated, there will be greater activation in the dominant (left) inferior frontal region while if ambiguity resolution is lexically mediated, there will be greater activation in the dominant posterior perisylvian region.

5.2.1 Imaging Methods

Participants. Twenty participants from the Massachusetts General Hospital community (18 female, 2 male; mean age 23.3 years, range 19-27; mean years of education 16, range 14-18) were paid for their involvement (see Appendix M). All were right-handed, native speakers of English and naïve as to the purposes of the study. Informed consent was obtained for all participants.

Materials & Design. The experimental items consisted of 32 sets of RR and RC syntactic structures with 4 conditions: MV plausible / implausible matrix subject noun phrases by ambiguous / unambiguous structures as illustrated in (37a-d) below. A fifth condition (37e) was included as a foil for all the presented RR and RC structures and consisted of the MV syntactic structure alternative of the animate matrix subject noun phrase. See Appendix F for a list of the items presented, half of which are adapted from Trueswell et al. (1994). The remaining items were chosen based on a

combination of plausibility ratings of the subject noun and verb (e.g. “the courier” and “transported”) and the ambiguity effect of the item in a reading time pre-test.

(37)

- a. MV Plausible Subject Noun, Ambiguous as RR

The courier transported by the armored car was safe.

- b. MV Plausible Subject Noun, Unambiguous as RC

The courier who was transported by the armored car was safe.

- c. MV Implausible Subject Noun, Ambiguous as RR

The treasure transported by the armored car was safe.

- d. MV Implausible Subject Noun, Unambiguous as RC

The treasure that was transported by the armored car was safe.

- e. Foil of MV Plausible Subject Noun condition

The courier transported the treasure by armored car.

Sentence (37a) is ambiguous between the RR and MV syntactic structures. Furthermore, the subject noun phrase “the courier” is highly plausible as the AGENT of the verb “transported” and thus favors the MV structure. Sentence (37b), however, is unambiguously the RC structure. Sentence (37c) is also ambiguous between the RR and MV syntactic structures, but the subject noun phrase “the treasure” is implausible as the AGENT and favors the RR structure. Once again, the unambiguous RC structure of sentence (37d) is used as the baseline. The foil sentence (37e), as mentioned above, was included as an attempt to counter participant strategies that may develop as a result of the continued presentation of RR and RC structures by maintaining the possibility that the plausible noun phrase is indeed the start of a MV structure. However, the ratio of MV to RR and RC structures is still relatively large at 4:1.

For some items, subject nouns were rated on their typicality or plausibility as AGENTS and PATIENT/THEMES for the matching verbs in a MV structure. Unfortunately, ratings were not available for all subject nouns in the items of Appendix F due to changes to the final stimuli. As such, the mean values for MV plausible subject nouns are based on the ratings of 27 items while values for MV implausible subject nouns are based on the ratings of 31 items. All available ratings for MV plausible subject nouns originated from a plausibility norming study with 17 participants previously performed in this lab. Ratings for MV implausible subject nouns for the first 16 items are also from this norming study while ratings for the last 16 items originate from Trueswell et al. (1994), which were based on a separate norming study with 107 participants. Based on a scale from 1 to 7 where 7 was “very typical”, MV plausible subject nouns had an average rating of 5.63 (standard deviation = 1.13) by items as being a typical AGENT for the verb they were paired with. The MV plausible subject nouns were also fairly typical as PATIENT/THEMES of the verbs with an average rating of 4.78 (1.44). Meanwhile, MV implausible subject nouns had an average rating of 1.83 (0.80) as being a typical AGENT and an average rating of 5.26 (1.29) as being a typical PATIENT/THEME.

Procedures:

Stimulus Presentation. Each stimulus sentence item and comprehension question was visually displayed in its entirety on a single line in the center of the screen. A given experimental trial consisted of a sentence item presented for 4 sec, followed by a brief 300 ms blank screen, then a comprehension question for 3.5 sec, and a final 200 ms inter-trial interval for a total trial length of 8 sec. The task for the participants during the experimental trial was to read the sentence and then to answer a yes / no comprehension question. The accuracy of the comprehension question answer and the reaction time for the response from the onset of the sentence trial were recorded.

The pseudo-randomization in the event-related design was according to a computer program developed to randomize trial types and vary the duration of the fixation trials for optimum efficiency in the deconvolution and estimation of the hemodynamic response (Burock et al., 1998; Dale, 1999). Thus, randomly interspersed between each 8 sec sentence trial was a 2, 4, 6, 8, 10 or 12 sec fixation trial.

The 160 items (32 sets of items x 5 conditions) interspersed with the fixation trials were divided into 8 runs. No two sentences from a single set were presented in the same run. Participants were allowed a short break between each run. The sentences were projected to the back of the scanner using a Sharp LCD projector and viewed by the participants as a reflection in a mirror attached to the head coil. Responses were recorded via a custom-designed, magnet compatible button box. A Dell Inspiron 4000 computer running the E-Prime v1.0 software package (Psychology Software Tools, Inc., Pittsburgh, PA) was used to both present the stimuli and record the comprehension question response.

MR Imaging Parameters. The imaging parameters are identical to those of Experiment 2 except for a total of 135 volume acquisitions.

Cortical Surface Reconstruction & Functional Pre-processing. The cortical surface reconstruction and pre-processing procedures are identical to those described for Experiment 2, as presented in Chapter 4.

Functional Statistical Analysis. The analysis procedures are nearly identical to those of Experiment 2 except that for each voxel on the spherical surface, the BOLD signal value was averaged across 3 functional volume acquisitions, from 6 to 12 seconds after the onset of the sentence trial. (See Appendix G for an overview of the FS-FAST analysis stream and Appendix K

for a summary of the FS-FAST commands). Also, cluster level analyses were based on simulations of 20 participants with 9492 repetitions, resulting in a vertex-level p-value threshold ($p < 0.012589$) for a 200 mm² minimum area that was identical to Experiment 2 (see Appendix I).

5.2.2 Imaging Results

Behavioral Results:

Comprehension Question Accuracy. Mean percent correct and the standard error for the comprehension question task are shown in Figure 5-1.

In the 2x2 ANOVA (MV plausible / implausible subject noun x ambiguous / unambiguous structure), there was a significant main effect of MV plausibility of the subject noun by participants

Figure 5-1

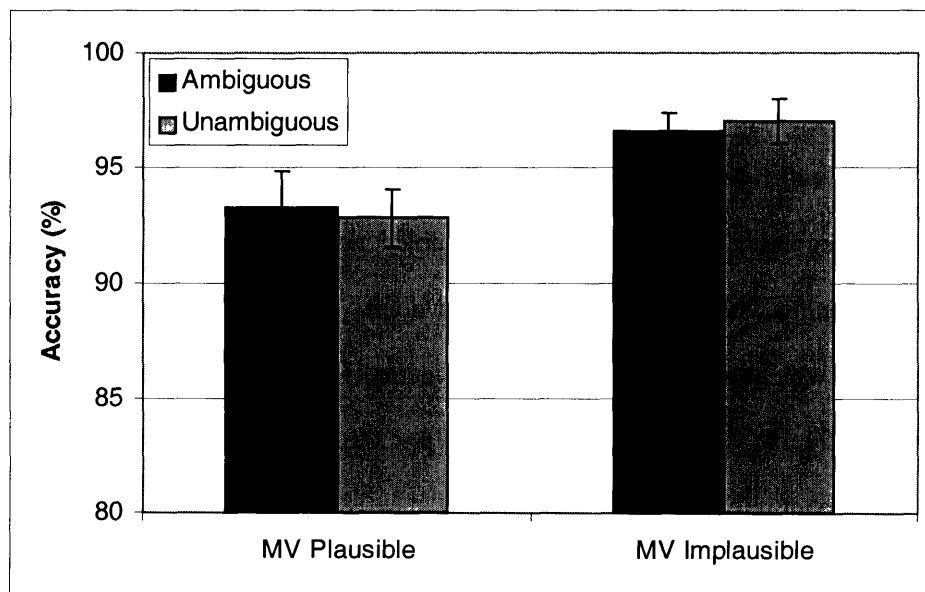


Figure 5-1: Bar graph of mean (standard error) response accuracy for each condition to the comprehension question task in the neuroimaging portion of Experiment 4, by participants. MV Plausible / Implausible refers to the plausibility of the initial noun as the subject of the initial verb in a Main Verb syntactic structure.

and by items ($F(1,19) = 13.12$, $MS_{\text{within}} = 0.0021$, $p < 0.002$; $F(1,31) = 12.00$, $MS_{\text{within}} = 0.0037$, $p < 0.002$) in which sentences beginning with a MV plausible subject noun had a lower average percent correct score than sentences beginning with an MV implausible subject noun. No significant effect of ambiguity was noted ($F_s < 1$) and the interaction was also not significant by participants or by items ($F_s < 1$). These results were unexpected, as they do not show an ambiguity effect for the MV plausible subject nouns. In other words, for an ambiguity effect we would anticipate significantly greater accuracy for the unambiguous condition of the MV plausible subject nouns as compared to the ambiguous. Instead, we see a main effect of MV plausibility indicating that in even syntactically unambiguous RC structures, the possibility of the MV structure is still considered.

Comprehension Question Reaction Times. Mean reaction times and the standard error for the comprehension question task are shown in Figure 5-2. Only the reaction times of correctly judged items are included in the analysis and no other trimming was employed.

In the 2x2 ANOVA (MV plausible / implausible subject noun x ambiguous / unambiguous structure), there was once again a significant main effect of MV plausibility both by participants and by items ($F(1,19) = 149.1$, $MS_{\text{within}} = 2392$, $p < 0.0001$; $F(1,31) = 7.678$, $MS_{\text{within}} = 78720$, $p < 0.01$) in which sentences beginning with an MV plausible subject noun had a slower average reaction time than sentences beginning with an implausible subject noun. A significant effect of ambiguity was also noted, one in which unambiguous sentences actually had slower reaction times, but by participants only ($F(1,19) = 4.437$, $MS_{\text{within}} = 6060$, $p < 0.05$; $F(1,31) < 1$). The interaction was also significant by participant only ($F(1,19) = 6.174$, $MS_{\text{within}} = 1870$, $p < 0.03$; $F(1,31) < 1$). Once again we see only a main effect of MV plausibility, and not an ambiguity effect in the MV plausible conditions in which we would have expected to see significantly faster reaction times for the unambiguous condition as compared to the ambiguous condition. These results mirrored the accuracy results described above.

Figure 5-2

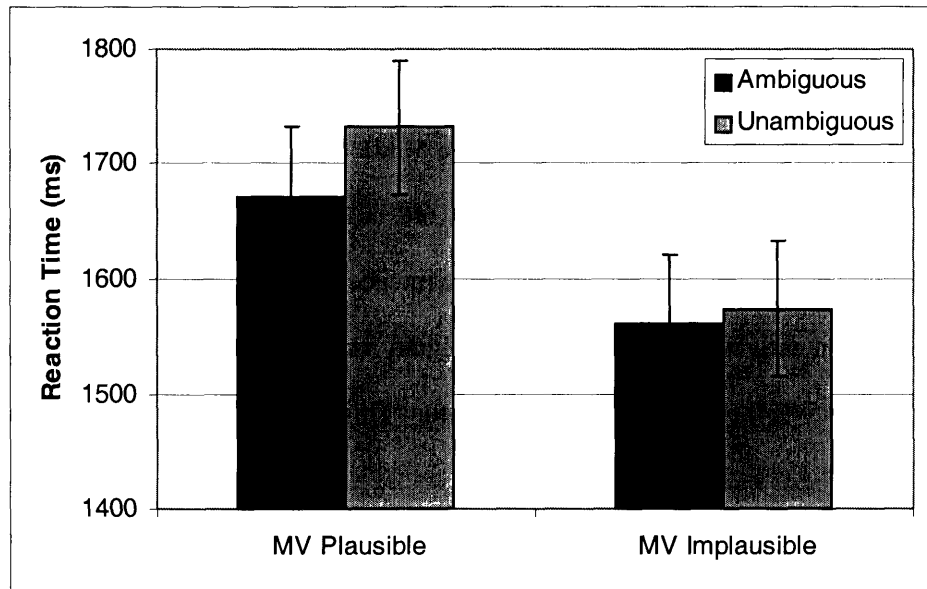


Figure 5-2: Bar graph of mean (standard error) reaction times for each condition to the comprehension question task in the neuroimaging portion of Experiment 4, by participants.

fMRI Results:

Activation Maps. Figure 5-3 shows the statistical activation maps of the various contrasts. Table 5-1 is a listing of the Talairach coordinate locations of peak activation and the corresponding p-values.

For the contrast of subject noun MV plausibility averaged across ambiguity (Figure 5-3a), multi-focal areas of significant activation in the left hemisphere were seen broadly across the inferior frontal region, peaking in the pars orbitalis (Cluster #1; BA 47) but also including the pars opercularis and pars triangularis (BA 44 & 45). Anterior perisylvian activation also included the premotor region (BA 6) both on the dorsal lateral (Cluster #3) and medial (Cluster #4) portions of the hemisphere. Posterior perisylvian activation of the left hemisphere was seen primarily in the region of the anterior occipital sulcus, the horizontal posterior segment of the superior temporal sulcus,

peaking in the posterior portion of the middle temporal gyrus (#2; BA 37/21) but also including activation in the angular gyrus (BA 39). A similar activation pattern in the left hemisphere, except for the lack of premotor activity, is seen in the pair-wise comparisons of subject noun plausibility averaged across unambiguous sentences only (Figure 5-3d). A cluster of activation is also seen in the right hemisphere for the medial aspect of BA 6/8 (Cluster #3). No significant clusters of activation are seen when averaged across ambiguous sentences only (Figure 5-3c).

No significant activation is seen for the ambiguity contrast averaged across all subject noun plausibilities (Figure 5-3b), only plausible subject nouns (Figure 5-3e), or only implausible subject nouns (Figure 5-3f).

Finally, for the crossed comparison of plausible ambiguous vs. implausible unambiguous (Figure 5-3g), there is significant activation only in the left inferior frontal region peaking in BA 44 (Cluster #1). For the crossed comparison of plausible unambiguous vs. implausible ambiguous (Figure 5-3h) there is significant activation in the left hemisphere very much like the activation pattern of the plausible vs. implausible subject noun contrast averaged across all ambiguities (Figure 5-3a). Left inferior frontal activation again peaks in BA 47 (Cluster #1) and significant clusters of activation are seen in the premotor region (#2 & #3; BA 6). Left posterior perisylvian activation is split into two clusters with peaks adjoining the angular gyrus (#4 & #5; BA 39). In the right hemisphere, one cluster of activation is seen in the superior parietal region (#8; BA 7).

Table 5-1

Contrast	Left Hemisphere				Right Hemisphere							
	Cluster #	Region	BA	Talairach (x,y,z)	Size (mm2)	p-value	Cluster #	Region	BA	Talairach (x,y,z)	Size (mm2)	p-value
A. Plausible vs. Implausible	1	IFG	47	(-56, 27, -4)	974	0.001585						None
	2	MTG	37/21	(-49, -54, 2)	716	0.000200						
	3	Premotor	6/4	(-47, 0, 44)	222	0.000079						
	4		6m/32	(-19, 12, 48)	323	0.001995						
B. Ambiguous vs. Unambiguous				None								None
C. Plausible vs. Implausible (Ambig)				None								None
D. Plausible vs. Implausible (Unambig)	1	IFG	45/47	(-60, 20, 2)	290	0.00158	3		6m/8m	(1, 17, 44)	305	0.001585
	2	MTG	21/37	(-50, -54, 2)	424	0.00126						
E & F. Ambiguous vs. Unambiguous (Plaus & Implaus)				None								None
G. Plausible Ambig vs. Implausible Unambig	1	IFG	44/6	(-49, 7, 8)	351	0.000398						None
H. Plausible Unambig vs. Implausible Ambig	1	IFG	47	(-56, 21, -3)	518	0.000063	8	Superior		(11, -76, 48)	243	0.00158
	2	Premotor	6	(-18, 5, 56)	494	0.000126		Parietal				
	3	Premotor	6	(-33, -5, 41)	298	0.001000						
	4	Inferior	39/22	(-55, -58, 14)	481	0.000631						
		Parietal										
	5	Inferior	39/19/18	(-57, -68, 13)	292	0.002512						
		Parietal										
6		FU (INS?)	(-28, 16, -11)	326	0.000025							
7		ventrical/callosum	(-15, -2, 9)	230	0.003981							

Table 5-1. Talairach coordinates of peak activation corresponding to the local minimum p-values for each cluster of activated vertices. Cluster # for each contrast corresponds directly to cluster labels on the corresponding contrasts in Figure 5-3.



5.2.3 Behavioral Methods

Participants. Thirteen participants from the MIT community who did not take part in the imaging experiment were paid for their involvement. All were native speakers of English and were naïve as to the purposes of the study.

Materials & Design. Identical items as in the imaging experiment were used, presented in the same pseudo-randomized order but without the presence of fixation trials. The 8 runs were concatenated into a single continuous session for the participants. As in the imaging experiment, target items were not split into separate lists and instead, all items were presented to each subject. Furthermore, there were no filler items.

Procedure. The procedure was the same as in Experiment 1 and the comprehension task that followed each item was a simple yes / no question about the contents of the preceding sentence.

5.2.4 Behavioral Results

Comprehension Task Performance. All participants had comprehension task performances greater than 67% accuracy overall. On average, the comprehension tasks for the experimental items were answered correctly in 92.9% of the trials. The percentages of correct answers for each condition are presented in Table 5-2.

An omnibus 2x2 ANOVA (MV plausible / implausible subject noun x ambiguous / unambiguous structure) showed a significant plausibility effect by participants and by items ($F(1, 12) = 18.25$, $MS_{\text{within}} = 0.0016$, $p < 0.002$; $F(1, 124) = 9.271$, $MS_{\text{within}} = 0.0080$, $p < 0.003$). There was no significant effect of ambiguity ($F_s < 1$) or interaction ($F_s < 2.8$).

Table 5-2

	MV Plausible	MV Implausible
Ambiguous	89.66 (2.44)	96.15 (1.51)
Unambiguous	91.35 (2.24)	94.47 (1.28)

Table 5-2. Mean (standard error) comprehension question performance in percent correct for the behavioral portion of Experiment 4 as a function of MV plausibility and structure ambiguity, by participants.

Reading Times in the Critical Region. Only items with correctly answered comprehension questions were analyzed. Reading time (RT) data points that were greater than 3 standard deviations from the mean were excluded from the analysis, affecting less than 1.3% of the data points in the behavioral experiment overall. Figure 5-4 presents the mean residual RTs (msec/word) across the four conditions in this experiment. The critical region in the ambiguity contrast consists of the disambiguating PP “by the armored car” (see (37)). Table 5-3 presents the mean RTs per word for all regions, in both residual and raw forms.

Across the critical material “by the armored car”, residual RTs were slower for the MV plausible condition than MV implausible both by participants and by items ($F(1, 12) = 13.16$, $MS_{\text{within}} = 100.9$, $p < 0.004$; $F(1, 124) = 4.021$, $MS_{\text{within}} = 897.6$, $p < 0.05$). Furthermore, there was no significant ambiguity effect ($F_s < 1$) and the interaction was significant by participants though not by items ($F(1, 12) = 5.028$, $MS_{\text{within}} = 61.73$, $p < 0.05$; $F_2 < 1.2$). Pair-wise comparisons revealed no significant effects except for the MV plausibility contrast for unambiguous sentences only which was significant by participants and items ($F(1, 12) = 19.99$, $MS_{\text{within}} = 73.08$, $p < 0.0009$; $F(1, 62) = 5.017$, $MS_{\text{within}} = 834.4$, $p < 0.03$). A 2x2 ANOVA examination of the subset of items from Trueswell et al. (1994) produced the same pattern of results in the critical region: a significant main effect of MV plausibility but insignificant main effects of ambiguity and interaction. Thus, the RT results in the critical region directly reflect the comprehension question results of the imaging experiment.

Figure 5-4

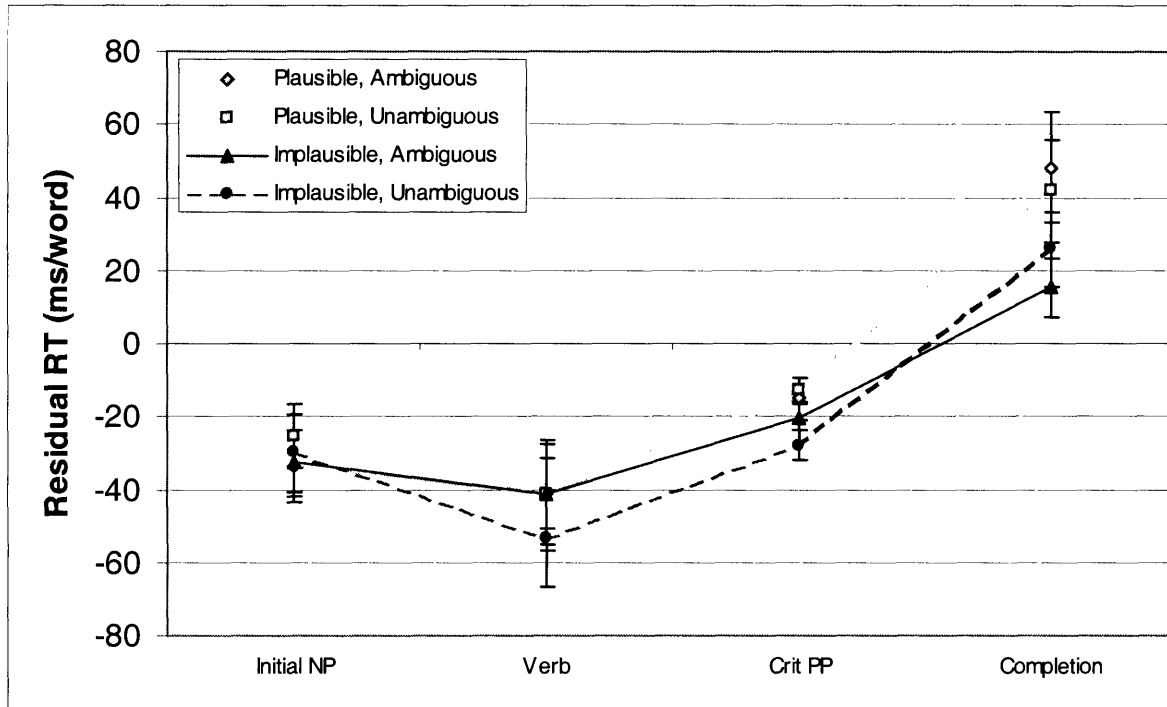


Figure 5-4. Plot of mean (standard error) residual RTs per word for each region of Experiment 4 as a function of MV plausibility of the subject noun and structural ambiguity, by participants.

Table 5-3

	Plausible Subject Noun		Implausible Subject Noun	
	Ambiguous	Unambiguous	Ambiguous	Unambiguous
Matrix Subj	-34 (273)	-26 (282)	-33 (267)	-30 (274)
that was		0 (283)		-22 (268)
Verb	-41 (300)	-41 (296)	-42 (295)	-54 (283)
Critical PP	-15 (275)	-13 (278)	-20 (271)	-28 (263)
Completion	48 (364)	42 (358)	15 (335)	26 (343)

Table 5-3. Mean residual RTs (msec / word) for each region of Experiment 4 as a function of MV plausibility of the subject noun and structural ambiguity, by participants (raw RTs in parentheses).

5.3 Discussion

The MV/RR ambiguity has been very prominent in the psycholinguistic literature and has led to significant insights into the sentence processing mechanism including structural preference, verb information, thematic fit, and referential context (for a brief review, see Tanenhaus & Trueswell, 1995). It is therefore valuable to image the resolution of the MV/RR ambiguity for the potential insight that the neuroimaging data can provide. However, this experiment showed an unexpected result and not the direct look into the ambiguity resolution mechanism that was expected.

In the Trueswell et al. (1994) examination of on-line ambiguity resolution, reaction times representing immediate first stage processing were derived from eye-tracking data at the disambiguating PP. The longest reaction times were found for the MV plausible subject noun, ambiguous sentence condition while the other three conditions had shorter reaction times, all similar in value to each other. Statistically, this was apparent in a significant interaction and a significant main effect of ambiguity. The main effect of subject noun plausibility was marginal however. The longer reaction times were theorized to reflect the parser's attempt to reconcile the conflict between the disambiguating PP indicating a RR structure and the preference for the MV structure based on the plausibility of the initial noun as the subject of the initial verb. The results demonstrated that the manipulation of subject noun plausibility can influence syntactic parsing and that plausibility can be a strong signal for the RR structure.

In our experiment however, both the behavioral measures of the neuroimaging study and the separate self-paced reading study did not demonstrate the Trueswell et al. (1994) pattern of results. In fact, though the MV plausible subject noun, ambiguous sentence condition of the neuroimaging study had a high reaction time and low accuracy percentage, as one would expect, the baseline unambiguous sentence condition with MV plausible subject nouns had similar values for reaction times and accuracy. A comparable pattern was seen in the reading times of the critical

disambiguating PP for the self-paced reading study. There was no ambiguity effect or interaction evident at all and only a main effect of subject noun plausibility.

The main effect of MV plausibility is reminiscent of results seen in Just & Carpenter (1992) in which high span participants, those with greater memory resources for sentence processing, show a main effect of subject noun animacy at the disambiguating PP in an eye-tracking experiment, even for the unambiguous sentences. The results were thought to provide evidence that high span participants had such vast resources available to them that they considered even syntactically impossible interpretations of a sentence and attempted to assign improper θ -roles. Therefore, in the case of unambiguous RC sentences where the structure of the sentence should be well determined, an animate initial noun still results in longer reading times, presumably because of the conflicting indication of a MV structure that the animacy implies and which the high span participants still consider. The suggestion then, is that the participants of this experiment belong to the high span group and, given the fast reaction times (on the order of 1.6 sec) and high accuracy (over 90%), this group of participants does fit into the category of “fast responders” (Waters et al., 2003; Caplan et al., 2003).

The kind of improper θ -role assignments described above has also been seen for syntactically unambiguous simple active and passive sentences as well as the more complex cleft-subject and cleft-object structures (Ferreira, 2003). In a series of experiments, Ferreira (2003) compared actives, passives, cleft-subject, and cleft-object sentences in which the subject and object nouns were symmetrical (e.g. “the man visited the woman” vs. “the woman visited the man”), biased but reversible (e.g. “the dog bit the man”), or irreversible (e.g. “the mouse ate the cheese”). In a task that included asking participants to identify the “DO-ER” and the “ACTED-ON” described in the target items, it was found that participants had more difficulty performing the task for passive and object-cleft sentences than actives and subject-cleft sentences, respectively, regardless of the reversibility of the nouns. The proposal was that the sentence comprehension mechanism makes use

of a quick-and-dirty Noun-Verb-Noun (NVN) strategy, in which the initial subject noun is the DO-ER while the second object noun is the ACTED-ON, in addition to a more detail syntactic analysis to determine the θ -role assignments in a typical sentence. In the case of non-canonical, yet still syntactically unambiguous sentences like the passive and object-clefts, conflict between the NVN strategy and the full parse results in greater difficulty and more mis-assignments of θ -roles. In the case of the RR structures presented in this experiment, the nouns are also in a non-canonical order such that the initial noun is the ACTED-ON and not the DO-ER for the embedded clause. In light of the Ferreira (2003) results, it is conceivable that, even with syntactically unambiguous RC structures, the canonical MV structure was still considered a possible interpretation.

It should be noted that nature of the comprehension question task might have especially encouraged participants to consider the option of a MV interpretation. For the task of this experiment, the MV interpretation may have been considered even if syntactically impossible because of the need to exhaustively determine the veracity of the probe question, especially when given the chance with a plausibility cue in the target sentence. Similarly, for Ferreira (2003), the task also focused on θ -role assignments, which may have encouraged participants to consider a more complete search of possible candidates in the presented sentences than would be normal. Regardless of the unexpected results for unambiguous sentences, the main effect of subject noun plausibility does show the influence that the semantic information has on syntactic parsing and the imaging data provided here reflects the resolution of that conflict.

Turning to the issues concerning the localization of the processes described above, one of the first things noticed from the activation maps (Figure 5-3) is the lack of significant activation in any of the ambiguous vs. unambiguous contrasts (Figure 5-3b, e, f) as reflected in the behavioral results. On the other hand, MV plausible vs. implausible subject noun contrast revealed robust activation in three primary locations, all in the left hemisphere: the inferior frontal cortex, the dorsal prefrontal

cortex / premotor strip, and the anterior occipital sulcus (Figure 5-3a). The presence of a similar activation pattern in the MV plausibility contrast for unambiguous sentences only (Figure 5-3d), and the crossed comparison of MV plausible, unambiguous vs. MV implausible, ambiguous (Figure 5-3h) would imply that the MV plausible, unambiguous condition has the greatest BOLD signal strength of the four conditions in these three regions. That the MV plausible, unambiguous condition imposes the greatest load on the parser was an unexpected result because the syntactic structure in this condition should be unequivocal and no ambiguity resolution should be necessary. As discussed above, it is possible that the activation reflects a combination of the high span of the participants, conflict with the NVN template, and the nature of the comprehension question task. The activation from the subject noun MV plausibility contrast likely reflects the overall difficulty that the ambiguity resolution mechanism has as it deals with the conflicting semantic plausibility information of the subject noun indicating a possible MV structure and the ambiguous or unambiguous RC syntactic structure. If this is the case, the idea of “ambiguity” must be expanded such that sentences that are traditionally considered syntactically unambiguous can indeed still be ambiguous as to their ultimate meanings. And if these results are replicable, we must reconsider what the boundaries are for the consideration of possible θ -role or other propositional feature assignments, as syntax does not appear to be the final deciding factor in these data.

The left inferior frontal cortex presents broad activation for the MV plausibility contrast across the entire gyrus and peaks in the pars orbitalis (BA 47). An examination of the literature seems to indicate that the inferior frontal gyrus is the neural substrate for several language comprehension processes, including working memory that possibly mediates syntactic processes, lexico-semantics, and phonological processing. In fact, in a continued vein of anatomical modularity, several papers have proposed that the inferior frontal gyrus can be subdivided for these processes (Dapretto & Bookheimer, 1999; Poldrack et al., 1999; Muller et al., 2003). A meta-analysis of these works suggests that the anterior portions of the inferior temporal gyrus (BA 47, 45)

may be specialized for lexico-semantic processes while posterior portions (BA 44, 45) are specialized for syntactic and phonological processes. Because of the broad activation across the left inferior frontal cortex, the results here could indicate that both syntactically- and lexically-mediated ambiguity resolution processes are involved in the MV plausibility contrast.

The dorsal prefrontal cortex however, has been implicated in the motivation for lexico-semantic information retrieval (Binder & Price, 2001). If this is the case, greater activation for the condition of MV plausible subject noun could be representative of more effortful processing in which the ambiguity resolution mechanism must change the salient verb argument structure from the preferred MV structure to the RR structure. Thus, the ambiguity resolution mechanism may need to motivate for re-retrieval of the relevant lexico-semantic information. However, it must be noted that similar dorsal prefrontal cortex activation has also been seen with syntactic processing associated with working memory loads (Experiment 2; Chen et al., in preparation). Similarly, the angular gyrus and middle temporal gyrus (BA 39, 37/21) included in the occipital sulcus activation have both been implicated in lexico-semantic processing. Therefore, the activation in the occipital sulcus could also represent the more effortful re-retrieval of the lexico-semantic information related to the less preferred RR argument structure of the past participle verb. Though angular gyrus activation has been seen in studies of syntactic processing (see Chapter 4), the activation here is more inferior and includes the middle temporal gyrus, and so may be functionally distinct.

The results presented here were unexpected but are an important first step in the use of imaging techniques to gain insight into a fundamental psycholinguistic contrast. The comprehension question task and the high spans of the participants encouraged the consideration of the MV interpretation based on subject noun MV plausibility information, despite the unambiguous RC syntactic structure. Thus, no ambiguity effect was evident. The overall pattern of activation is most consistent with a theory of lexically-mediated ambiguity resolution in the MV/RR ambiguity. However, it is not possible to entirely rule out syntactically mediated processing as many of the same

anatomical regions are involved, especially in the case of the inferior frontal gyrus. It is clear that further experiments will be needed, including the use of a different task and controlling for the span of the participants, to better isolate the processes of MV/RR ambiguity resolution.

Chapter 6

Conclusion

This thesis was an exploration into two important constraints in the syntactic processing mechanism, syntactic storage costs and plausibility information. It is also one in a series of studies that begin to use neuroimaging as a technique for exploring the issues of syntactic processing by providing information on the anatomical locations involved and the possible cognitive processes that they support.

Experiment 1 behaviorally demonstrated the online presence of storage costs for predictions of verbs, filler-gaps, and subcategorized prepositional phrases. For Experiment 1a, the data showed a three-tiered response for zero, one, and two predicted verbs across the critical region. The greater the number of verb predictions the parser had to maintain, the greater the load and the slower the reading times. In regions other than the critical one, the data that did not satisfy the predictions of models of storage costs could be explained by confounding factors such as interference effects. In Experiments 1b and 1c, slower reading times were demonstrated for the storage costs of filler-gap predictions and for subcategorized prepositional phrase predictions. As a result, it was determined that the data supported the Gibson's (1998) Dependency Locality Theory (DLT) theory of stored predicted heads as well as a theory of syntax that includes empty categories as syntactic heads.

Experiment 2 demonstrated the regions of the brain important to processes associated with the maintenance and integration of predicted syntactic heads. By measuring the BOLD signal response to so-called subject-object (SO) and object-subject (OS) sentences, the experiment reinforced results that showed the consistent presence of Broca's area activation, but the inconsistent presence of posterior perisylvian activation, in this case the bilateral inferior parietal cortex. The results seem to indicate that the inferior parietal cortex is part of a network of cortex including Broca's area that is involved in the processing of SO vs. OS sentences, but not involved in an identical manner. The exact nature of the activation pattern is still not clear and will require further investigation.

Experiment 3 was an attempt at developing a neuroimaging experiment to determine the regions of activation associated with the storage costs of filler-gap predictions found in Experiment 1b. Unfortunately, the items that were developed, while reinforcing the results of the behavioral experiments, were not robust enough to warrant expending further neuroimaging resources.

Finally, Experiment 4 demonstrated regions of the brain involved in the resolution of the main verb / reduced relative (MV/RR) ambiguity. A pattern that seemed to reflect lexically-mediated processing showed activation in Broca's area as well as the posterior portion of the dominant (left) medial temporal cortex. Yet, activation was seen only for the subject noun MV plausibility contrast regardless of whether the sentences were syntactically ambiguous or not, apparently indicating that the MV interpretation was still considered, when it was plausible, even for the syntactically unambiguous RC structure. The results demonstrate that Broca's area and the dominant posterior middle temporal gyrus are involved in the resolution of the conflict between the plausibility of the alternative structures and Ferreira's (2003) NVN template cues on the one hand, and the full syntactic parse on the other. While there was no evidence on the anatomical localization of the ambiguity resolution process, the experiment does provide insight into the use of plausibility cues and the interaction of structural and pragmatic information.

This thesis has demonstrated that neuroimaging promises to be a fruitful technique that can provide great insight into the sentence comprehension mechanism. However, as we have seen, it is not a straightforward path and much work in the fundamental contrasts still needs to be done. Furthermore, in addition to the neuroimaging work, this thesis has helped develop a model of storage costs, which combined with our understanding of ambiguity resolution has the promise to aid in the exploration other fundamental questions in sentence processing including the serial vs. parallel nature of the parser.

References

- Abney SP, Johnson M (1991): Memory requirements and local ambiguities of parsing strategies. *J Psycholing Research* 20: 233-250.
- Altmann G, Steedman M (1988): Interaction with context during human sentence processing. *Cognition* 30: 191-238.
- Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993): Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine* 30: 161-173.
- Bates E, Goodman JC (1997): On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Lang Cogn Processes* 12: 507-584.
- Bavelier D, Corina D, Jezzard P (1997): Sentence reading: a functional MRI study at 4 Tesla. *J Cogn Neurosci* 9: 664-686.
- Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, Grodzinsky Y (in press): The neural reality of syntactic transformations: Evidence from fMRI. *Psych Sci*.
- Berndt R, Mitchum C, Haendiges A (1996): Comprehension of reversible sentences in agrammatism: a meta-analysis. *Cognition* 58: 289-308.
- Bever TG (1970): *The cognitive basis for linguistic structure*. New York: Wiley.
- Binder J, Price CJ (2001): Functional neuroimaging of language. In: Cabeza R, Kingstone A (Eds). *Handbook of Functional Neuroimaging of Cognition*. Bradford Book, Massachusetts. Pp. 187-251.
- Boland JE (1997): The relationship between syntactic and semantic processes in sentence comprehension. *Lang Cogn Processes* 12(4): 423-484.
- Bookheimer SY, Zeffiro TA, Gallard W, Theodore W (1993): Regional cerebral blood flow changes during the comprehension of syntactically varying sentences. *Neuroscience Society Abstracts* 347.5: 843.
- Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM (1998): Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *Neuroreport* 9: 3735-3739.
- Burock MA, Dale AM (2000): Estimation and detection of event-related fMRI signals with temporally correlated noise: A statistically efficient and unbiased approach. *Human Brain Map* 11: 249-260.
- Caplan D (1995): Issues arising in contemporary studies of disorders of syntactic processing in sentence comprehension in agrammatic patients. *Brain Lang* 50: 325-338.
- Caplan D, Alpert N, Waters G (1998): Effects of syntactic structure and propositional number of patterns of regional cerebral blood flow. *J Cogn Neurosci* 10: 541-552.

- Caplan D, Alpert N, Waters G (1999): PET studies of sentence processing with auditory sentence presentation. *Neuroimage* 9: 343-351.
- Caplan D, Alpert N, Waters G, Olivieri A (2000): Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Map* 9: 65-71.
- Caplan D, Baker C, Dehaut F (1985): Syntactic determinants of sentence comprehension in aphasia. *Cognition* 21: 117-175.
- Caplan D, Hildebrandt N, Makris N (1996): Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* 119: 933-949.
- Caplan D, Vijayan S, Kuperberg G, West C, Waters G, Greve D, Dale AM (2002): Vascular responses to syntactic processing: Event-related fMRI study of relative clauses. *Human Brain Map* 15: 26-38.
- Caplan D, Waters G, Alpert N (2003): Effects of age and speed of processing on rCBF correlates of syntactic processing in sentence comprehension. *Human Brain Map* 19: 112-131.
- Caplan D, Waters G, Hildebrandt N (1997): Determinants of sentence comprehension in aphasic patients in sentence-picture matching tasks. *J Speech Hear Res* 40: 542-555.
- Caramazza A, Capitani E, Rey A, Berndt RS (2001): Agrammatic Broca's Aphasia is not associated with a single pattern of comprehension performance. *Brain Lang* 76: 158-184.
- Carpenter PA, Just MA, Keller TA, Eddy WF, Thulborn KR (1999): Time course of fMRI-activation in language and spatial networks during sentence comprehension. *Neuroimage* 10: 216-224.
- Caviness VS, Meyer J, Makris N, Kennedy DN (1996): MRI-based topographic parcellation of human neocortex: An anatomically specified method with estimate of reliability. *J Cogn Neurosci* 8: 566-587.
- Chen E, West WC, Waters G, Caplan D (in preparation): BOLD signal correlates of sentence comprehension: Effects of syntactic processing and related cognitive processes.
- Chomsky N (1981): *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky N, Miller GA (1963): Introduction to the formal analysis of natural languages. In: Luce RD, Bush RR, Galanter E (Eds). *Handbook of Mathematical Psychology, Volume 2*. Wiley, New York. Pp. 269-321.
- Cooke A, Zurif EB, DeVita C, Alsop D, Koenig P, Detre J, Gee J, Pinãngo M, Balogh J, Grossman M (2001): Neural basis for sentence comprehension: grammatical and short-term memory components. *Human Brain Map* 15: 80-94.
- Cox RW (1996): AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers & Biomedical Research* 29: 162-173.

- Crain S, Steedman M (1985): On not being led up the garden path: The use of context by the psychological parser. In: Dowty D, Karttunen L, Zwicky A (Eds). *Natural Language Parsing*. Cambridge, UK: Cambridge University Press.
- Crocker M (1999): Mechanisms for sentence processing. *Language Processing*. In: Garrod S, Pickering M (Eds). Cambridge, MA, MIT Press.
- Dale AM (1999): Optimal experimental design for event-related fMRI. *Human Brain Map* 8: 109-114.
- Dale AM, Buckner RL (1997): Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Map* 5: 329-340.
- Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis - I: Segmentation and surface reconstruction. *Neuroimage* 9:179-194.
- Damasio AR (1992): Aphasia. *The New England Journal of Medicine* 326: 531-539.
- Dapretto M, Bookheimer SY (1999): Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron* 24: 427-432.
- Dick F, Bates E, Wulfeck B, Aydelott Utman J, Dronkers N, and Gernsbacher MA (2001): Language Deficits, Localization, and Grammar: Evidence for a Distributive Model of Language Breakdown in Aphasic Patients and Neurologically Intact Individuals. *Psych Review* 108: 759-788.
- Eberhard K, Sedivy JC, Spivey-Knowlton MJ, Tanenhaus MK (1995): Eye movements as a window into real-time spoken language processing in natural contexts. *Journal of Psycholinguistic Research* 24: 409-436.
- Ferreira F (2003): The misinterpretation of noncanonical sentences. *Cogn Psych* 47: 164-203.
- Ferreira F, Clifton C (1986): The independence of syntactic processing. *J Mem Lang* 25: 348-368.
- Fiebach CJ, Schlesewsky M, Friederici AD (2001a): Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *J Psycholing Research* 30: 321-338.
- Fiebach CJ, Schlesewsky M, Friederici AD (2001b): Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *J Mem Lang* 47: 250-272.
- Fischl B, Sereno MI, Dale AM (1999a): Cortical surface-based analysis - II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9: 195-207.
- Fischl B, Sereno MI, Tootell RBH, Dale AM (1999b): High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Map* 8: 272-284.
- Fodor JA (1983): *Modularity of mind*. Cambridge, MA: MIT Press.

Ford M (1983): A method for obtaining measures of local parsing complexity throughout sentences. *J Verbal Learn Verbal Behav*, 22, 203-218.

Frazier L (1978): *On comprehending sentences: Syntactic parsing strategies*. Bloomington: Indiana University Linguistics Club.

Frazier L, Clifton C (1996): *Construal*. Cambridge, MA, MIT Press.

Frazier L, Fodor JD (1978): The sausage machine: A new two-stage parsing model. *Cognition* 6: 291-326.

Frazier L, Rayner K (1982): Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentence. *Cog Psych* 14: 178-210.

Garnsey SM, Pearlmutter NJ, Myers E, Lotocky MA (1997): The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentence. *J Mem Lang* 37: 58-93.

Gazdar G, Klein E, Pullum GK, Sag IA (1985): *Generalized phrase structure grammar*, Basil Blackwell, Oxford, UK.

Gibson E (1991): *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

Gibson E (1998): Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68: 1-76.

Gibson E (2000): The dependency locality theory: A distance-based theory of linguistic complexity. In: Miyashita Y, Marantz A, O'Neil W (Eds). *Image, language, brain*. Cambridge, MA: MIT Press. Pp. 95-126.

Gibson E, Desmet T, Grodner D, Watson D, Ko K (2002): Reading relative clauses in English. MIT manuscript under review.

Gibson E, Hickok G (1993): Sentence processing with empty categories. *Lang and Cogn Processes* 8: 147-161.

Gibson E, Pearlmutter NJ (1998): Constraints on sentence comprehension. *Trends in Cog Sci* 2: 262-268.

Gibson E, Thomas J (1997): The complexity of nested structures in English: Evidence for the syntactic prediction locality theory of linguistic complexity. Unpublished paper. MIT, Cambridge, MA

Gibson E, Warren T (1999): The psychological reality of intermediate linguistic structure in long-distance dependencies. MIT manuscript under review.

Gordon PC, Hendrick R, Johnson M (2001): Memory interference during language processing. *J Expt Psych - Learning, Memory and Cognition* 27(6): 1411-1423.

- Grodner D, Gibson E, Tunstall S (2002): Syntactic complexity in ambiguity resolution. *J Mem Lang* 46: 267-295.
- Grodner D, Gibson E, Watson D (2001): The influence of contextual contrast on syntactic processing: Evidence for strong-interaction in sentence comprehension. Poster presented at the 14th Annual CUNY Conference on Human Sentence Processing in Philadelphia, PA.
- Grodzinsky Y (2000): The neurology of syntax: Language use without Broca's area. *Behav Brain Sci* 23: 47-117.
- Hakuta K (1981): Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition*, 9, 197-236.
- Hashimoto R, Sakai KL (2002): Specialization in the left prefrontal cortex for sentence comprehension. *Neuron* 35: 589-597.
- Holmes V (1973): Order of main and subordinate clauses in sentence perception. *J Verbal Learn Verbal Behav* 12: 285-293.
- Holmes VM, O'Regan JK (1981): Eye fixation patterns during the reading of relative clause sentences. *J Verbal Learn Verbal Behav* 20: 417-430.
- Homae F, Hashimoto R, Nakajima K, Miyashita Y, Sakai KL (2002): From perception to sentence comprehension: the convergence of auditory and visual information of language in the left inferior frontal cortex. *Neuroimage* 16: 883-900.
- Juliano C, Tanenhaus MK (1994): A constraint-based lexical account of the subject/object attachment preference. *J Psycholing Research* 23: 459-472.
- Just MA, Carpenter PA (1992): A capacity theory of comprehension: Individual differences in working memory. *Psych Review* 99: 122-149.
- Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR (1996): Brain activation modulated by sentence comprehension. *Science* 274: 114-116.
- Just MA, Carpenter PA, Woolley JD (1982): Paradigms and processing in reading comprehension. *J Expt Psych: General* 111: 228-238.
- Kann E, Swaab T (2002): The brain circuitry of syntactic comprehension. *Trends in Cogn Sci* 6: 350-356.
- Keller TA, Carpenter PA, Just MA (2001): The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cereb Cortex* 11: 223-237.
- Kennison SM (2000): Two assessments of American usage frequencies for ninety-seven sentence complement-taking nouns. *J Psycholing Research* 29: 249-264.
- Kimball J (1973): Seven principles of surface structure parsing in natural language. *Cognition* 2: 15-47.

King J, Just MA (1991): Individual differences in syntactic processing: the role of working memory. *J Mem Lang* 30: 580-602.

Konieczny L (2000): Locality and parsing complexity. *J Psycholing Research* 29(6): 627-645.

Konieczny L, Doring P (2003): Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. *Proceedings of the ICCS/ASCS-2003 Joint International Conference on Cognitive Science*. Sydney, Australia 13 - 17 July 2003

Lewis RL (1996): A theory of grammatical but unacceptable embeddings. *J Psycholing Research* 25: 93-116.

Lewis RL (2000): Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *J Psycholing Research* 29(2): 241-248.

MacDonald MC (1994a): Probabilistic constraints and syntactic ambiguity resolution. *Lang Cogn Processes* 9: 15-202.

MacDonald MC (1994b): Lexical representations and sentence processing: An introduction. *Lang Cogn Processes* 12: 121-136.

MacDonald MC, Pearlmutter NJ, Seidenberg MS (1994): Lexical nature of syntactic ambiguity resolution. *Psych Review* 101: 676-703.

Marslen-Wilson WD (1973): Linguistic structure and speech shadowing at very short latencies. *Nature* 244: 522-523.

Mazoyer B, Tzourio N, Frak V, Syrota A, Mrayama N, Levrier O, Salamon G (1993): The cortical representation of speech. *J Cogn Neurosci* 5: 467-479.

McElree B, Griffith T (1995): Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *J Expt Psych - Learning, Memory and Cognition* 21(1): 134-157.

Mesulam M-M (1990): Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology* 28(5): 597-613.

Miller GA, Isard S (1964): Free recall of self-embedded English sentences. *Information and Control* 7: 292-303.

Miyamoto E (2002): Case markers as clause boundary inducers in Japanese. *J Psycholing Research* 31: 307-347.

Muller R-A, Kleinhans N, Courchesne E (2003): Linguistic theory and neuroimaging evidence: An fMRI study of Broca's area in lexical semantics. *Neuropsychologia* 41: 1199-1207.

Nakatani K, Gibson E (2003): An on-line study of Japanese nesting complexity. Poster presented at the 16th Annual CUNY Conference on Human Sentence Processing, MIT, Cambridge, March 2003.

- Ni W, Constable RT, Mencl WE, Pugh KR, Fulbright RK, Shaywitz SE, Shaywitz BA, Gore JC, Shankweiler D (2000): An event-related neuroimaging study distinguishing form and content in sentence processing. *J Cogn Neurosci* 12: 120-133.
- Noguchi Y, Takeuchi T, Sakai KL (2002): Lateralized activation in the inferior frontal cortex during syntactic processing: event-related optical topography study. *Human Brain Map* 17: 89-99.
- Pearlmutter NJ, Mendelsohn A (1999): Serial versus parallel sentence comprehension. Northeastern University, Department of Psychology, manuscript submitted for publication.
- Pickering M, Barry G (1991): Sentence processing without empty categories. *Lang Cogn Processes* 6: 229-259.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JDE (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10: 15-35.
- Pollard C, Sag IA (1994): *Head-driven phrase structure grammar*, University of Chicago Press, Chicago, IL.
- Rademacher J, Galaburda AM, Kennedy DN, Filipek PA, Caviness VS (1992): Human cerebral cortex localization, parcellation, and morphometry with magnetic resonance imaging. *J Cogn Neurosci* 4: 352-373.
- Rayner K, Duffy SA (1986): Lexical complexity and fixation times in reading - effects of word frequency, verb complexity, and lexical ambiguity. *Mem & Cogn* 14: 191-201.
- Rayner L, Carlson M, Frazier L (1983): The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *J Verbal Learn & Verbal Behav* 22: 358-374.
- Röder B, Stock O, Neville H, Siegfried B, Rösler F (2002): Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: a functional magnetic resonance imaging study. *Neuroimage* 15: 1003-1014.
- Schmauder AR, Kennison SM, Clifton C (1991): On the conditions necessary for obtaining argument structure complexity effects. *J Expt Psych - Learning, Memory and Cognition* 17: 1188-1192.
- Shapiro LP, Zurif E, Grimshaw J (1987): Sentence processing and the mental representation of verbs. *Cognition* 27: 219-246.
- Speer SR, Bernstein MK (1992): The influence of prosodic structure on processing temporary syntactic ambiguity. *Bulletin of the Psychonomic Society* 30: 478.
- Spivey-Knowlton MJ, Trueswell JC, Tanenhaus MK (1993): Context effects in syntactic ambiguity resolution: Parsing reduced relative clauses. *Canadian J Psych* 47: 276-309.

- Stabler EP (1994): The finite connectivity of linguistic structures. In: Clifton C, Frazier L, Rayner K (Eds). *Perspectives on sentence processing*. Hillsdale, NJ: Erlbaum. Pp. 303-336.
- Stolz WS (1967): A study of the ability to decode grammatically novel sentences. *J Verbal Learn Verbal Behav* 6: 867-873.
- Stowe LA (2000): Sentence comprehension and the left inferior frontal gyrus: storage, no computation. *Behav and Brain Sci* 23: 51.
- Stowe LA, Broere CAJ, Paans AMJ, Wijers AA, Mulder G, Vaalburg W, Zwarts F (1998): Localizing components of a complex task: sentence processing and working memory. *Neuroreport* 9: 2995-2999.
- Stowe LA, Wijers AA, et al. (1994): PET-studies of language: an assessment of the reliability of the technique. *J Psycholing Res* 23: 499-527.
- Stromswold K, Caplan D, Alpert N, Rauch S (1996): Localization of syntactic comprehension by positron emission tomography. *Brain Lang* 52: 452-473.
- Swinney D, Zurif E (1995): Syntactic processing in aphasia. *Brain Lang* 50: 225-239.
- Swinney D, Zurif E, Prather P, & Love T (1996): Neurological distribution of processing resources underlying language comprehension. *J Cog Neurosci* 8: 174-184.
- Talairach J, Tournoux P (1988): *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers, Inc.
- Tanenhaus MK, Trueswell JC (1995): Sentence Comprehension. In: Miller JL, Eimas PD (Eds). *Speech, Language, and Communication*. Academic Press, Inc.
- Trueswell JC, Tanenhaus MK, Garnsey SM (1994): Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *J Mem Lang* 33: 285-318.
- Turner R (1997): Signal sources in BOLD contrast fMRI. *Adv Exp Med Biol* 413: 19-25.
- Ullman MT (2001): A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience* 2: 717-726.
- Vasishth S (2002): *Working memory in sentence comprehension: Processing Hindi center embeddings*. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.
- Wanner E, Maratsos M (1978): An ATN approach in comprehension. In: Halle M, Bresnan J, Miller G (Eds). *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press. Pp. 119-161.
- Warren T, Gibson E (2002): The influence of referential processing on sentence complexity. *Cognition* 85: 79-112.

Waters G, Caplan D, Alpert N, Stanczak L (2003): Individual differences in rCBF correlates of syntactic processing in sentence comprehension: Effects of working memory and speed of processing. *Neuroimage* 19: 101-112.

Yngve VH (1960): A Model and an Hypothesis for Language Structure. *Proceedings of the American Philosophical Society* 104: 444-466.

Appendix A: Experiment 1a Materials

In the experimental items below, the four conditions for stimulus (1) of Experiment 1a can be derived as follows. For the 0 incomplete verb dependency condition, choose the verb form for the first two clauses (e.g. “the detective suspected” and “the thief knew”) and the first completion (e.g. “and so he reported immediately to the museum curator”). For the 1 Late condition pick the verb form for the first clause and the nominalized form for the second (e.g. “the knowledge”) and the second completion (e.g. “came from an insider”). For the 1 Early condition pick the nominalized form in the first clause (e.g. “the suspicion”), the verb form of the second clause (e.g. “the thief knew”), and the third completion (e.g. “worried the museum curator”). For the 2 incomplete verb dependencies condition, pick the nominalized form of the two initial clauses and the last completion (e.g. “came from an insider worried the museum curator”).

1. (The detective suspected / The suspicion) that (the thief knew / the knowledge) that the guard protected the jewels (and so he reported immediately to the museum curator. / came from an insider. / worried the museum curator. / came from an insider worried the museum curator.)
2. (The dictator insisted / The insistence) that (the country acknowledged / the acknowledgement) that the army violated the treaty (because he felt enraged. / be made public. / threatened the unstable peace. / be made public threatened the unstable peace.)
3. (The employee realized / The realization) that (the boss implied / the implication) that the company planned a layoff (and so he sought alternative employment. / had been unintentional. / caused a panic. / had been unintentional caused a panic.)
4. (The clerk recommended / The recommendation) that (the customer should complain / the complaint) that the product was faulty (and so he pointed to Customer Service. / should be filed immediately. / was taken seriously. / should be filed immediately was taken seriously.)
5. (The mother sensed / The sense) that (the child feared / the fear) that a monster might eat little boys (and so she kept a light on. / kept the boy up at night. / worried the parents greatly. / kept the boy up at night worried the parents greatly.)
6. (The psychiatrist worried / The worry) that (the patient felt / the feeling) that everyone deserved to die (and so he quickly called hospital security. / could not be helped. / concerned the family. / could not be helped concerned the family.)
7. (The lawyer acknowledged / The acknowledgement) that (the defendant had hinted / the hint) that the mob bribed the official (and so he went with a new line of questioning. / had been ignored completely. / was recorded by the court. / had been ignored completely was recorded by the court.)
8. (The producer doubted / The doubt) that (the director would realize/ the realization) that the actress hated the lead actor (and so he immediately sent a messenger. / might eventually cause trouble. / puzzled the motion picture executives. / might eventually cause trouble puzzled the motion picture executives.)
9. (The counselor implied / The implication) that (the teacher should know / the knowledge) that the student had a disability (and so he took matters into his own hand. / should help the school in dealing with the student. / was not given enough emphasis. / should help the school in dealing with the student was not given enough emphasis.)
10. (The author contended / The contention) that (the publisher predicted / the prediction) that the novel would be a success (and so he planned to buy a new car. / was nothing more than adulation. / amused the public. / was nothing more than adulation amused the public.)
11. (The banker worried / The worry) that (the investor sensed / the sense) that the company was performing poorly (and so he called to offer reassurance. / might influence stock prices. / was well founded. / might influence stock prices was well founded.)
12. (The king doubted / The doubt) that (the squire felt / the feeling) that the knight would win the joust (and so he reprimanded the young boy. / was purely emotional. / proved to be true. / was purely emotional proved to be true.)
13. (The historian hypothesized / The hypothesis) that (the emperor believed / the belief) that the elixir gave long life (and so he wrote an essay on the topic. / was promoted by alchemists. / was not a big surprise. / was promoted by alchemists was not a big surprise.)

14. (The lifeguard thought / The thought) that (the swimmer implied / the implication) that a shark was in the waters (and so he yelled for everyone to get out. / might have been a joke. / made everyone angry. / might have been a joke made everyone angry.)
15. (The family knew / The knowledge) that (the weatherman predicted / the prediction) that the storm would cause severe damage (and so they boarded up their windows. / would be accurate. / caused some panic. / would be accurate caused some panic.)
16. (The homeowner suggested / The suggestion) that (the contractor should request / the request) that the carpenter use maple instead (and so the project was delayed. / was an excellent decision. / made everyone happy. / was an excellent decision made everyone happy.)
17. (The magician realized / The realization) that (the assistant feared / the fear) that the guillotine was actually real (and so he removed the trick from the act. / was not an act. / came almost too late. / was not an act came almost too late.)
18. (The conductor claimed / The claim) that (the engineer indicated / the indication) that the train was in good condition (and so the accident was not his fault. / was a lie. / became a front page story. / was a lie became a front page story.)
19. (The brother thought / The thought) that (the sister hoped / the hope) that the couple would break up (and so everyone was soon gossiping. / was fruitless. / made everyone laugh. / would be fruitless made everyone laugh.)
20. (The chauffeur suspected / The suspicion) that (the executive wished / the wish) that the documents would be shredded (and so he discreetly took them. / was part of the scandal. / would become a huge story. / was part of the scandal would become a huge story.)
21. (The paramedic feared / The fear) that (the fireman thought / the thought) that the wound was not serious (and so he injected a strong sedative. / would result in further injury. / was precautionary. / would result in further injury was precautionary.)
22. (The tourist implied / The implication) that (the waiter suggested / the suggestion) that the tip was too small (and so the manager apologized profusely. / was inappropriate. / concerned the manager. / was inappropriate concerned the manager.)
23. (The athlete contended / The contention) that (the trainer claimed / the claim) that the injury was not too severe (and so the coach put him back in the game. / was gross malpractice. / concerned the general manager. / was gross malpractice concerned the general manager.)
24. (The photographer sensed / The sense) that (the model worried / the worry) that the photos were not tasteful (and so he tried to convince her. / could not be helped. / would delay the shoot. / could not be changed would delay the shoot.)
25. (The movers wished / The wish) that (the family knew / the knowledge) that the boxes were too heavy (and so they complained loudly while working. / would mean a bigger tip. / went unfulfilled. / would mean a bigger tip went unfulfilled.)
26. (The caddie hoped / The hope) that (the golfer sensed / the sense) that the wind was blowing eastwards (and so he pulled out a pitching wedge. / was correct. / turned out to be in vain. / was correct turned out to be in vain.)
27. (The butler recommended / The recommendation) that (the gardener should insist / the insistence) that the master relandscape the property (and so he set up a meeting. / should be considered. / fell on deaf ears. / should be considered fell on deaf ears.)
28. (The sailor indicated / The indication) that (the diver complained / the complaint) that the tanks were half full (and so the captain ordered new ones brought up. / had been noted already. / worried the admiral. / had been noted already worried the admiral.)
29. (The brewer hoped / The hope) that (the merchant would sense / the sense) sense that the beer was made from fine ingredients (and so he poured a sample glass. / meant more business. / was perfectly within reason. / meant more business was perfectly within reason.)
30. (The grocer hinted / The hint) that (the chef thought / the thought) that the produce was the best in town (and so he beamed a big smile. / was completely accurate. / seemed arrogant. / was completely accurate seemed arrogant.)
31. (The farmer thought / The thought) that (the winemaker recommended / the recommendation) that the grapes should be harvested (and so he warmed up the tractor. / was premature. / had been correct. / was premature had been correct.)

32. (The parishioner insinuated / The insinuation) that (the bishop knew / the knowledge) that the priest had a criminal history (and so there was a public outcry. / did not result in action. / made the public angry. / did not result in action made the public angry.)
33. (The jockey feared / The fear) that (the veterinarian implied / the implication) that the horse was not fit to race (and so he consulted with the owner. / meant disaster. / was an over-reaction. / meant disaster was an over-reaction.)
34. (The lawyer indicated / The indication) that (the brother claimed / the claim) that the will had been tampered with (and so the papers were examined closely. / was completely untrue. / calmed everyone down. / was completely untrue calmed everyone down.)
35. (The aide speculated / The speculation) that (the president would acknowledge / the acknowledgment) acknowledge that the situation had become worse (and so a speech was prepared. / would come that night. / was more than rumor. / would come that night was more than rumor.)
36. (The principal acknowledged / The acknowledgment) that (the librarian complained / the complaint) that the books were falling apart (and so he poured over the budget for more money. / could not be addressed. / spurred the parents to donate. / could not be addressed spurred the parents to donate.)
37. (The chiropractor realized / The realization) that (the patient feared / the fear) that the injury might be incurable (and so he tried to be reassuring. / was causing the unpleasantness. / came as an epiphany. / was causing the unpleasantness came as an epiphany.)
38. (The demonstrator doubted / The doubt) that (the executive felt / the feeling) that the environment truly was important (and so he hurled rotten eggs at the car. / was sincere at all. / had only been expected. / was sincere at all had only been expected.)
39. (The authority sensed / The sense) that (the smuggler hoped / the hope) that the stash would not be found (and so he searched harder. / was becoming desperate. / had only been natural. / was becoming desperate had only been natural.)
40. (The astronomer requested / The request) that (the technician should promise / the promise) that the telescope would get daily maintenance (and so he arranged a schedule. / be in writing. / was unreasonable. / be in writing was unreasonable.)

Appendix B: Experiment 1b Materials

The four conditions for item (1) of Experiment 1b can be derived as follows. For the ambiguous conditions, always include “that”. For the unambiguous SC condition, include the present participle (e.g. “implying”) plus “that” while for the unambiguous RC condition, include “which” instead of “that”. Regarding the region following the embedded verb (e.g. “ridiculed”), do as follows: for the SC conditions, include the embedded object (e.g. “the newscaster”); for the RC conditions, omit the object.

The number following each item number is the percent SC continuations in Pearlmutter & Mendelsohn’s (1999) completion norming data.

1. (93%) The reason (implying) (that/which) the comedian who the network fired ridiculed (the newscaster) was kept a secret.
2. (48%) The revelation (showing) (that/which) the executive who the company employed belittled (the secretary) had to be completely ignored.
3. (59%) The perception (suggesting) (that/which) the photographer who the magazine hired abused (the model) was an issue at the agency.
4. (60%) The verification (confirming) (that/which) the ranger who the tourists trusted checked (the canoe) came over the walkie-talkie.
5. (44%) The conclusion (reporting) (that/which) the spy who the CIA pursued reached (the checkpoint) has not yet been confirmed.
6. (91%) The belief (asserting) (that/which) the terrorists who the UN denounced held (the hostage) was depressing to the negotiators.
7. (60%) The allegation (stating) (that/which) the senator who the army supported neglected (the danger) was leaked to the press.
8. (60%) The theory (claiming) (that/which) the chief who the lawyers advised defended (the tribe) was proven to be untrue.
9. (68%) The claim (alleging) (that/which) the cop who the mobster attacked ignored (the informant) might have affected the jury.
10. (55%) The discovery (revealing) (that/which) the physicist who the FBI funded publicized (the project) should be the top story.
11. (84%) The hope (suggesting) (that/which) the teenager who the teachers applauded encouraged (the boy) was shared by both parents.
12. (63%) The accusation (implying) (that/which) the guard who the warden questioned refused (the request) was absurd and completely unfounded.
13. (71%) The implication (stating) (that/which) the philosopher who the speaker cited refuted (his opponents) had to be considered illogical.
14. (46%) The threat (suggesting) (that/which) the men who the priest protected disregarded (the laws) might have scared the sheriff.
15. (63%) The rule (stating) (that/which) the kids who the teenager watched should follow (the clown) was part of a game.
16. (67%) The confirmation (indicating) (that/which) the parent who the teacher recruited made (the costumes) was sent to the PTA.
17. (54%) The acknowledgment (mentioning) (that/which) the fireman who the station contacted got (the truck) did not reach his partner.
18. (47%) The concern (stating) (that/which) the instructor who the students disliked raised (the grades) was discussed by the board.
19. (45%) The opinion (indicating) (that/which) the runner who the trainer rejected challenged (his opponent) was good for his reputation.
20. (73%) The assumption (implying) (that/which) the girl who the woman brought along made (the dress) had misled the sewing teacher.

Appendix C: Experiment 1c Materials

The three storage conditions and two additional integration conditions for item (1) of Experiment 1c can be derived as follows. The first verb that is provided is the obligatory ditransitive condition (e.g. “gave”). The second verb is from the optional ditransitive condition (e.g. “read”) and the third verb is from the obligatory transitive condition (e.g. “publish”). All three storage conditions included both of the RCs that followed the object NP. The obligatory transitive condition ended after these RCs. The other two storage conditions finished with the first option for the final segment of the sentence (e.g. nothing, as indicated by “0”). The three integration conditions all included the first verb, the obligatory ditransitive verb. The first of these conditions also functioned as one of the storage conditions. This version – the long integration condition – contains both RCs following the direct object. The medium integration condition contained only the first of the RCs following the direct object (e.g. only “which had impressed some critics”). This condition finished with the middle option for the final segment of the sentence (e.g. “who was very grateful for the present”). The short integration condition contained no RCs following the direct object, and finished with the final option for the final segment of the sentence.

Ditransitive completion biases listed as percentages are provided for each main verb in each condition. High percentages indicate a high bias to include a PP following the direct object NP. Plausibility ratings are also included for each verb, as a rating from 1 (unnatural) to 5 (natural). Items which were excluded from analyses because of either inappropriate completion biasing for the condition or unmatched plausibilities are marked with an asterisk.

1. Mary (gave (89%; 4.3) / read (0%; 4.6) / published (0%; 4.6)) a book (which had impressed some critics / who worked for a magazine) to a young child (0 / who was very grateful for the present / who was very grateful for the present and read it quickly).
2. Joe (placed (95%; 4.6) / moved (21%; 4.5) / broke (0%; 4.7)) the dishes (which had pleased the hostess / who catered for some friends) into the cabinet (above the counter / in the dining room / which had the beautiful stained glass doors).
3. Jane (put (100%; 4.5) / took (0%; 3.4) / hid (0%; 4.2)) the ball (which had struck the boy / who cried after the accident) into the closet (in the hallway / that contained all the other toys / after deciding the garage was an unsuitable place to store it).
4. Chris (handed (95%; 4.4) / mentioned (41%; 3.9) / copied (0%; 4.5)) the drawing (which had amazed the teacher / who studied at art school) to a student (of art history / who studied art history / who looked at it critically and decided that the perspective was poorly done).
5. Joe (positioned (60%; 4.2) / arranged (40%; 4.2) / held (42%; 3.9)) the pieces (which had fascinated the artist / who talked about wood work) onto the chess board (carefully / with care / after determining that he would play with the black pieces).
6. Sue (slung (55%; 4.3) / raised (45%; 4.1) / carried (16%; 4.7)) the bag (which had strained the grandma / who suffered from chronic pain) into the overhead compartment (0 / and closed the compartment securely / with surprising ease and grace).
7. Pete (handed (100%; 4.3) / returned (45%; 4.1) / found (20%; 4.7)) the binder (which had upset the engineer / who complained about budget cuts) to the clerk (who put it away / who took it without question / who took it without question or complaint and placed it on the shelf).
8. Margarite (gave (95%; 4.5) / took (32%; 3.9) / wanted (5%; 4.4)) the present (which had interested the girl / who played with the doll) to the toddler (who appreciated the gift / who smiled and opened the gift / who smiled and thanked her for the gift).
- *9. Joey (passed (35%; 4.2) / served (20%; 4.8) / ate (0%; 4.7)) the muffins (which had interested the child / who sat at the table) to the woman (who asked for them / who slowly sipped her tea / who slowly sipped her tea and admired the expensive china).
10. Marcie (placed (100%; 3.8) / raised (30%; 2.6) / dried (15%; 4.2)) the vest (which had saved the man / who fell into the water) into the storage area (0 / with the other life vests / which was located below the main deck of the ship).

11. Kate (put (100%; 5.0) / moved (32%; 4.3) / ate (0%; 4.8)) the food (which had quieted the children / who whined about being hungry) into the fridge (to save for later / on the middle shelf / to prevent it from going stale before tomorrow's lunch).
12. Nancy (put (100%; 4.4) / lowered (30%; 3.7) / purchased (16%; 4.3)) the cup (which had impressed the women / who drank the green tea) onto the saucer (to prevent it from spilling / after she had finished drinking her coffee / after she had finished drinking her coffee with cream).
13. De (handed (95%; 4.5) / recited (20%; 3.8) / reviewed (10%; 4.6)) the script (which had relieved the director / who worried about minute details) to the actor (who took it nonchalantly / who read his lines quickly / who immediately began to practice and memorize his lines).
14. Ed (passed (80%; 4.4) / recommended (32%; 3.5) / disliked (0%; 4.8)) the book (which had angered the protesters / who believed in other ideals) to the clerk (who read it quietly / who noted the title / who noted the title, author and copyright date).
- *15. Jan (bequeathed (60%; 3.8) / took (32%; 3.8) / admired (0%; 4.6)) the sculpture (which had inspired the author / who wrote about the artist) to the Museum of Fine Arts (to / who put the statue in storage / who placed the statue in the basement of their art gallery).
- *16. David (loaned (68%; 2.3) / lost (0%; 4.0) / pocketed (0%; 4.7)) the cash (which had helped the company / who expanded into new markets) to the casino owners (from Las Vegas / who used it to start other businesses / who used it to start other lucrative businesses of questionable moral value).
- *17. Eric (allocated (35%; 4.1) / mentioned (53%; 3.1) / borrowed (35%; 4.2)) some funds (which had pleased the executive / who spoke at the event) to a charity (for the homeless / in order to improve the company image / because it seemed the appropriate thing to do at the time).
18. Molly (crammed (80%; 4.6) / folded (0%; 4.1) / ignored (0%; 2.9)) the clothes (which had angered the roommate / who wished for a maid) into the drawer (which she then tried to shut / not caring that most of the clothes were dirty / without realizing that she would wrinkle her brand new skirt).
19. Eric (lodged (89%; 3.0) / lowered (25%; 3.4) / sharpened (0%; 4.7)) the knife (which had frightened the girl / who yelled about the noise) into the rack (which held the other kitchen utensils / with the other kitchen utensils / with the other kitchen utensils even though he knew that this was not a good arrangement).
20. Brenda (stuffed (85%; 4.1) / brought (32%; 3.4) / washed (0%; 4.6)) the towel (which had dried the children / who swam at the beach) into the closet (and shut the door / and shut the door quickly / and shut the door before the contents of the closet could overflow onto the floor).

Appendix D: Experiment 2 Materials

In the experimental items below, the two conditions of Experiment 2 can be derived as follows. For the center-embedded SO condition, choose item #a. For the right-branching OS condition, choose item #b. Plausibility is denoted in the parenthesis by “Y” for plausible and by “N” for implausible. For the first quarter of the items (1-36) both items are plausible, for the next quarter (37-72) they are both implausible, for the following quarter (73-108) the SO sentence is plausible while the OS sentence is implausible, and finally for the last quarter (109-144) the SO sentence is implausible while the OS sentence is plausible.

- 1a. (Y) The limerick that the boy recited appalled the priest.
- 1b. (Y) The boy recited the limerick that appalled the priest.
- 2a. (Y) The strike that the activist planned began the rebellion.
- 2b. (Y) The activist planned the strike that began the rebellion.
- 3a. (Y) The patient that the handcuff restrained bit the orderly.
- 3b. (Y) The handcuff restrained the patient that bit the orderly.
- 4a. (Y) The soldier that the bush concealed captured the enemy.
- 4b. (Y) The bush concealed the soldier that captured the enemy.
- 5a. (Y) The deputy that the newspaper identified chased the mugger.
- 5b. (Y) The newspaper identified the deputy that chased the mugger.
- 6a. (Y) The man that the show disturbed demanded the refund.
- 6b. (Y) The show disturbed the man that demanded the refund.
- 7a. (Y) The mural that the artist painted depicted the battle.
- 7b. (Y) The artist painted the mural that depicted the battle.
- 8a. (Y) The fire that the arsonist set destroyed the warehouse.
- 8b. (Y) The arsonist set the fire that destroyed the warehouse.
- 9a. (Y) The curfew that the mayor announced ended the riot.
- 9b. (Y) The mayor announced the curfew that ended the riot.
- 10a. (Y) The spice that the chef added flavored the soup.
- 10b. (Y) The chef added the spice that flavored the soup.
- 11a. (Y) The woman that the crime alarmed hired the bodyguard.
- 11b. (Y) The crime alarmed the woman that hired the bodyguard.
- 12a. (Y) The scenery that the woman admired inspired the painter.
- 12b. (Y) The woman admired the scenery that inspired the painter.
- 13a. (Y) The story that the biographer omitted insulted the queen.
- 13b. (Y) The biographer omitted the story that insulted the queen.
- 14a. (Y) The suspect that the evidence incriminated left the country.
- 14b. (Y) The evidence incriminated the suspect that left the country.
- 15a. (Y) The candidate that the newspaper endorsed lost the race.
- 15b. (Y) The newspaper endorsed the candidate that lost the race.
- 16a. (Y) The car that the mechanic owned needed the tune-up.
- 16b. (Y) The mechanic owned the car that needed the tune-up.
- 17a. (Y) The man that the sound deafened operated the jackhammer.
- 17b. (Y) The sound deafened the man that operated the jackhammer.
- 18a. (Y) The child that the book bored pestered the mother.
- 18b. (Y) The book bored the child that pestered the mother.
- 19a. (Y) The nail that the man hammered pierced the pipe.
- 19b. (Y) The man hammered the nail that pierced the pipe.
- 20a. (Y) The sailor that the lighthouse guided piloted the boat.
- 20b. (Y) The lighthouse guided the sailor that piloted the boat.
- 21a. (Y) The actress that the award thrilled praised the producer.
- 21b. (Y) The award thrilled the actress that praised the producer.
- 22a. (Y) The drug that the doctor discovered prevented the disease.

- 22b. (Y) The doctor discovered the drug that prevented the disease.
23a. (Y) The jockey that the bell alerted raced the horse.
23b. (Y) The bell alerted the jockey that raced the horse.
24a. (Y) The woman that the toothache annoyed saw the dentist.
24b. (Y) The toothache annoyed the woman that saw the dentist.
25a. (Y) The bride that the weather depressed scheduled the wedding.
25b. (Y) The weather depressed the bride that scheduled the wedding.
26a. (Y) The shack that the inspector condemned sheltered the drunk.
26b. (Y) The inspector condemned the shack that sheltered the drunk.
27a. (Y) The sculpture that the artist exhibited shocked the minister.
27b. (Y) The artist exhibited the sculpture that shocked the minister.
28a. (Y) The song that the father hummed soothed the child.
28b. (Y) The father hummed the song that soothed the child.
29a. (Y) The butcher that the knife cut summoned the doctor.
29b. (Y) The knife cut the butcher that summoned the doctor.
30a. (Y) The well that the man dug supplied the water.
30b. (Y) The man dug the well that supplied the water.
31a. (Y) The golfer that the lightning struck survived the incident.
31b. (Y) The lightning struck the golfer that survived the incident.
32a. (Y) The commuter that the construction delayed took the detour.
32b. (Y) The construction delayed the commuter that took the detour.
33a. (Y) The plane that the pilot flew transported the king.
33b. (Y) The pilot flew the plane that transported the king.
34a. (Y) The recession that the economist predicted worried the man.
34b. (Y) The economist predicted the recession that worried the man.
35a. (Y) The gun that the detective found wounded the victim.
35b. (Y) The detective found the gun that wounded the victim.
36a. (Y) The critic that the opera amazed wrote the review.
36b. (Y) The opera amazed the critic that wrote the review.
37a. (N) The wine that the hostess intoxicated accosted the guest.
37b. (N) The hostess accosted the guest that intoxicated the wine.
38a. (N) The tailor that the bow altered adorned the dress.
38b. (N) The bow adorned the dress that altered the tailor.
39a. (N) The ballet that the girl enchanted became the dancer.
39b. (N) The girl enchanted the dancer that became the ballet.
40a. (N) The man that the study commissioned analyzed the economy.
40b. (N) The study analyzed the economy that commissioned the man.
41a. (N) The bill that the activist angered organized the march.
41b. (N) The activist organized the march that angered the bill.
42a. (N) The verdict that the criminal astonished cursed the judge.
42b. (N) The criminal cursed the judge that astonished the verdict.
43a. (N) The rosebush that the bee attracted stung the gardener.
43b. (N) The bee stung the gardener that attracted the rosebush.
44a. (N) The heat that the man bothered installed the fan.
44b. (N) The man installed the fan that bothered the heat.
45a. (N) The storm that the man drenched bought the umbrella.
45b. (N) The man bought the umbrella that drenched the storm.
46a. (N) The video that the businessman showed bribed the senator.
46b. (N) The businessman showed the senator that bribed the video.
47a. (N) The plumber that the hair extracted clogged the sink.
47b. (N) The hair clogged the sink that extracted the plumber.
48a. (N) The janitor that the snow shoveled coated the sidewalk.
48b. (N) The snow coated the sidewalk that shoveled the janitor.
49a. (N) The woman that the envelope sealed contained the check.

- 49b. (N) The envelope contained the check that sealed the woman.
- 50a. (N) The woman that the cloth wove covered the table.
- 50b. (N) The cloth covered the table that wove the woman.
- 51a. (N) The tailor that the suit sewed delighted the customer.
- 51b. (N) The suit delighted the customer that sewed the tailor.
- 52a. (N) The mailman that the box delivered intrigued the employee.
- 52b. (N) The box intrigued the employee that delivered the mailman.
- 53a. (N) The towel that the athlete dried swam the channel.
- 53b. (N) The athlete swam the channel that dried the towel.
- 54a. (N) The wife that the secret revealed embarrassed the husband.
- 54b. (N) The secret embarrassed the husband that revealed the wife.
- 55a. (N) The magician that the trick performed entertained the girl.
- 55b. (N) The trick entertained the girl that performed the magician.
- 56a. (N) The book that the doctor mentioned eradicated the disease.
- 56b. (N) The doctor eradicated the disease that mentioned the book.
- 57a. (N) The butler that the shirt ironed fit the man.
- 57b. (N) The shirt ironed the man that fit the butler.
- 58a. (N) The actor that the performance gave humiliated the director.
- 58b. (N) The performance humiliated the director that gave the actor.
- 59a. (N) The heater that the woman warmed hemmed the skirt.
- 59b. (N) The woman hemmed the skirt that warmed the heater.
- 60a. (N) The weather that the diver hindered sought the treasure.
- 60b. (N) The diver sought the treasure that hindered the weather.
- 61a. (N) The thief that the gun loaded intimidated the victim.
- 61b. (N) The gun intimidated the victim that loaded the thief.
- 62a. (N) The mother that the sweater knitted tickled the child.
- 62b. (N) The sweater tickled the child that knitted the mother.
- 63a. (N) The paper that the actress libeled sued the publisher.
- 63b. (N) The actress sued the publisher that libeled the paper.
- 64a. (N) The haircut that the boy pleased tipped the barber.
- 64b. (N) The boy tipped the barber that pleased the haircut.
- 65a. (N) The scandal that the politician ruined quit the race.
- 65b. (N) The politician quit the race that ruined the scandal.
- 66a. (N) The martini that the businessman relaxed watched the show.
- 66b. (N) The businessman watched the show that relaxed the martini.
- 67a. (N) The legend that the knight described rescued the king.
- 67b. (N) The knight described the king that rescued the legend.
- 68a. (N) The alarm that the robber surprised seized the hostage.
- 68b. (N) The robber seized the hostage that surprised the alarm.
- 69a. (N) The cook that the knife sharpened sliced the meat.
- 69b. (N) The knife sliced the meat that sharpened the cook.
- 70a. (N) The child that the juice spilled stained the rug.
- 70b. (N) The juice stained the rug that spilled the child.
- 71a. (N) The clerk that the cash register used calculated the tax.
- 71b. (N) The cash register calculated the tax that used the clerk.
- 72a. (N) The comedian that the joke told upset the woman.
- 72b. (N) The joke told the woman that upset the comedian.
- 73a. (Y) The girl that the thorn pricked applied the bandaid.
- 73b. (N) The girl applied the bandaid that pricked the thorn.
- 74a. (Y) The child that the toy excited ate the popcorn.
- 74b. (N) The child ate the popcorn that excited the toy.
- 75a. (Y) The bear that the trap caught attacked the hunter.
- 75b. (N) The bear attacked the hunter that caught the trap.
- 76a. (Y) The cake that the boy baked impressed the girl.

- 76b. (N) The cake impressed the girl that baked the boy.
77a. (Y) The lawyer that the bar suspended betrayed the client.
77b. (N) The lawyer suspended the client that betrayed the bar.
78a. (Y) The key that the girl borrowed opened the door.
78b. (N) The key opened the door that borrowed the girl.
79a. (Y) The fence that the carpenter built surrounded the yard.
79b. (N) The fence surrounded the yard that built the carpenter.
80a. (Y) The wood that the man chopped heated the cabin.
80b. (N) The wood heated the cabin that chopped the man.
81a. (Y) The wolf that the campfire scared circled the area.
81b. (N) The wolf circled the area that scared the campfire.
82a. (Y) The precedent that the lawyer cited supported the case.
82b. (N) The precedent supported the case that cited the lawyer.
83a. (Y) The baby that the lullaby comforted sucked the pacifier.
83b. (N) The baby sucked the pacifier that comforted the lullaby.
84a. (Y) The pupil that the homework confused despised the tutor.
84b. (N) The pupil despised the tutor that confused the homework.
85a. (Y) The food that the caterer cooked poisoned the guest.
85b. (N) The food poisoned the guest that cooked the caterer.
86a. (Y) The patient that the drug cured thanked the doctor.
86b. (N) The patient thanked the doctor that cured the drug.
87a. (Y) The monument that the artist designed honored the hero.
87b. (N) The monument honored the hero that designed the artist.
88a. (Y) The device that the scientist invented detected the comet.
88b. (N) The device detected the comet that invented the scientist.
89a. (Y) The student that the grade discouraged dropped the course.
89b. (N) The student dropped the course that discouraged the grade.
90a. (Y) The food that the cook prepared displeased the heiress.
90b. (N) The food prepared the heiress that displeased the cook.
91a. (Y) The noise that the prisoner made distracted the guard.
91b. (N) The noise distracted the guard that made the prisoner.
92a. (Y) The secret that the aide divulged infuriated the banker.
92b. (N) The secret infuriated the banker that divulged the aide.
93a. (Y) The law that the worker disliked favored the millionaire.
93b. (N) The law disliked the millionaire that favored the worker.
94a. (Y) The dealer that the forgery fooled notified the policeman.
94b. (N) The dealer notified the policeman that fooled the forgery.
95a. (Y) The child that the movie frightened grabbed the usher.
95b. (N) The child grabbed the usher that frightened the movie.
96a. (Y) The detective that the crime frustrated uncovered the lead.
96b. (N) The detective uncovered the lead that frustrated the crime.
97a. (Y) The fort that the soldier guarded housed the ammunition.
97b. (N) The fort housed the ammunition that guarded the soldier.
98a. (Y) The tenant that the leak irritated harassed the landlord.
98b. (N) The tenant harassed the landlord that irritated the leak.
99a. (Y) The silo that the farmer repaired held the grain.
99b. (N) The silo held the grain that repaired the farmer.
100a. (Y) The coin that the man sold interested the collector.
100b. (N) The coin interested the collector that sold the man.
101a. (Y) The sheriff that the scandal involved investigated the crime.
101b. (N) The sheriff involved the crime that investigated the scandal.
102a. (Y) The foreman that the call interrupted managed the factory.
102b. (N) The foreman interrupted the factory that managed the call.
103a. (Y) The diplomat that the article discussed met the dictator.

- 103b. (N) The diplomat discussed the dictator that met the article.
104a. (Y) The woman that the comment offended slapped the man.
104b. (N) The woman slapped the man that offended the comment.
105a. (Y) The button that the maid pressed started the dishwasher.
105b. (N) The button started the dishwasher that pressed the maid.
106a. (Y) The person that the store rewarded recovered the jewelry.
106b. (N) The person recovered the jewelry that rewarded the store.
107a. (Y) The boat that the captain sailed won the race.
107b. (N) The boat won the race that sailed the captain.
108a. (Y) The novel that the writer submitted satisfied the editor.
108b. (N) The novel satisfied the editor that submitted the writer.
109a. (N) The millionaire that the money donated aided the orphan.
109b. (Y) The millionaire donated the money that aided the orphan.
110a. (N) The soldier that the barrier erected blocked the road.
110b. (Y) The soldier erected the barrier that blocked the road.
111a. (N) The pitcher that the ball threw broke the window.
111b. (Y) The pitcher threw the ball that broke the window.
112a. (N) The candy that the dieter tempted called the friend.
112b. (Y) The candy tempted the dieter that called the friend.
113a. (N) The cape that the bull provoked charged the matador.
113b. (Y) The cape provoked the bull that charged the matador.
114a. (N) The court that the investor indicted cheated the widow.
114b. (Y) The court indicted the investor that cheated the widow.
115a. (N) The actor that the scene rehearsed concluded the play.
115b. (Y) The actor rehearsed the scene that concluded the play.
116a. (N) The doctor that the test ordered confirmed the diagnosis.
116b. (Y) The doctor ordered the test that confirmed the diagnosis.
117a. (N) The computer that the man bewildered consulted the expert.
117b. (Y) The computer bewildered the man that consulted the expert.
118a. (N) The repairman that the machine fixed copied the paper.
118b. (Y) The repairman fixed the machine that copied the paper.
119a. (N) The waiter that the meal recommended disgusted the child.
119b. (Y) The waiter recommended the meal that disgusted the child.
120a. (N) The merchant that the candy displayed enticed the child.
120b. (Y) The merchant displayed the candy that enticed the child.
121a. (N) The man that the bill debated established the foundation.
121b. (Y) The man debated the bill that established the foundation.
122a. (N) The mother that the book read fascinated the child.
122b. (Y) The mother read the book that fascinated the child.
123a. (N) The boy that the garbage collected filled the can.
123b. (Y) The boy collected the garbage that filled the can.
124a. (N) The rain that the woman soaked hailed the cab.
124b. (Y) The rain soaked the woman that hailed the cab.
125a. (N) The clock that the student awoke hit the snooze button.
125b. (Y) The clock awoke the student that hit the snooze button.
126a. (N) The teenager that the miniskirt wore horrified the mother.
126b. (Y) The teenager wore the miniskirt that horrified the mother.
127a. (N) The nurse that the shot administered hurt the patient.
127b. (Y) The nurse administered the shot that hurt the patient.
128a. (N) The soldier that the weapon fired injured the baby.
128b. (Y) The soldier fired the weapon that injured the baby.
129a. (N) The armor that the knight protected killed the opponent.
129b. (Y) The armor protected the knight that killed the opponent.
130a. (N) The gardener that the hedge planted lined the driveway.

- 130b. (Y) The gardener planted the hedge that lined the driveway.
131a. (N) The hearing that the woman implicated married the man.
131b. (Y) The hearing implicated the woman that married the man.
132a. (N) The man that the salt sprinkled melted the ice.
132b. (Y) The man sprinkled the salt that melted the ice.
133a. (N) The review that the pupil benefited neglected the course.
133b. (Y) The review benefited the pupil that neglected the course.
134a. (N) The reminder that the man goaded paid the bill.
134b. (Y) The reminder goaded the man that paid the bill.
135a. (N) The water that the jogger cooled ran the marathon.
135b. (Y) The water cooled the jogger that ran the marathon.
136a. (N) The storekeeper that the package sent reached the customer.
136b. (Y) The storekeeper sent the package that reached the customer.
137a. (N) The bomb that the soldier blinded received the medal.
137b. (Y) The bomb blinded the soldier that received the medal.
138a. (N) The bullet that the gangster grazed stole the car.
138b. (Y) The bullet grazed the gangster that stole the car.
139a. (N) The school that the teacher employed taught the class.
139b. (Y) The school employed the teacher that taught the class.
140a. (N) The cowboy that the whip cracked terrified the horse.
140b. (Y) The cowboy cracked the whip that terrified the horse.
141a. (N) The fire that the wolf repelled threatened the child.
141b. (Y) The fire repelled the wolf that threatened the child.
142a. (N) The test that the teacher challenged tutored the student.
142b. (Y) The test challenged the teacher that tutored the student.
143a. (N) The sunlamp that the woman burned visited the doctor.
143b. (Y) The sunlamp burned the woman that visited the doctor.
144a. (N) The delay that the man inconvenienced wanted the ticket.
144b. (Y) The delay inconvenienced the man that wanted the ticket.

Appendix E: Experiment 3 Materials

The four conditions for item (1) of Experiment 3 can be derived as follows. For the ambiguous SC condition, always include “that” and the embedded object (e.g. “the award”) while for the unambiguous RC condition, include “which” instead of “that” and omit the object.

The number following each item number is the percent SC continuations in Study 2 of Kennison (2000), used because knowledge of Pearlmuter & Mendelsohn’s (1999) completion norming data in Experiment 1b was not available at the time.

1. (82%) The announcement (that/which) the baker from a small bakery in New York City received (the award) helped the business of the owner.
2. (82%) The announcement (that/which) the official from the labor union of the controversial company delivered (the contract) concerned the lawyers of the executives.
3. (82%) The announcement (that/which) the actress from the new movie about the dance industry signed (the autographs) excited the fans of ballroom dancing.
4. (82%) The announcement (that/which) the ranger with the inexperienced tourists in the blue tent received (the injury) worried the supervisor of the park.
5. (86%) The assumption (that/which) the reporter from the controversial magazine with the liberal bias made (the remark) disgruntled some members of the board.
6. (86%) The assumption (that/which) the guard at the new jail with the hardened criminals made (the mistake) caused many problems for the warden.
7. (86%) The assumption (that/which) the parents at the important meeting about the school performance made (the costumes) confused the director of the play.
8. (86%) The assumption (that/which) the colonel from the army base with the secret facilities made (the promise) misled the members of the community.
9. (82%) The belief (that/which) the philosopher from the historical period of the bronze age destroyed (the emperor) interested the class of eager students.
10. (82%) The belief (that/which) the terrorists in the nondescript building with many broken windows held (the hostage) concerned the authorities of Homeland Security.
11. (82%) The belief (that/which) the bishop at the international conference in a quiet town encouraged (the contributions) brought much hope to the community.
12. (81%) The claim (that/which) the advocate for the needy people in the impoverished neighborhood made (a difference) convinced the mayor of the city.
13. (81%) The claim (that/which) the reaction of the dangerous chemicals from the unmarked bottles disproved (the theory) confused the assistant of the researcher.
14. (81%) The claim (that/which) the newscaster from the major network with an unusual logo ridiculed (the politician) destroyed the reputation of the show.
15. (81%) The claim (that/which) the cop with the bushy mustache in the old photo ignored (the informant) influenced the decision of the jury.
16. (80%) The discovery (that/which) the friends of the teenage girl at the department store made (the commotion) surprised the shoppers and the salesmen.
17. (80%) The discovery (that/which) the spy from the secret agency of the national government fabricated (the data) infuriated the officers of the military.
18. (80%) The discovery (that/which) the archeologist from the famous expedition to the Egyptian pyramids ignored (the instructions) angered the sponsor of the project.
19. (80%) The discovery (that/which) the official from the biology lab at the secret facility publicized (the project) made the headlines of the newspaper.
20. (100%) The fact (that/which) the countries in the western hemisphere with the popular presidents ignored (the treaty) angered the environmentalists of the world.
21. (100%) The fact (that/which) the veteran from the European country in the World War recorded (some stories) aided the research of the historian.
22. (100%) The fact (that/which) the attorney on the strong defense of the innocent victim presented (the evidence) aided the decision of the jury.

23. (86%) The hope (that/which) the charity for the tragedy survivors from the impoverished town gave (the money) warmed the hearts of the families.
24. (86%) The hope (that/which) the fireman at the remote station in a dangerous district inspired (the teenagers) encouraged the parents of the delinquents.
25. (86%) The hope (that/which) the book about the complicated politics of the developing world gave (constructive criticisms) started serious discussions among the politicians.
26. (80%) The indication (that/which) the student with a Master's degree from a prestigious university received (the fellowship) preceded the awarding of the certificates.
27. (80%) The indication (that/which) the waiter at the Italian restaurant in the Main Square ignored (the request) upset the patron and the manager.
28. (80%) The indication (that/which) the hacker from a secret network in a European country sent (the virus) concerned the head of internet security.
29. (88%) The insistence (that/which) the chief from an Indian reserve with a prosperous casino maintained (the traditions) became a point of serious contention.
30. (82%) The misconception (that/which) the designer of the unique building in the arts district created (the memorial) confused the critics from the newspapers.
31. (82%) The misconception (that/which) the scientists from the large organization with lots of money developed (the toxins) alarmed the president of the country.
32. (82%) The misconception (that/which) the runner at the track meet for the state championship challenged (the decision) tarnished the reputation of the team.
33. (90%) The perception (that/which) the magician in the Vegas showcase at the impressive hotel created (the display) fascinated the crowd of eager tourists.
34. (90%) The perception (that/which) the photographer from the fashion magazine with a global audience ignored (the model) fascinated the readers of the article.
35. (100%) The possibility (that/which) the doctor from the cardiology department at the teaching hospital ignored (the warnings) shocked the head of the university.
36. (100%) The possibility (that/which) the boy with the basketball logo on the oversized shirt discussed (the secret) started an argument in the family.
37. (100%) The possibility (that/which) the instructor from the state university with the rich history raised (the grades) troubled the dean of the department.
38. (90%) The realization (that/which) the monk in the old monastery on the mountain top achieved (true enlightenment) delighted the members of the order.
39. (90%) The realization (that/which) the spy from the security agency of the secretive nation achieved (the objective) validated the years of hard work.
40. (90%) The realization (that/which) the kids at the birthday party in the beautiful park brought (the gifts) aroused the memories of happier times.
41. (86%) The requirement (that/which) the children from the expensive preschool in the rural suburbs should meet (the standards) satisfied the members of the PTA.
42. (86%) The requirement (that/which) the secretary at the software company with the famous CEO should pass (the test) satisfied the manager of the firm.
43. (86%) The requirement (that/which) the principal from the magnet school in the inner city must consider (the applicants) decreased the chances of early admissions.
44. (86%) The requirement (that/which) the prospector with the dangerous equipment for the gold mine should meet (the sheriff) ensured the safety of the townsfolk.
45. (84%) The revelation (that/which) the patient in the rehabilitation program for continued substance abuse shared (their happiness) lifted the spirit of the group.
46. (84%) The revelation (that/which) the clown on the circus tour from the Laughing Academy experienced (multiple seizures) upset the performers in the show.
47. (84%) The revelation (that/which) the victim in the emergency room at a London hospital experienced (serious shock) surprised the nurse at the desk.
48. (84%) The revelation (that/which) the executive from the large company with the clean record belittled (the secretary) surprised the employees in the office.
49. (90%) The rumor (that/which) the foreman at the successful bakery beside the grocery store overheard (the negotiations) shocked the stockholders of the company.

50. (90%) The rumor (that/which) the philanthropist from the wealthy family in the tiny town created (the problem) hurt the charity for needy families.
51. (90%) The rumor (that/which) the candidate in the long debate before the close election disputed (the bill) influenced the opinions of the voters.
52. (90%) The rumor (that/which) the senator in the special committee on the terrorist activities ignored (the warnings) ruined any possibility for a reelection.
53. (88%) The speculation (that/which) the politician from the important district in New York state questioned (the reports) infuriated the voters of the electorate.
54. (88%) The speculation (that/which) the daughter of the school headmaster in the conservative neighborhood offered (the bribe) shocked the elders of the community.
55. (82%) The suspicion (that/which) the girl on the speeding train to the college town developed (some homesickness) worried the mother and the father.
56. (82%) The suspicion (that/which) the man in the old building with the drug dealers shared (the money) confused the judge and the jury.
57. (82%) The suspicion (that/which) the woman at the resort town on the Mediterranean coast ignored (the warning) disconcerted the mother of the boyfriend.
58. (80%) The thought (that/which) the mother of the newborn child at the suburban clinic ignored (the advice) worried the nurse and the doctor.
59. (80%) The thought (that/which) the patient with symptoms of depression at the new clinic suppressed (the rage) frightened the nurse of the psychiatrist.
60. (80%) The thought (that/which) the man in the support group from the small town shared (his feelings) encouraged the members with similar problems.

Appendix F: Experiment 4 Materials

In the experimental items below, the four conditions for item (1) of Experiment 4 can be derived as follows. For the plausible subject noun, ambiguous structure condition, choose the first NP of the matrix subject (e.g. “the demolitionist”) and omit the next words (e.g. “who was” and “that was”). For the plausible subject noun, unambiguous structure condition, choose the first NP again and the first form of the next two words (e.g. “who was”). For the implausible subject noun, ambiguous structure condition, choose the second NP of the matrix subject (e.g. “the building”) and omit the next words. For the implausible subject noun, unambiguous structure condition, choose the second NP again and the second form of the next two words (e.g. “that was”).

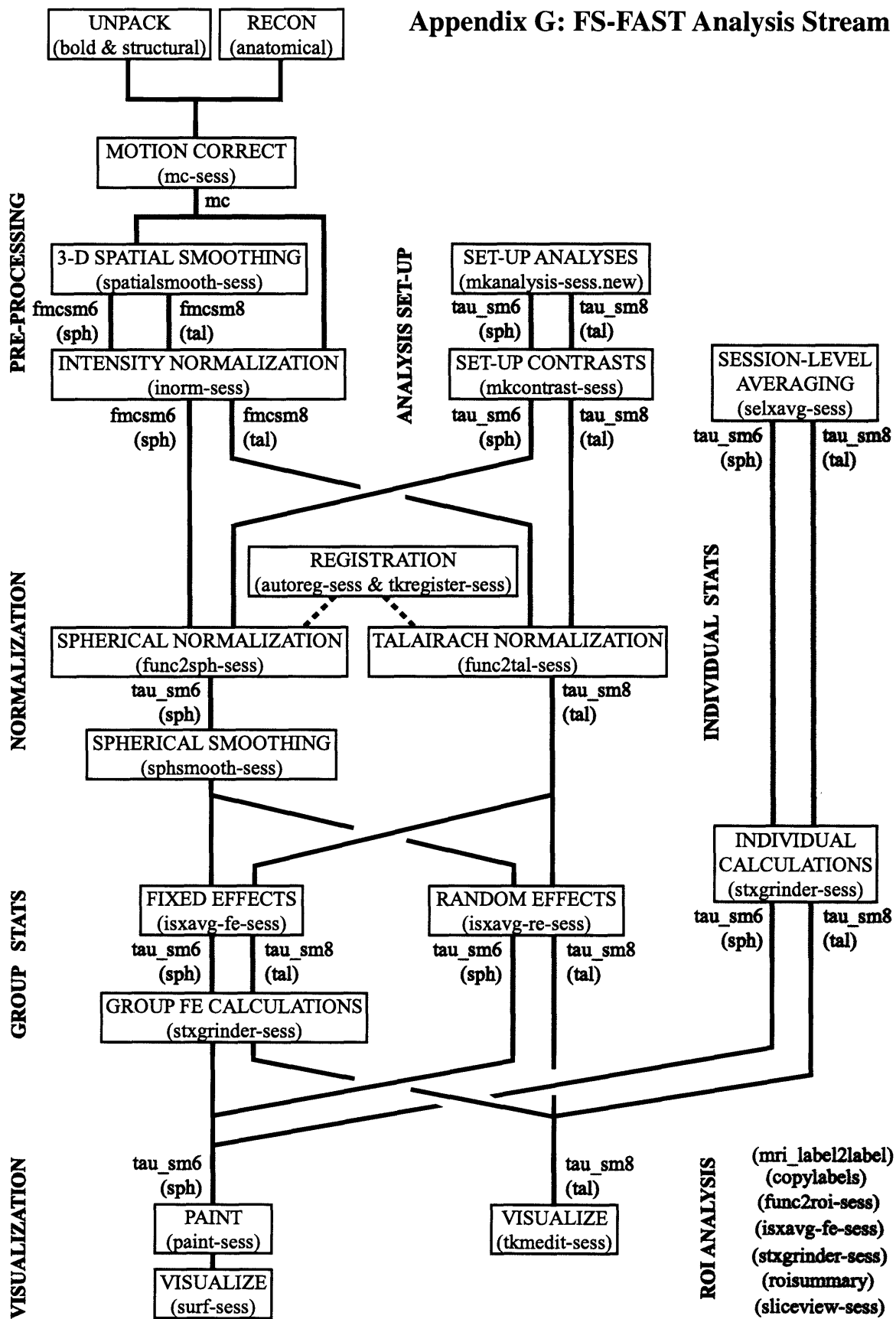
1. (The demolitionist/ the building) (who was / that was) blasted by the explosives was old.
2. (The housemaid/ the foyer) (who was / that was) cleaned by the water was filthy.
3. (The surgeon/ the cheese) (who was / that was) cut by the knife fell down.
4. (The waitress/ the flatbed) (who was / that was) delivered by the service was helpful.
5. (The searchers/ the glasses) (who were / that was) found by the policeman were missing.
6. (The sergeant/ the snack) (who was / that was) ordered by the general disappeared quickly.
7. (The strategist/ the vase) (who was / that was) positioned by the president complimented the cabinet.
8. (The scientist/ the discovery) (who was / that was) researched by the reporter was famous.
9. (The physician/ the tumor) (who was / that was) scanned by the x-ray machine was worse.
10. (The servant/ the feast) (who was / that was) served by the maids was cold.
11. (The huntsman/ the marshmallow) (who was / that was) speared by the stake was burnt.
12. (The courier/ the treasure) (who was / that was) transported by the armored car was safe.
13. (The applicant/ the trophy) (who was / that was) accepted by the college was outstanding.
14. (The puppy/ the retirement) (that was / that was) announced by the host surprised everyone.
15. (The vice/ the knob) (that was / that was) clasped by the child was cold.
16. (The sword/ the cheese) (that was / that was) cut by the metalsmith looked incredible.
17. (The defendant/ the evidence) (who was / that was) examined by the lawyer was unreliable.
18. (The prisoner/ the gold) (who was / that was) transported by the guards was secured.
19. (The teacher/ the textbook) (who was / that was) loved by the class was understandable.
20. (The workers/ the bricks) (who were / that was) lifted by the crane were heavy.
21. (The student/ the paper) (who was / that was) graded by the professor received an F.
22. (The contestant/ the recipe) (who was / that was) selected by the judges was favored.
23. (The specialist/ the equipment) (who was / that was) requested by the hospital finally arrived.
24. (The thief/ the jewelry) (who was / that was) identified by the victim was detained.
25. (The soldier/ the valley) (who was / that was) captured by the enemy was guarded.
26. (The troops/ the power plant) (who were / that was) attacked by the terrorists suffered losses.
27. (The artist/ the painting) (who was / that was) studied by the historian was obscure.
28. (The boy/ the necklace) (who was / that was) described by the lady was handsome.
29. (The mailman/ the package) (who was / that was) expected by the secretary arrived late.
30. (The woman/ the sofa) (who was / that was) scratched by the cat was injured.
31. (The man/ the van) (who was / that was) recognized by the spy fled suddenly.
32. (The client/ the account) (who was / that was) wanted by the advertiser was worth millions.

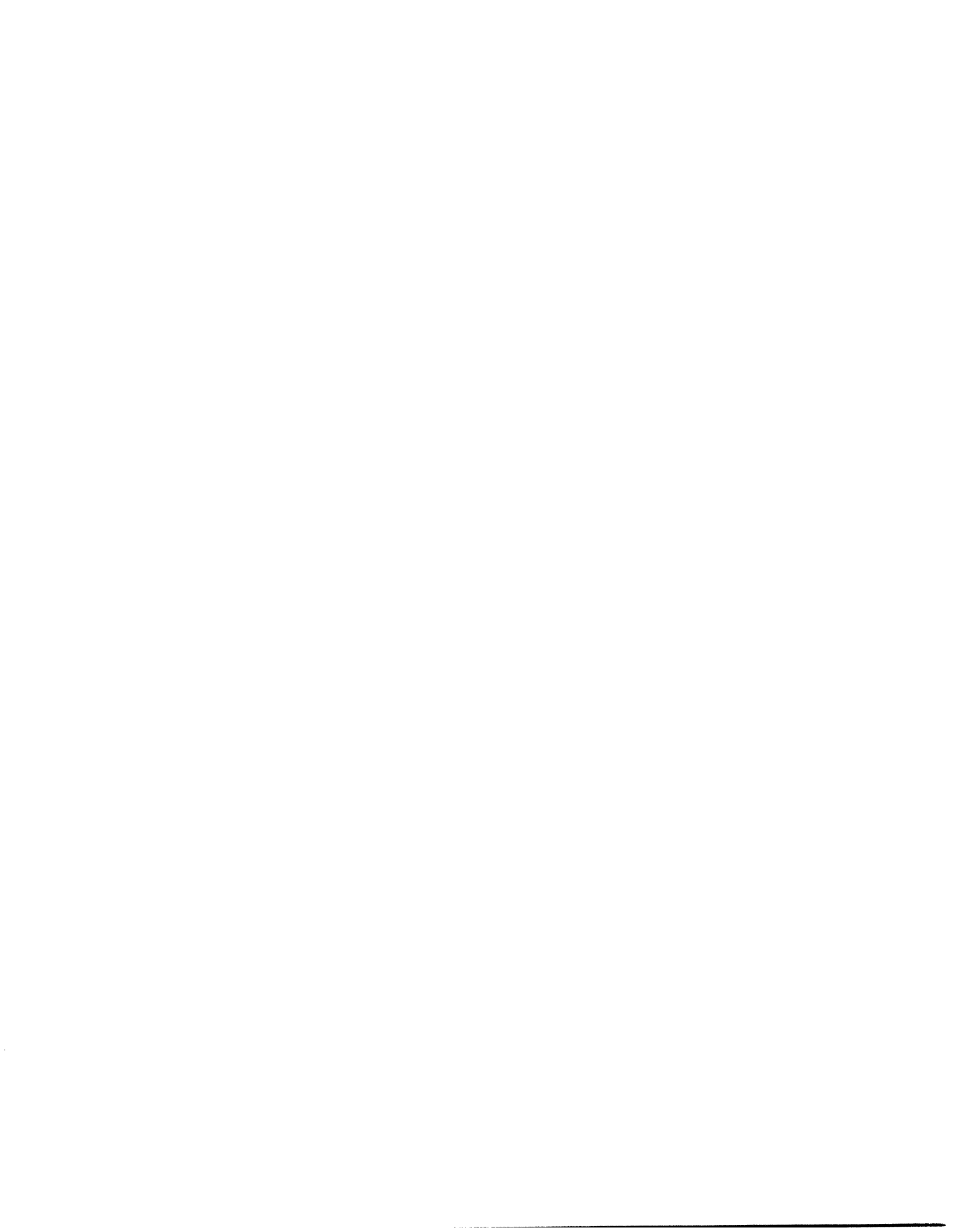
The Main Verb foil condition items were presented as listed below.

1. The demolitionist blasted the old building to rubble.
2. The housemaid cleaned the foyer after the storm.
3. The surgeon cut the cheese in the kitchen.
4. The waitress delivered the food to the customer.
5. The searchers found the glasses on the shelf.

6. The sergeant ordered the snack before the drills.
7. The strategist positioned the vase perfectly on the shelf.
8. The scientist researched the discovery using new technologies.
9. The physician scanned the tumor to check its growth.
10. The servant served the feast with style.
11. The huntsman speared the marshmallow before roasting it.
12. The courier transported the treasure by armored car.
13. The applicant accepted the trophy in a special ceremony.
14. The puppy announced his presence with a bark.
15. The vice clasped the stuck knob and opened it.
16. The sword cut the cheese in one stroke.
17. The defendant examined the evidence while under oath.
18. The prisoner transported the gold to nearby caves.
19. The teacher loved the textbook because of its clarity.
20. The workers lifted the bricks slowly and carefully.
21. The student graded the paper for the professor.
22. The contestant selected the delicious recipe for the competition.
23. The specialist requested the equipment for the surgery.
24. The thief identified the jewelry as extremely valuable.
25. The soldier captured the valley with his troops.
26. The troops attacked the power plant at night.
27. The artist studied the painting for some inspiration.
28. The boy described the necklace to the store manager.
29. The mailman expected the package to be heavier.
30. The woman scratched the sofa accidentally while moving.
31. The man recognized the van from the robbery.
32. The client wanted the account to boost business.

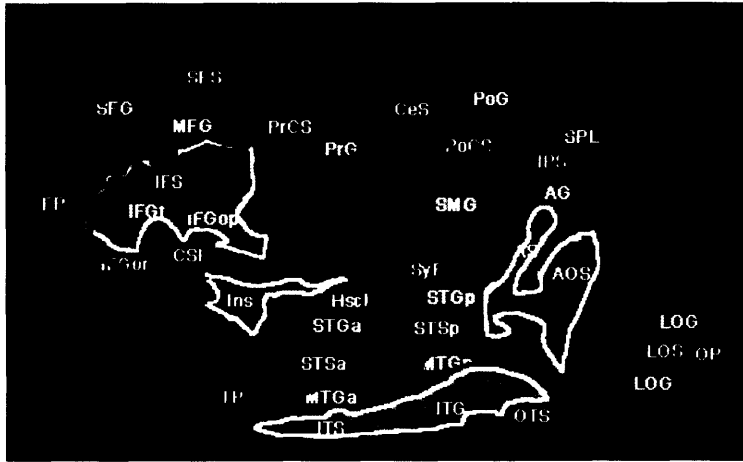
Appendix G: FS-FAST Analysis Stream



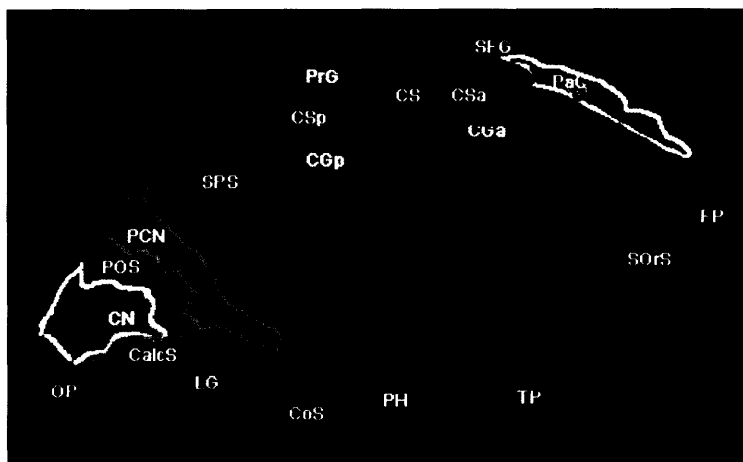


Appendix H: Region of Interest Definitions

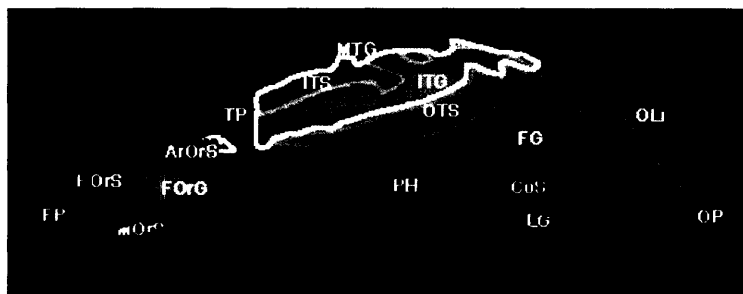
Lateral



Medial



Ventral

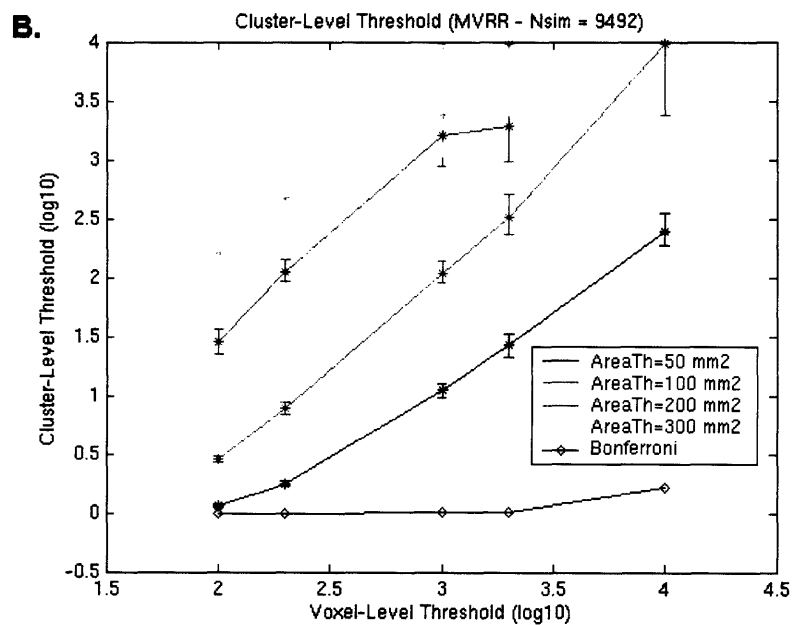
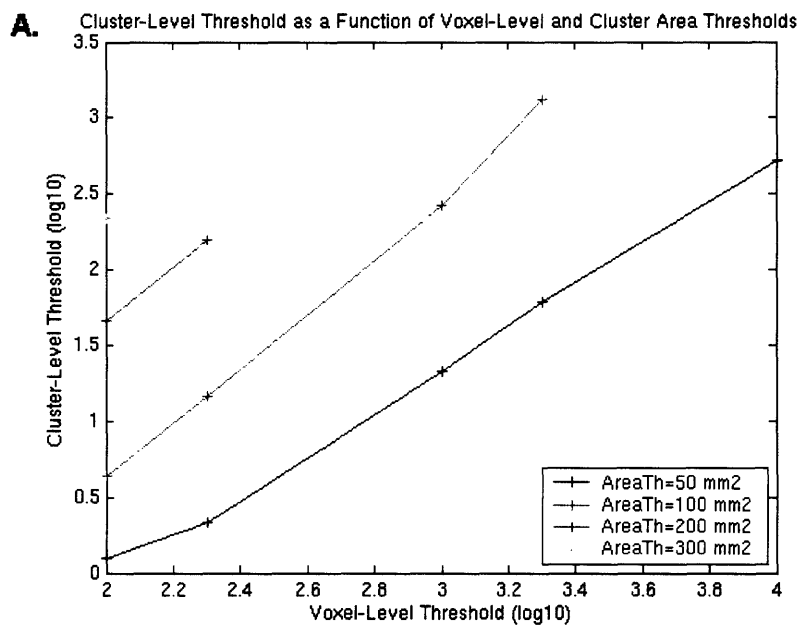


- AG: angular gyrus
- AOS: anterior occipital sulcus
- ArOrS: arcuate orbital sulcus
- AS: angular sulcus
- CalcS: calcarine sulcus
- CeS: central sulcus
- CG: cingulate gyrus
- CGa: anterior cingulate gyrus
- CGp: posterior cingulate gyrus
- CN: cuneus
- CoS: collateral sulcus
- CS: cingulate sulcus
- CSa: anterior cingulate sulcus
- CSp: posterior cingulate sulcus
- CSI: circular sulcus of insula
- FG: fusiform gyrus
- ForG: frontoorbital gyrus
- FP: frontal pole
- HscI: Heschl's gyrus
- HorS: H-shaped orbital sulcus
- IFGop: inferior frontal gyrus pars oper
- IFGtri: inferior frontal gyrus pars tri
- IFS: inferior frontal sulcus
- Ins: insula
- IPL: inferior parietal lobule
- IPS: intraparietal sulcus
- ITG: inferior temporal gyrus
- ITS: inferior temporal sulcus
- Lateral prefrontal cortex: * MFG+ SFS
- LG: lingual gyrus
- LOG: lateral occipital gyrus
- LOS: lateral occipital sulcus
- MFG: middle frontal gyrus
- MorS: medial orbital sulcus
- MTG: middle temporal gyrus
- OLi: inferior occipital lateral gyrus
- OP: occipital pole
- OTS: occipitotemporal sulcus
- PAC: paracingulate cortex
- PCN: precuneus
- PH: parahippocampal gyrus
- PoCS: postcentral sulcus
- PoG: postcentral gyrus
- POS: parieto-occipital sulcus
- PrCS: precentral sulcus
- PrG: precentral gyrus
- SFG: superior frontal gyrus
- SFS: superior frontal sulcus
- SOrS: superior orbital sulcus
- SPL: superior parietal lobule
- SMG: supramarginal gyrus
- SPS: subparietal sulcus
- STG: superior temporal gyrus
- STS: superior temporal sulcus
- SyF: sylvian fissure
- TP: temporal pole



Appendix I: Cluster Analysis Simulation Results

Results of cluster analysis simulations for Experiment 2 (Figure A) and Experiment 4 (Figure B). For each minimum area (colored lines) the x-axis value represents the minimum vertex-level p-value (\log_{10}) that is considered significantly active for the cluster and the y-axis is the resulting cluster-level p-value (\log_{10}).





Appendix J: Experiment 2 FS-FAST Commands

ACTIVATION MAPS

Motion Correction:

```
mc-sess -method afni -toff 0 -sf subjects-id-final12 -df dir -umask 2
```

3-D Spatial Smoothing (empirical ideal for spherical):

```
spatialsMOOTH-sess -i fmc -o fmcsM6 -fwhm 6 -sf subjects-id-final12 -df dir
```

Intensity Normalization:

```
inorm-sess -umask 2 -sf subjects-id-final12 -df dir
```

Set-up Analysis (for spherical):

```
mkanalysis-sess.new -analysis tau_sm6 -paradigm cerb.par -designType event-related -motioncor -inorm -force  
-sf subjects-id-final12 -df dir -nconditions 4 -prestim 4 -TR 2 -TER 2 -timewindow 26 -taumax 30 -funcstem  
fmcsM6
```

Set-up Contrasts:

```
#Clause Type Plausibility
```

```
#-----
```

```
#CE (y)es
```

```
#RB (n)o
```

```
#2 x 2 = 4
```

```
#
```

```
#Condition ID Code
```

```
#-----
```

```
#0 fixation
```

```
#1 center-embedded, plausible (cey)
```

```
#2 center-embedded, implausible (cen)
```

```
#3 right-branching, plausible (rby)
```

```
#4 right-branching, implausible (rbn)
```

```
#
```

```
#Delays
```

```
#-----
```

```
# 4-10: 0 0 0 0 1 1 1 1 0 0 0 0
```

```
#
```

```
#For spherical (cerb.par; 4-10sum)
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -a 2 -a 3 -a 4 -c 0 -contrast  
all_fix_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -a 2 -c 0 -contrast ce_fix_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 3 -a 4 -c 0 -contrast rb_fix_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 2 -a 4 -c 0 -contrast n_fix_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -a 3 -c 0 -contrast y_fix_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -a 2 -c 3 -c 4 -contrast ce_rb_4-  
10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 2 -a 4 -c 1 -c 3 -contrast n_y_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -c 3 -contrast cey_rby_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 2 -c 4 -contrast cen_rbn_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 2 -c 1 -contrast cen_cey_4-10sum
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 4 -c 3 -contrast rbn_rby_4-10sum
```

Registration:

```
autoreg-sess -sf subjects-id-final12 -df dir
```

```
tkregister-sess -s 000000AA -df dir
```

Spherical Normalization:

```
func2sph-sess -analysis tau_sm6 -sf subjects-id-final12 -df dir
```

Spherical Smoothing:

```
sphsmooth-sess -smoothsteps 20 -analysis tau_sm6 -insphdir sph -outsphdir sphsm_20 -sf subjects-id-final12 -df dir
```

Session-level Averaging:

```
selxavg-sess -analysis tau_sm6 -sf subjects-id-final12 -df dir
```

Individual Calculations (for each contrast and participant):

```
stxgrinder-sess -contrast ce_rb_4-10sum -analysis tau_sm6 -spacedir sphsm_20 -space sph -umask 0 -s 000000AA -df dir
```

Random Effects Analysis (for each contrast):

```
isxavg-re-sess -contrast ce_rb_4-10sum -group isxavg-final12 -analysis tau_sm6 -spacedir sphsm_20 -space sph -sf subjects-id-final12 -df dir
```

Paint (for each contrast):

```
paint-sess -isxavg random -map minsig -contrast ce_rb_4-10sum -analysis tau_sm6 -spacedir sphsm_20 -space sph -subject margaret -s isxavg-final12 -df dir
```

Surf (for contrast of interest):

```
# pthresh = 10^-(fthresh)
```

```
# pmax = 10^-(fmid+(1/fslope))
```

```
#
```

```
surf-sess -contrast ce_rb_4-10sum -map minsig -offset 0 -hemi lh -analysis tau_sm6 -spacedir sphsm_20 -space sph -fthresh 1.3 -fmid 2.0 -fslope 1.0 -smoothsteps 0 -subject margaret -avgcurv -isxavg random -s isxavg-final12 -df dir
```

CLUSTER ANALYSIS

Cluster Identification (for each contrast and hemisphere):

```
# cluster threshold 1.3; vertex threshold 2.38; minarea 100 (from plot)
```

```
#
```

```
mri_surfcluster --hemi lh --src minsig-0-margaret-lh.w --srcsubj margaret --thmin 2.38 --thmax infinity --thsign pos --minarea 100 --sum ce_rb_4-10sum_clth13_thmin24_minarea100_lh --o ./ce_rb_4-10sum_clth13_thmin24_minarea100_lh
```

ROI ANALYSIS

Individual Participant Label Transformation (for each participant, hemisphere & label; functional mask of $p < 0.01$):

```
mri_label2label --srcsubject margaret --regmethod surface --hemi lh --srclabel $SUBJECTS_DIR/margaret/label/ L_inf_front_cortex_020327_avgcurv.label --trglabel L_inf_front_cortex_020327_gr2_4-10sum --trgsubject west_000000AAa --srcmask /space/invaders/4/users/caroline/cerb/isxavg-final12/bold/tau_sm6/sphsm_20-rfx/ce_rb_4-10sum/sig-0-margaret-lh.w 2 paint
```

Copy Labels to Participant Folders (for each participant):

```
cp $SUBJECTS_DIR/west_000000AAa/label/*.label ~/subjects/000000AA/labels/
```

Compile ROI Values (for each participant, hemisphere & label):

```
func2roi-sess -s 000000AA -df dir -roidef L_inf_front_cortex_020327_gr2_4-10sum -sesslabel L_inf_front_cortex_020327_gr2_4-10sum -analysis tau_sm6
```


Fixed Effect Analysis (for each hemisphere & label):

```
isxavg-fe-sess -analysis tau_sm6 -group isxavg-final12 -sf subjects-id-final12 -df dir -space  
L_inf_front_cortex_020327_gr2_4-10sum  
stxgrinder-sess -analysis tau_sm6 -s isxavg-final12 -df dir -contrast omnibus -space  
L_inf_front_cortex_020327_gr2_4-10sum
```

Output Summary (for each hemisphere & label):

```
roisummary-sess -sumfile L_inf_front_cortex_020327_gr2_4-10sum .dat -roidef  
L_inf_front_cortex_020327_gr2_4-10sum -analysis tau_sm6 -sf subjects-id-final12 -transpose -df dir
```

Visualization (for label of interest):

```
sliceview-sess -slice mos -analysis tau_sm6 -contrast ce_rb_4-10sum -map sig -s isxavg-final12 -df dir -space  
L_inf_front_cortex_020327_gr2_4-10sum -isxavg fixed
```


Appendix K: Experiment 4 FS-FAST Commands

ACTIVATION MAPS

Motion Correction:

```
mc-sess -method afni -toff 0 -sf subjects-20 -df dir -umask 2
```

3-D Spatial Smoothing (empirical ideal for spherical):

```
spatialsmooth-sess -i fmc -o fmcsm6 -fwhm 6 -sf subjects-20 -df dir
```

Intensity Normalization:

```
inorm-sess -umask 2 -sf subjects-20 -df dir
```

Set-up Analysis (for spherical):

```
mkanalysis-sess.new -analysis tau_sm6 -paradigm mvrr.par -designtype event-related -motioncor -inorm -force  
-sf subjects-20 -df dir -nconditions 5 -prestim 8 -TR 2.0000 -TER 2 -timewindow 30 -taumax 30 -funcstem  
fmcsm6
```

Set-up Contrasts:

```
#MVRR
```

```
#
```

```
#animacy    ambiguity
```

```
#-----
```

```
#An        Amb
```

```
#In        Unamb
```

```
#2    x    2    =    4
```

```
#
```

```
#
```

```
#Condition  ID Code
```

```
#-----
```

```
#0        fix
```

```
#1        An-Amb
```

```
#2        An-Unamb
```

```
#3        In-Amb
```

```
#4        In-Unamb
```

```
#5        MV control
```

```
#
```

```
#Delays
```

```
#-----
```

```
#6-10: 0 0 0 0 0 0 1 1 1 0 0 0 0 0
```

```
#
```

```
#For spherical (mvrr.par; 6-10sum; rmprestim)
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -a 2 -c 3 -c 4 -contrast An_In_6-  
10sum_rmp
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -a 3 -c 2 -c 4 -contrast  
Amb_Unamb_6-10sum_rmp
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -c 2 -contrast
```

```
AnAmb_AnUnamb_6-10sum_rmp
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 3 -c 4 -contrast InAmb_InUnamb_6-  
10sum_rmp
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -c 4 -contrast AnAmb_InUnamb_6-  
10sum_rmp
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 2 -c 3 -contrast AnUnamb_InAmb_6-  
10sum_rmp
```

```
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 1 -c 5 -contrast AnAmb_Mvfoil_6-10sum_rmp
mkcontrast-sess -rmprestim -analysis tau_sm6 -setwdelay -sumdelays -a 2 -c 5 -contrast AnUnamb_Mvfoil_6-10sum_rmp
```

Registration:

```
autoreg-sess -sf subjects-20 -df dir
tkregister-sess -s 000000AA -df dir
```

Spherical Normalization:

```
func2sph-sess -analysis tau_sm6 -sf subjects-20 -df dir
```

Spherical Smoothing:

```
sphsmooth-sess -smoothsteps 20 -analysis tau_sm6 -insphdir sph -outsphdir sphsm_20 -sf subjects-20 -df dir
```

Session-level Averaging:

```
selxavg-sess -analysis tau_sm6 -sf subjects-20 -df dir
```

Individual Calculations (for each contrast and participant):

```
stxgrinder-sess -contrast an_in_6-10sum_rmp -analysis tau_sm6 -spacedir sphsm_20 -space sph -umask 0 -s 000000AA -df dir
```

Random Effects Analysis (for each contrast):

```
isxavg-re-sess -contrast an_in_6-10sum_rmp -group isxavg-20subj -analysis tau_sm6 -spacedir sphsm_20 -space sph -sf subjects-20 -df dir
```

Paint (for each contrast):

```
paint-sess -isxavg random -map minsig -contrast an_in_6-10sum_rmp -analysis tau_sm6 -spacedir sphsm_20 -space sph -subject margaret -s isxavg-20subj -df dir
```

Surf (for contrast of interest):

```
# pthresh = 10^-(fthresh)
# pmax = 10^-(fmid+(1/fslope))
#
surf-sess -contrast an_in_6-10sum_rmp -map minsig -offset 0 -hemi lh -analysis tau_sm6 -spacedir sphsm_20 -space sph -fthresh 1.3 -fmid 2.0 -fslope 1.0 -smoothsteps 0 -subject margaret -avgcurv -isxavg random -s isxavg-20subj -df dir
```

CLUSTER ANALYSIS

Cluster Identification (for each contrast and hemisphere):

```
# cluster threshold 1.3; vertex threshold 2.38; minarea 100 (from plot)
#
mri_surfcluster --hemi lh --src minsig-0-margaret-lh.w --srcsubj margaret --thmin 1.9 --thmax infinity --thsign pos --minarea 200 --sum an_in_6-10sum_rmp_clth13_thmin19_minarea200_lh --o ./ an_in_6-10sum_rmp_clth13_thmin19_minarea200_lh
```

ROI ANALYSIS

Individual Participant Label Transformation (for each participant, hemisphere & label; functional mask of $p < 0.01$):

```
mri_label2label --srcsubject margaret --regmethod surface --hemi lh --srclabel
$SUBJECTS_DIR/margaret/label/ L_inf_front_cortex_020327_avgcurv.label --trglabel
L_inf_front_cortex_020327_gr2_6-10sum --trgs subject west_000000AAa --srcmask
/space/evry/4/users/caroline/mvrr/isxavg-20subj/bold/tau_sm6/sphsm_20-rfx/an_in_6-10sum_rmp/sig-0-
margaret-lh.w 2 paint
```

Copy Labels to Participant Folders (for each participant):

```
cp $SUBJECTS_DIR/west_000000AAa/label/*.label ~/subjects/000000AA/labels/
```

Compile ROI Values (for each participant, hemisphere & label):

```
func2roi-sess -s 000000AA -df dir -roidef L_inf_front_cortex_020327_gr2_6-10sum -sesslabel  
L_inf_front_cortex_020327_gr2_6-10sum -analysis tau_sm6
```

Fixed Effect Analysis (for each hemisphere & label):

```
isxavg-fe-sess -analysis tau_sm6 -group isxavg-20subj -sf subjects-20 -df dir -space  
L_inf_front_cortex_020327_gr2_6-10sum  
stxgrinder-sess -analysis tau_sm6 -s isxavg-20subj -df dir -contrast omnibus -space  
L_inf_front_cortex_020327_gr2_6-10sum
```

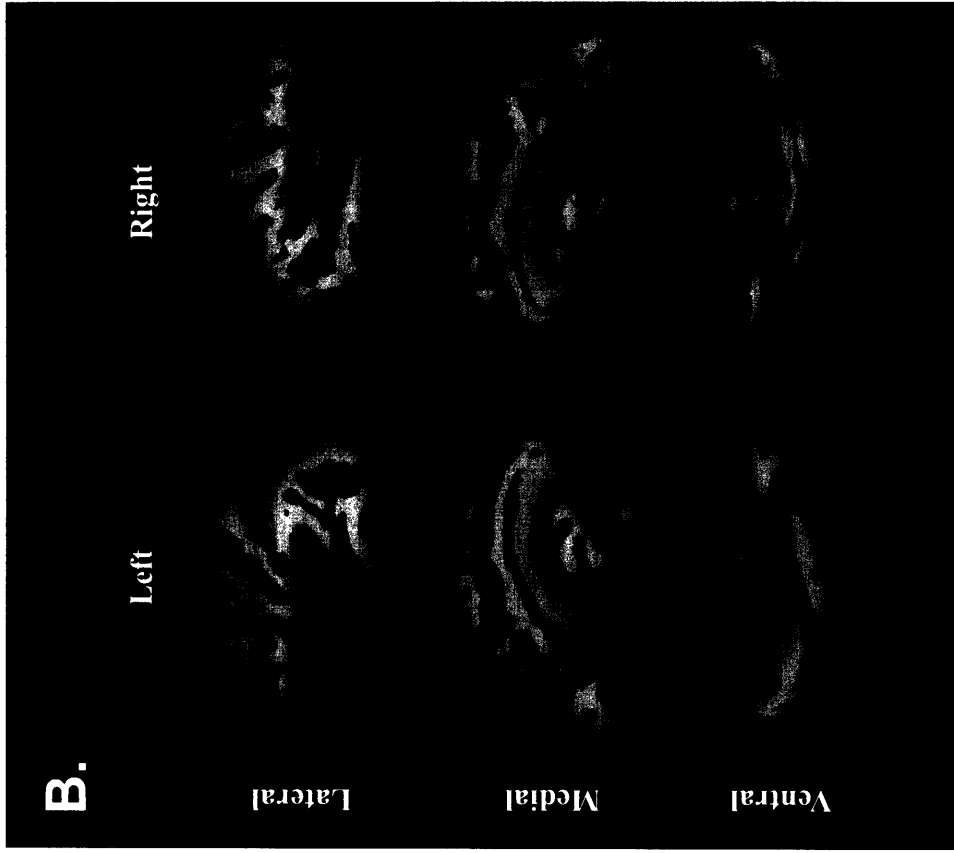
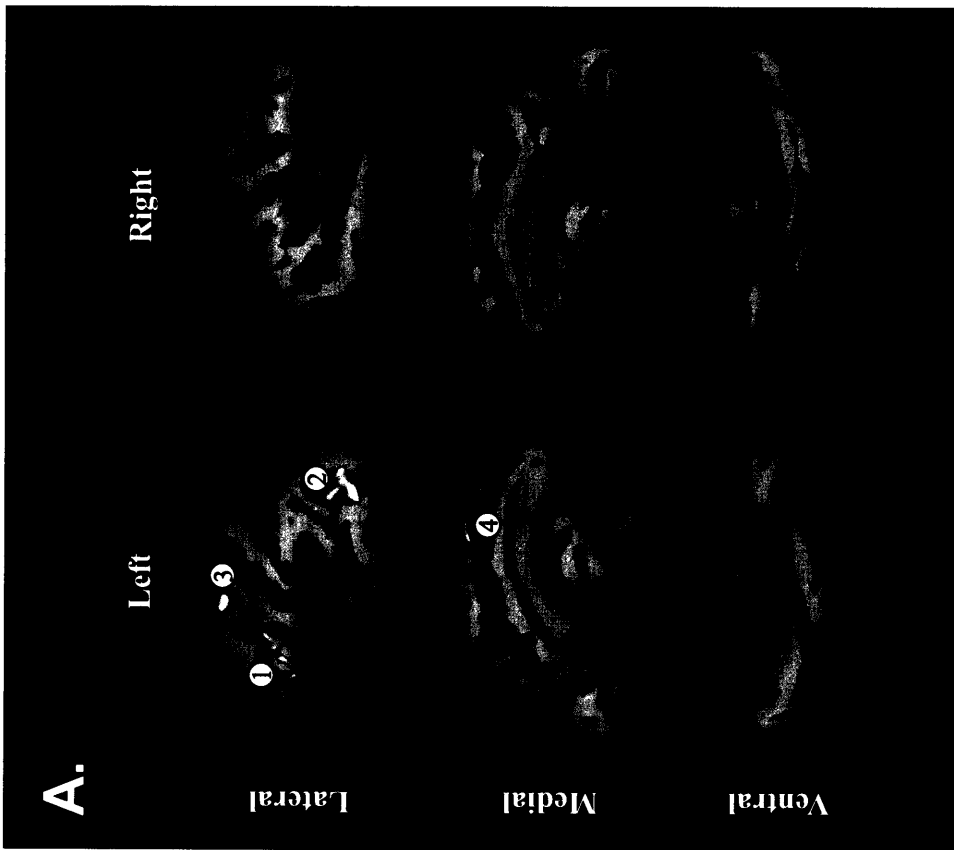
Output Summary (for each hemisphere & label):

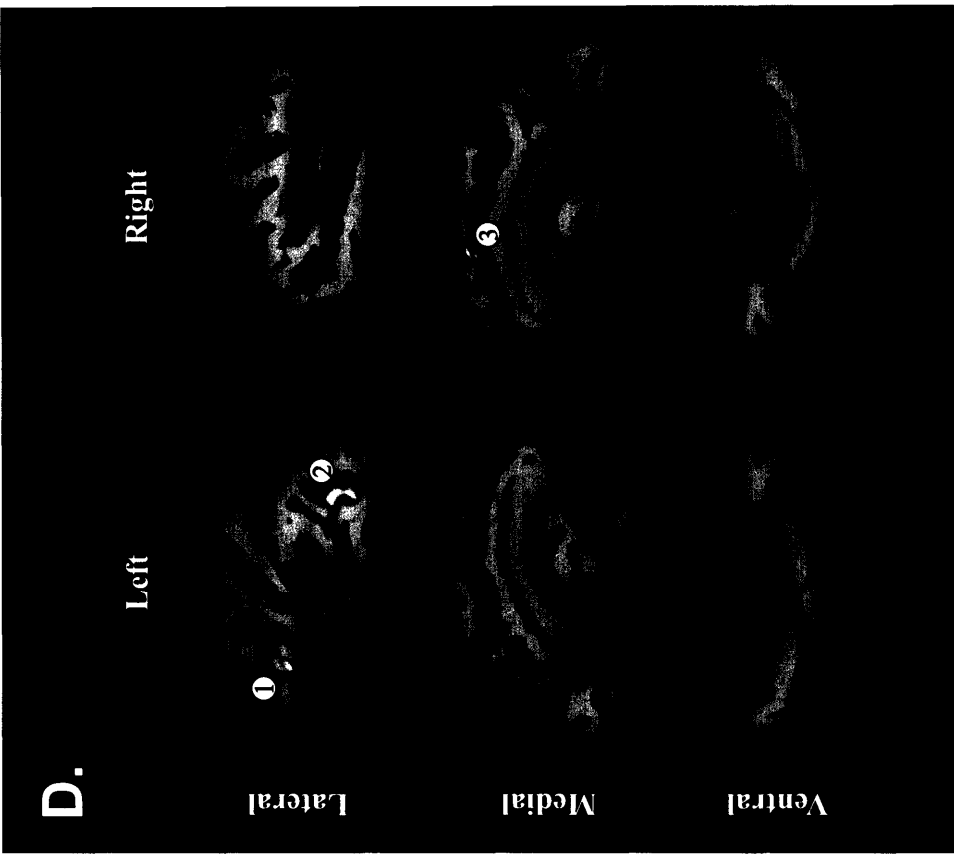
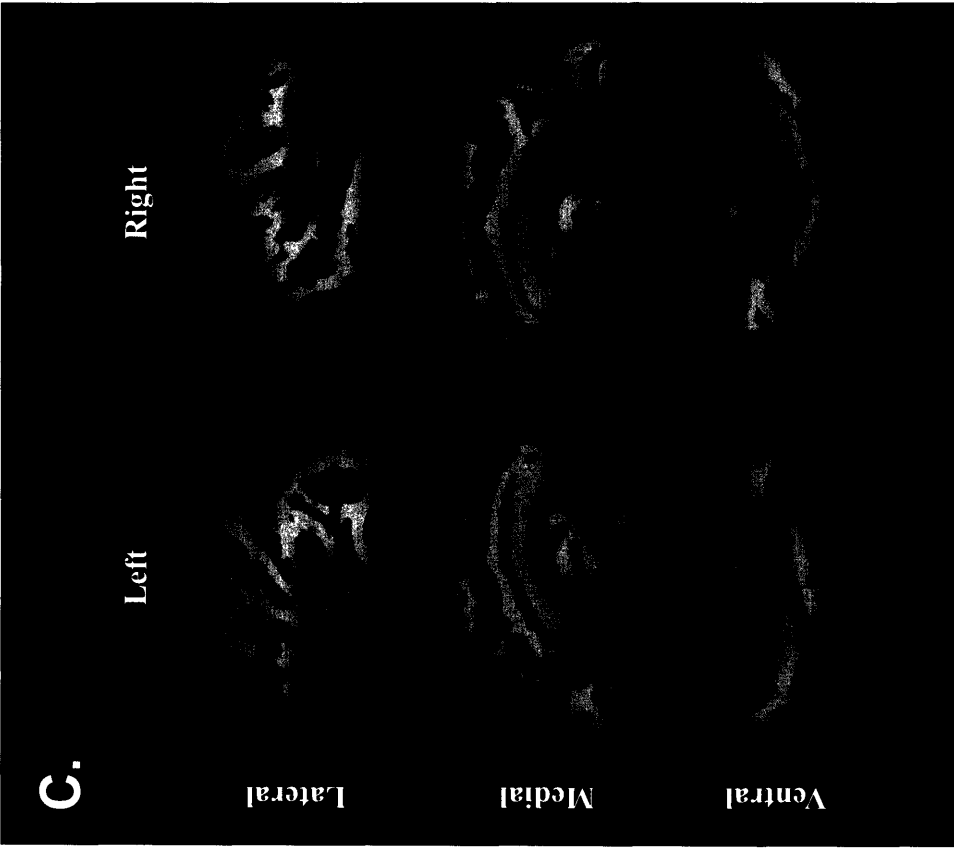
```
roisummary-sess -sumfile L_inf_front_cortex_020327_gr2_6-10sum .dat -roidef  
L_inf_front_cortex_020327_gr2_6-10sum -analysis tau_sm6 -sf subjects-20 -transpose -df dir
```

Visualization (for label of interest):

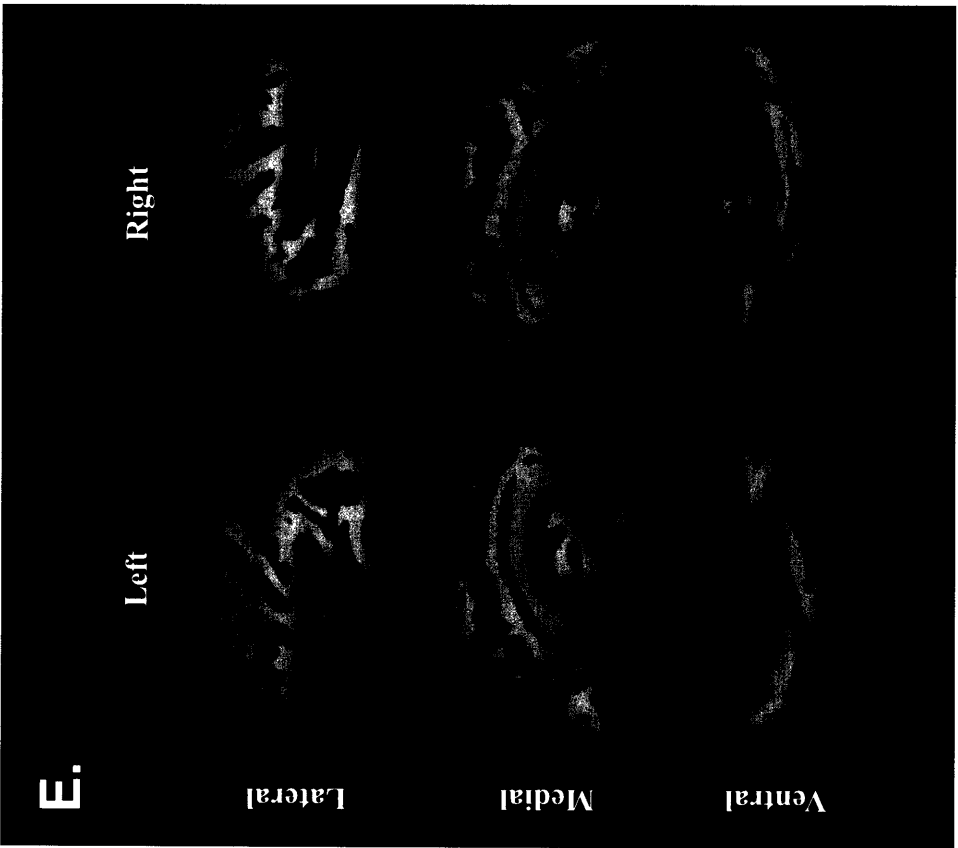
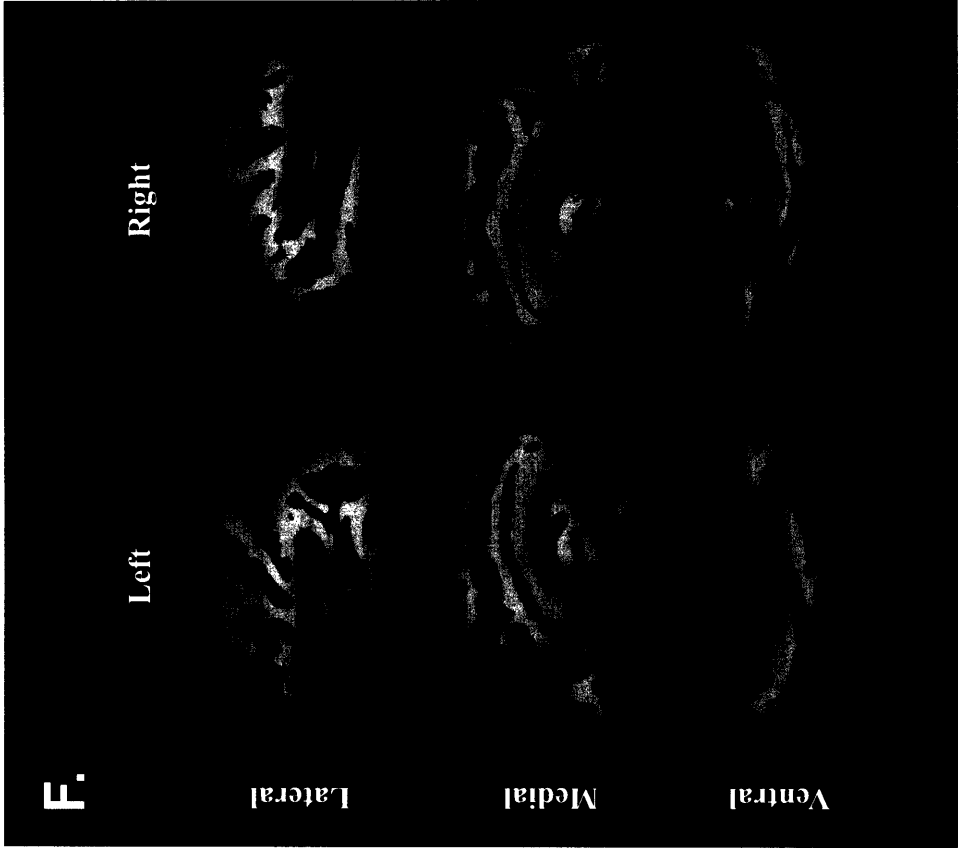
```
sliceview-sess -slice mos -analysis tau_sm6 -contrast an_in_6-10sum_rmp -map sig -s isxavg-20subj -df dir -  
space L_inf_front_cortex_020327_gr2_6-10sum -isxavg fixed
```


Figure 5-3











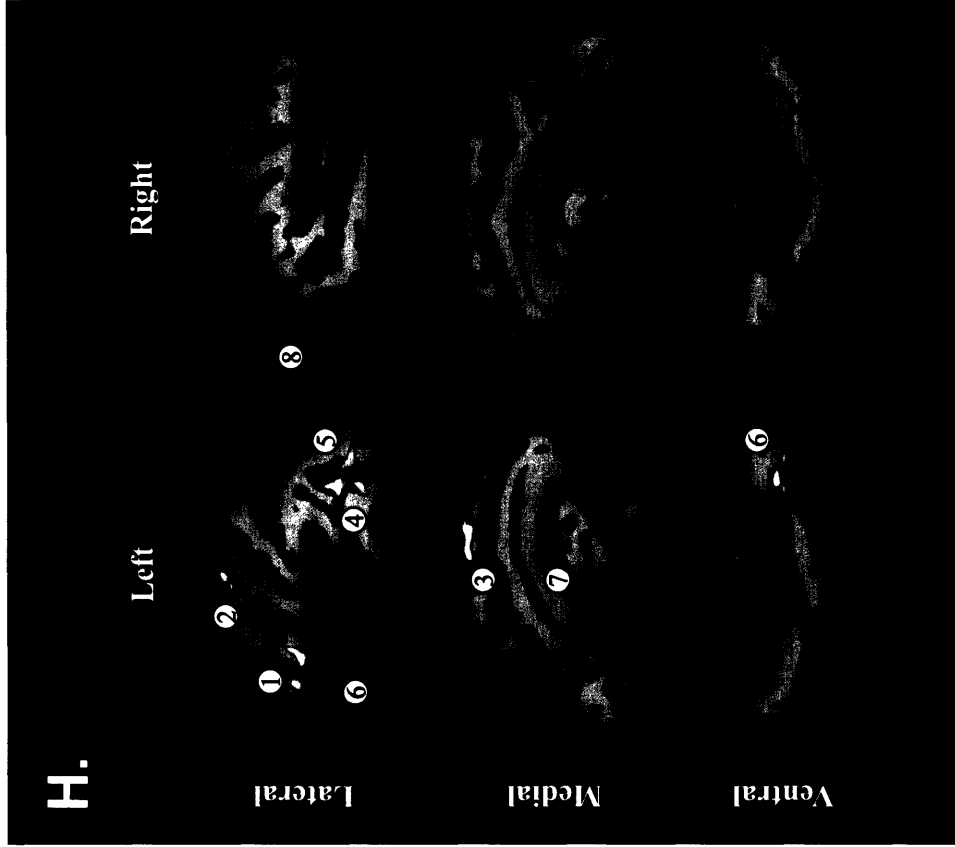
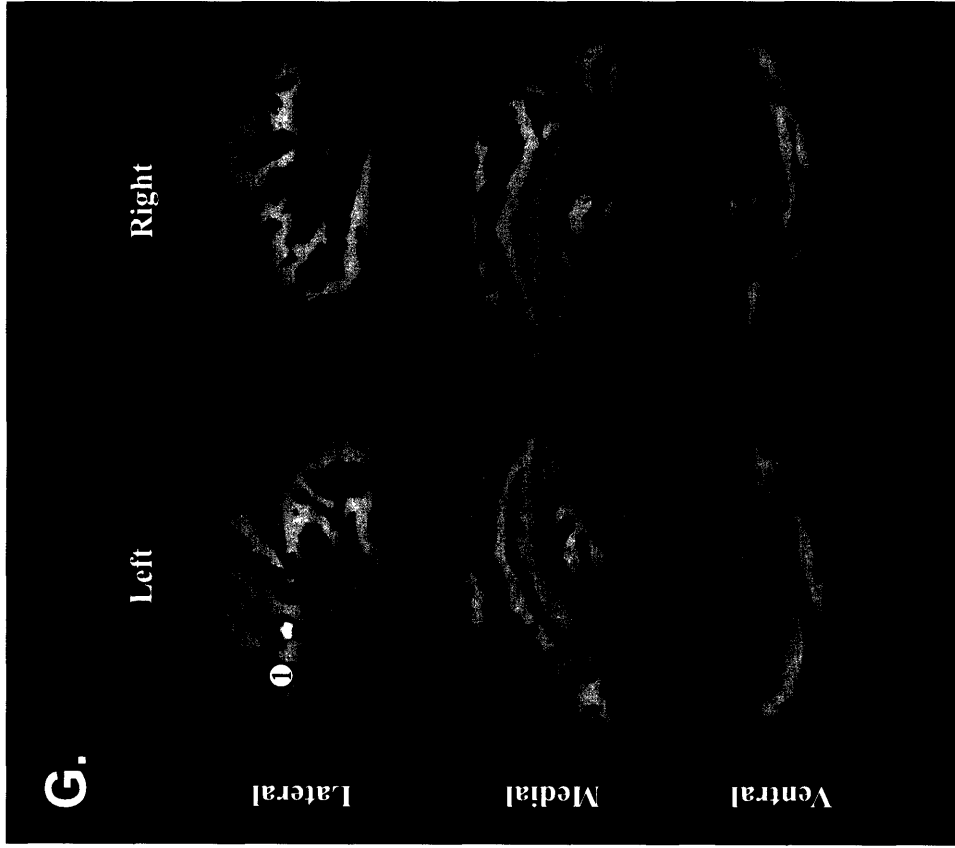


Figure 5-3. (A) Color overlays representing p-values of the subject noun MV plausibility contrast averaged across ambiguity. For all figures, color threshold (red) corresponds to $p = 0.012589$ and ceilings (yellow) at $p = 0.001259$. Each cluster has a minimum area of 200 mm^2 and a false-positive $p < 0.05$. The number label for each cluster corresponds to the cluster # of the same contrast in Table 5-1. (B) The ambiguity contrast for all plausibilities. (C) MV plausibility contrast for ambiguous sentences only. (D) MV plausibility contrast for unambiguous sentences only. (E) Ambiguity contrast for plausible sentences only. (F) Ambiguity contrast for implausible sentences only. (G) MV plausible subject noun, ambiguous sentence vs. implausible, unambiguous. (H) MV plausible subject noun, unambiguous sentence vs. implausible, ambiguous.



Appendix L: Experiment 2 Participant Summary

Participant #	EPRIME List	Age	Yrs of		Hand	Gender	ID	Recon ID	Notes
			Education	Education					
1	R	22	16	16	Left	Male	011002CE	west_020109CEa	Omitted; Left handed
2	L	23	15	15	Right	Female	011009JH	west_011019JHa	
3	R	22	16	16	Right	Female	011016MH	west_011024MHa	Omitted; Head movement
4	L	19	15	15	Right	Male	011023SL	west_011101SLa	
5	R	24	16?	16?	Right	Female	011106JW	west_011211JWa	JW1
6	-	22	15	15	Right	Male	011107AA	011107AA	Omitted; Did not return for functional
7	L	20	15	15	Left	Male	011113SP	west_020322SPa	Omitted; < 50% task performance
8	R	23	16	16	Right	Female	011113ME	west_011211MEa	
9	L	25	16?	16?	Right	Male	011127JG	west_011205JGa	
10	R	24	16	16	Right	Female	011127SC	west_011127SCa	
11	L	22	16	16	Left	Female	011204KL	west_011207KL a	Omitted; Left handed
12	R	22	16	16	Right	Female	011204NI	west_011219NIa	
13	L	25	16	16	Right	Female	011211JW	west_011218JWa	JW2
14	-	22	16	16	Right	Female	020315BC	011212EC	Omitted; Imager issues, didn't return
15	R	22	16	16	Right	Female	011218LD	west_011219LDa	
16	L	22	16?	16?	Right	Female	020115CP	west_011205CPa	Omitted; Head movement
17	-	23	16	16	Right	Female	020124FF	020124FF	Omitted; Claustrophobic, didn't return
18	R	24	16	16	Right	Female	020125SW	west_020220SWa	
19	L	23	16	16	Right	Female	020308NS	west_011024NSa	Omitted; Reading out loud?
20	L	23	16	16	Right	Female	020329KM	west_020207KMa	
21	L	18	13	13	Right	Female	020412KV	west_020325KV a	



Appendix M: Experiment 4 Participant Summary

Participant #	EPRIME Participant #	EPRIME List	Age	Yrs of Education	Hand	Gender	ID	Notes
1	1	89-R	26	>16?	Right	Female	020529SR	Omitted; Incorrect imaging sequence (25 slices)
2	2	78-R	24	16	Right	Male	020612BW	Display problems: left justified
3	3	78-L	23	16	Right	Female	020619ME	
4	4	78-L	22	16	Right	Female	020626VC	
5	5	78-R	23	16	Right	Female	020626LD	
6	6	78-L	26	16	Right	Female	020703DG	
7	7	78-R	22	16	Right	Female	020703MH	
8	8	78-L	26	16	Right	Male	020710JG	
9	9	78-R	25	18	Right	Female	020710SC	
10	-	-	24	16	Right	Female	-	Functionals incomplete, need plastic glasses
11	10	78-L	23	16	Right	Female	020724NS	
12	11	78-R	23	16	Right	Female	020724NI	
13	12	78-L	21	15	Right	Female	020731SF	
14	13	89-R	24	16	Right	Female	020807KM	Participant #10 returned to complete functionals
15	15	34-L	26	>16?	Right	Female	030226ML	
16	16	34-R	22	16?	Right	Female	030305AS	
17	17	34-L	24	16?	Right	Male	030305TG	Omitted; Poor task performance (<76% correct)
18	18	34-L	22	16?	Right	Female	030312SW	
19	19	34-R	27	>16?	Right	Female	030312SH	
20	20	34-L	19	14	Right	Female	030319KV	
21	21	34-R	24	16?	Right	Female	030326SP	
22	22	34-L	22	16?	Right	Female	030326AM	
23	23	34-R	22	16	Right	Female	030402SM	