GRAMMATICALITY, ACCEPTABILITY AND SENTENCE PROCESSING: A PSYCHOLINGUISTIC STUDY

Forrest David Braze, Ph.D.

University of Connecticut

Linguistic theory is built on an empirical foundation consisting largely of sentence acceptability judgments, deemed to reflect underlying grammaticality. This thesis focuses on extra-grammatical influences that sometimes obscure such judgments. For example, processing resource limitations may lead a perceiver to reject a grammatical sentence because its constituents cannot easily be recognized. Further, extra-linguistic processes may influence sentence ratings due to the analytic nature of the judgment task. Thus, In order for the researcher to accurately delineate the boundary between grammatical and ungrammatical sentences it is necessary to identify and evaluate non-grammatical influences that muddy acceptability judgments.

Three experiments, exploiting the technique of monitoring eye-movements during reading, probe readers' responses to various classes of unacceptable sentences. This technique was chosen because eye-movements serve as a rich source of information about how sentence features are assimilated as they unfold in time. The Referential Model of the human sentence processor (HSP) serves as a theoretical framework for interpreting results, paying careful attention both to computational devices and the resources available to them.

Forrest David Braze – University of Connecticut, 2002

Experiments 1 and 2 contrast ungrammatical sentences, containing morpho-syntactic anomalies, with others that are grammatical, yet pragmatically anomalous. The results illustrate that eye-movements are sensitive to constraints invoked by a range of morpho-syntactic and pragmatic anomalies, and that responses to the two classes of anomaly are distinct. These findings are argued to support a model of HSP that engages distinct devices in response to each class of anomalous sentence.

Experiment 3 examines the status of constraints on long-distance syntactic movement. Judgment satiation is a phenomenon in which some, but not all, classes of initially unacceptable sentences are judged increasingly acceptable with additional exposure. This study uncovers an on-line counterpart to judgment satiation, where the effort expended in processing (satiable) wh-island violations increases with repeated exposure. No such increase is seen in the case of (non-satiable) adjunct island violations. The initial unacceptability of wh-island violations is argued to stem from a processing resource limitation, while the satiation effect derives from the HSP's ability to dynamically shift resources to meet the moment-by-moment needs of its sub-components.

GRAMMATICALITY, ACCEPTABILITY AND SENTENCE PROCESSING: A PSYCHOLINGUISTIC STUDY

Forrest David Braze, Ph.D.

B.A. California State University, Fresno, 1992

M.A. California State University Fresno, 1994

M.A. University of Connecticut, 1998

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2002

UMI Number: 3062071

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

Approval Page

Doctor of Philosophy Dissertation

GRAMMATICALITY, ACCEPTABILITY AND SENTENCE PROCESSING: A PSYCHOLINGUISTIC STUDY

Presented by

Forrest David Braze, B.A., M.A.

Major Advisor _____
Diane Lillo-Martin

Associate Advisor _____
Howard Lasnik

Associate Advisor _____
Donald Shankweiler

Associate Advisor _____
William Snyder

University of Connecticut

2002

Acknowledgements

I am grateful to Diane Lillo-Martin for her patience and good humor; both of these qualities were of great value to me in her role as my advisor (and to her as well, I suspect). Her enthusiasm and gentle but incisive questioning pushed me to explore farther than I might have otherwise. I am no less indebted to the other members of my committee. Howard Lasnik exemplifies a devotion to the study of language coupled with a clear style of rhetoric that I can only hope to emulate. Donald Shankweiler deserves credit for opening my eyes to experimental methods in the study of language. William Snyder for helping me to stay focused on issues that matter. Together, they conveyed to me a solid foundation of linguistic theory, experimental technique and rhetorical skill.

I've also benefitted from the many other members of the Linguistics faculty at UConn, past and present. They all deserve credit for creating stimulating and friendly environment in which to learn and to grow as a researcher. I am extremely grateful to all of the other graduate students whose time in this department as overlapped with mine. I would not have made it through to this day without their constant help and companionship, and their occasional commiseration.

Beyond the Linguistics Department, I've benefitted from the opportunity to talk about my research from time-to-time at the UConn Psychology and Linguistics brown bag seminars. The financial support I've received from Haskins Laboratories made my stay at UConn much more comfortable than it would have been otherwise, and the time I've spent at Haskins has ensured that my graduate education was broader than it might have otherwise been. I owe a special debt of gratitude to Wei-jia Ni who blazed a path for me to follow both at UConn and at Haskins.

Finally, I owe my wife Sandra more than I can say for the support that she's given me and for all that she's done without while I've chased this white rabbit.

Table of Contents

List of Figures

List of Tables

Chapter 1: Preliminaries

1.1 Introduction and Overview

This thesis examines certain limitations of the data most often used to justify and test linguistic theory, and proposes ways that the data may be usefully supplemented. Generative linguistic theory concerns itself with the mental states that enable language use (Chomsky, 1965). These mental states define a speaker's underlying linguistic competence. Competence is a matter of knowledge, that which an individual must possess to speak and comprehend a particular language. The linguist's domain of inquiry, thus delineated, abstracts away from many influences on performative linguistic acts, either of comprehension or production. Chomsky recognized that real linguistic performance, putting aside the impossible "ideal" of a monolingual speaker in a homogeneous speech community, cannot provide a perfect image of linguistic competence.

What kinds of data, then, are useful in divining speakers' competence, when only their performance (behavior) is accessible? Judgments of sentence grammaticality have long been an important data source for linguists. Nonetheless, the nature of the relationship between such judgments and linguistic knowledge is not well understood. Consider, especially, that the act

of making grammaticality judgments is itself a form of linguistic behavior, and, as such, is subject to various constraints on performance. Chomsky (1965) recognized that speakers' linguistic intuitions, expressed as sentence judgments, ostensibly of grammaticality, are often colored by irrelevant details of performance. Hence, such judgments are best characterized, in a pre-theoretic sense, as acceptability judgments, grammaticality being only one of a number of criteria contributing to sentence acceptability. The fact that judgments are subject to a variety of influences does not mean that judgment data are hopelessly flawed, but it does argue that they should be interpreted with caution. How should a linguistic scientist proceed so as to avoid being misled? This thesis argues that one way is to seek external confirmation of sentence judgment validity. Certain measures of non-judgment language behavior are argued to be useful in this regard.

Limitations of sentence judgments as reflected in linguistic intuitions were recognized from the very earliest days of the Chomskyan program. Chomsky writes in LSLT "But intuition, of course, is an extremely weak support. The program of linguistic research would be a much clearer one if we could show experimentally that these intuitions have distinct behavioral correlates" (Chomsky, 1975, p 101). Objections to using experimental data to inform and test linguistic theory seem to revolve around the fact that these

data can often be ambiguous. But this should not preclude their use in testing linguistic theories, or in guiding their development.

In this thesis, it will be shown that a specific measure of linguistic performance, eye-movement patterns during reading, can contribute non-redundant information about the organization and employment of linguistic competence, and hence serve to complement linguistic intuitions. This possibility is important because, while theorizing grounded in acceptability judgment data has pushed our understanding of the human capacity for language to remarkable levels, the information that can be garnered from such data may, ultimately, not be sufficient to address all of the questions that theory leads us to ask.

I will not directly address problems surrounding the use of judgment data as such, as those issues have been extensively discussed elsewhere (Cowart, 1996; Schutze, 1996). The latter portion of this chapter will, however, review some findings that point to a rather indefinite relationship between sentence judgments and linguistic competence. Of particular concern is the possibility of non-grammatical factors introducing systematic biases into judgment data.

I will proceed as follows. The remainder of chapter 1 reviews a set of phenomena which demonstrate that extra-grammatical biases on sentence judgments do occur, and are, perhaps, more subtle than generally recognized.

Three classes of constructions, illustrated in (1), are discussed: multiply center-embedded clauses (1a), the stage/individual level contrast in secondary depictive predicates (1b), and island phenomena (1c). In each of these constructions, systematic, non-grammatical biases arguably contribute to diminished acceptability. This preliminary appraisal supports the idea that drawing upon a variety of data types may lead to greater insight into linguistic competence than can be obtained through use of sentence judgments alone.

(1) a. The rat that the cat that the dog chased killed had eaten the cheese.
b. John ate the carrot orange.
c. What does John wonder whether Mary likes?

Chapter 2 gives an overview of the relationship between linguistic performance as indexed by a record of eye-movements during reading, and some kinds of unacceptable, or anomalous, sentences. This overview motivates and introduces the experiments discussed in chapters 3 and 4. The two experiments in chapter 3 show that eye-movement patterns are sensitive to linguistically-relevant contrasts in sentence acceptability of both grammatical and non-grammatical genesis, and that these two broad classes can be distinguished in the eye-movement signal. The demonstration proceeds by comparing two types of (relatively) unacceptable sentences, one in which the unacceptability hinges on a grammatical anomaly and another in which it depends on a pragmatic anomaly, with a non-anomalous control. A doubly-

anomalous sentence condition, combining syntactic defect with pragmatic deviance, completes the paradigm. These experiments show that eye-movement patterns distinguish grammatical and non-grammatical anomalies from each other and from non-anomalous controls. Having established this point, chapter 4 uses eye-movement data to illuminate a matter of current theoretical interest in linguistic theory, the grammatical status of *wh*-islands. New facts about the processing status of *wh*-islands are developed, and a processing based account of these facts is proposed. Finally, Chapter 5 summarizes the results and offers some concluding remarks.

1.2 Sentence Judgments are Subject to Extra-Grammatical Influences

Considerable evidence exists to show that non-grammatical factors influence grammaticality judgments. Resource limitations on processing, for example, are commonly-invoked non-grammatical constraints on linguistic performance. A limitation in memory capacity has often been called upon to explain the difficulty of comprehending multiply center-embedded clauses (e.g. Chomsky & Miller, 1958; Gibson & Thomas, 1996; Gibson, 1998). While resource limitations on processing are commonly invoked to explain facts about sentence processing and comprehension, somewhat less attention is paid to the fact that such limitations also pose constraints on sentence judgments. Whatever the cause, there is evidence that individuals perceive difficult, but

putatively grammatical sentences, like (1a), as ungrammatical (Blumenthal, 1966; Marks, 1968).

Other differences in sentence makeup, which do not affect grammaticality, may also influence judgments. Kurtzman and Crawford (1990), for example, found that variations as irrelevant to well-formedness as the addition of a licit adverb can exert a bias on judgments even when subjects are explicitly requested to rate sentences for grammaticality. The genesis of this effect is not entirely clear. It could, in principle, result from increases in syntactic or semantic complexity introduced by the addition, or, alternately, the problem could be caused by pragmatic obstacles to interpreting such sentences out of context.[1] The latter possibility has greater intuitive appeal, given that no obvious explanation, in terms of either syntactic or semantic complexity, presents itself. Support for this surmise can be found in Ni et al. (1998) and Braze et al. (in press), where it is shown that substitution of one (subject) NP for another, less plausible one, can strongly influence sentence judgments. In this connection, McNally (1994) also argues that the infelicity of certain adjectives in the secondary depictive construction is attributable to the influence of pragmatic constraints on sentence interpretation.

Variations in sentence properties, as such, may not be the only source of variation in judgments. Ross (1979) suggests that individual differences in the sheer quantity of exposure to language (independent of qualitative, e.g.

dialectal, differences) can influence sentence judgments. He observes that differences in degree of linguistic sophistication, as in linguistic training, can act on judgments. It is an open question whether linguistic experience exerts its effects specifically on linguistic operations, or, perhaps, only on meta-linguistic behavior. Whatever the case, it seems plausible that these experiential factors are related to what has come to be known as the syntactic satiation effect (Hiramatsu, 1998; Snyder, 2000), in which repeated exposure to some classes of initially unacceptable sentences (violations of island constraints on long-distance movement, for example) causes them to become increasingly more acceptable.

A telling property of judgment satiation is that it does not occur with all types of putatively ungrammatical sentences. For example, Snyder found clear satiation effects for argument *wh*-extraction from complex NPs and *whether* complements to *wonder*, as well as some evidence of satiation for extraction from subject islands. This result contrasts with the response to violations of adjunct islands, which failed to show any sign of satiation. Hiramatsu's results largely parallel those of Snyder. She found robust satiation effects for extraction from *whether* complements and subject islands (although her findings differ from those of Snyder in that they exhibit no evidence of satiation for CNPC violations). Evidence of other behavioral dissociations among island constraints will be followed up in chapter 4.

1.2.1 Constraints on Center-embedding: grammar or memory

It is generally agreed that sentences containing multiply center-embedded clauses, like (2a), are highly unacceptable. The question, then, is how best to characterize that unacceptability. Is it a matter of grammar, or of something else? Answers to questions of this sort depend partly on the kinds of formal devices that can be motivated in the theory of grammar. In fact, theoretical considerations may over-ride intuitions about the grammatical status of such sentences. Blumenthal (1966), for example, presents evidence that individuals perceive sentences like (2a) as ungrammatical. Yet, center-embedded constructions, whether multiply-embedded like (2a), or singly embedded like (2b), are typically analyzed as being grammatical, but "difficult" for grammar-independent reasons. An informal review of the argument that leads to this somewhat counter-intuitive conclusion may be useful. I begin with an overview of relevant premises and goals of linguistic theory. What follows is based largely on Chomsky (1965, chapter 1).

(2) a. The rat that the cat that the dog chased killed had eaten the cheese.
b. The rat that the cat killed had eaten the cheese.
c. The dog chased the cat that killed the rat that had eaten the cheese.

In the generative tradition, we say that when an individual has learned a particular language, ***L***, he possesses a specific grammar characterizing its

grammatical sentences, a grammar of ***L***. Specific grammars arise in the individual's mind through a combination of linguistic experience and a system of innate universal principles. These principles serve two crucial functions. On the one hand, they constrain the range of linguistically relevant grammars, grammars of possible natural human languages. On the other, they serve as metrics that allow the learner to choose, from among the possible grammars, one that is optimally consistent with his linguistic input. It is the first of these functions that is of present interest.

The linguist's theory of these universal principles (universal grammar, or UG) is a theory both of those innate properties of human cognition that constrain the possible specific grammars for human languages, and of how the individual learner attains an optimal specific grammar on the basis of incomplete, and often defective, linguistic data. Additionally, a theory of UG must meet the sometimes opposing requirements of descriptive and explanatory adequacy. In order to satisfy the needs of descriptive adequacy, a theory must be powerful enough to accommodate a descriptively adequate specific grammar for the observable phenomena of every natural human language. The requirements of explanatory adequacy require that the theory of UG constrain the range of possible specific grammars to a set that will be learnable, given the linguistic data accessible to the learner.

In working toward a fully adequate theory of grammar, we should prefer to adopt processes and structures that have the greatest possible generality, and so contribute to our understanding of the relationships among apparently diverse linguistic phenomena. Returning to the question that prompted this excursion, we ask whether, given the expressive requirements imposed by the properties of ***L*** represented by the sentences in (2) and the constraints that we are prepared to admit into our theory of UG, it is possible to restrict our theory of the specific grammar of ***L*** in a way that will admit sentences (2b) and (2c), while at the same time excluding (2a). Such a restriction, if properly motivated, will provide a grammatical explanation for the unacceptability of (2a).

Of two possibilities for realizing such a constraint, neither, ultimately, is sustainable. First, we might consider a blanket prohibition against center-embedded recursive structures, but the generally acceptable (2b) shows that this is not a viable option. As an alternative we might posit a limit on the maximum depth of embedding, but there are two problems with this approach. On the one hand, the contrast between the unacceptable (2a) and the acceptable (2c) shows that any such restriction could not be fully general. It would need to be characterized as a constraint on only some recursive structures. Perhaps an even greater problem is that this approach would require the grammar to incorporate a counting mechanism such that it could

keep track of the degree to which any structure is embedded. To complicate the theory of grammar with such a powerful mechanism, for the sole purpose of ruling out the construction exemplified in (2a), would seem unjustified. This leads us to the conclusion that the unacceptability of (2a) and similar sentences is due not to the application of grammatical principles, as such, but rather to the operation of extra-grammatical constraints.

In fact, the unacceptability of (2a) has been attributed to limitations of memory available to the parsing mechanism, although specific proposals vary considerably in detail and specificity (e.g. Miller & Chomsky, 1963; Lukatela, Shankweiler, & Crain, 1995; Gibson, 1998). In the case of center-embedding a non-grammatical explanation falls out straightforwardly, but, in other cases, discussed below, the judgment is not as easy to make.

1.2.2 Secondary Depictives: grammatical licensing or pragmatic violation?

The choice between grammatical and non-grammatical explanations of constraints on the distribution of adjectives as secondary depictive predicates (SDP) is not straightforward. Not all adjectives are equally felicitous as SDPs, as shown in (3).

(3) a. John bought the carrots fresh.
b. *John bought the carrots orange.

The question, then, is whether the correct analysis of the contrast in (3) will appeal to grammatical or non-grammatical principles. McNulty (1988) discusses in detail the syntactic properties of SDPs, giving a lucid account of their external distribution (the relationships they bear to other phrasal elements in the clause) and internal structure, but her story falls short of explaining the contrast between (3a) and (3b).

Rapoport (1990; 1991) leverages the observation that adjectives in acceptable SDPs, like (3a), are stage-level predicates, denoting relatively temporary properties, while those in unacceptable SDPs, like (3b), are individual-level predicates, denoting longer-lived properties. She suggests that restrictions on SDPs cannot be stated in terms of syntax, developing an account of the acceptability contrast grounded in the Kratzer (1995) theory of the semantic constitution of stage and individual level predicates.

Kratzer (1995) maintains that argument structures of stage-level predicates include an event argument, which individual-level predicates lack. The significance of this is that a predicate with an event-place in its argument structure can individuate over times or occasions, while one without an event-place cannot. Kratzer exploits this difference in semantic makeup to explain a variety of phenomena which are observed to be sensitive to the stage/individual level contrast. These include bare plural subjects, (4), and existential *there* constructions, (5).

(4) a. Firemen are available.
b. Firemen are altruistic.

(5) a. There are firemen available.
b. *There are firemen altruistic.

In (4a), the relevant observation is that a stage level predicate applied to a bare-plural subject is ambiguous between generic and existential interpretations of the subject, but if an individual level predicate is substituted, as in (4b), only the generic interpretation is available (Carlson, 1977). The contrast in (5) shows that existential *there* sentences containing individual level predicates are considerably less acceptable than similar sentences containing stage level predicates (Milsark, 1974).

Both Kratzer (1995) and Rapoport (1991, fn 16) acknowledge that the classification of a predicate as either stage level or individual level cannot be absolute. If the stage/individual level contrast reflects a true distinction in the lexical makeup of predicates, then we must allow for the existence of some process that shifts predicates from one class to the other. For example, given an appropriate context, putatively individual level predicates, like *altruistic*, can denote transient properties of an individual, as shown in (6). Kratzer proposes that when an individual level predicate is used in a way that requires a stage level interpretation, it somehow acquires an event argument.

(6) John joined the police force altruistic and full of optimism.
He left it five years later intolerant and cynical.

Rapoport (1990) initially argues for a constraint on the acceptability of secondary predicates like (3), such that only verbs which entail a change of state or location in their object can participate in SDPs.[2] This aspect of meaning is assumed to be encoded in the lexical semantic properties of the verb. Rapoport (1991) elaborates on this observation, claiming that the lexical semantic properties of the adjunct predicate interact with those of the verb to constrain the acceptability of SDPs. Specifically, an association between the event arguments of the main verb and secondary predicate serve to license the latter. A similar account can be found in Miyamoto (1994). Both Rapoport and Miyamoto argue that the presence of an event-place in the argument structures of both main and secondary predicates is necessary to structurally license the SDP.

McNally (1994), however, proposes an alternative account of the contrast in (3). She argues that the lack of event-place arguments is not the underlying cause of the unacceptability of sentences like (3b). She asserts that for those individual level predicates that are not felicitous as SDPs, an explanation in terms of pragmatic infelicity is available. One objection she makes to Rapoport's analysis is that, arguably, even individual level predicates must possess event arguments in order to receive a correct interpretation. She points to the fact that, while such predicates typically denote relatively persistent states, these states are, nonetheless, generally not eternal. There

are temporal bounds for individual level predicates, even where, in the extreme case, the intervals over which they apply extend to the entire existence of their bearers. If both stage and individual level predicates possess event arguments, then an explanation that relies on this dimension to differentiate their putative acceptability in SDPs cannot go through.

McNally proposes a two-part explanation. First, individual level predicates are associated with an inference of temporal persistence. The intervals over which they hold are construed to be some large portion of the whole existence of their bearers. Stage level predicates lack this characteristic. She then leverages the fact that the interpretations of SDPs require that the properties denoted by the main predicate and the secondary predicate be coextensive. In some sense, McNally's proposal treats the SDP as a species of temporal modifier, or restrictor, on the main predicate. This pair of assumptions allows McNally to appeal to the pragmatic concepts of informativeness and relevance (Grice, 1975) to explain the contrast in (3). The heart of McNally's explanation lies in the fact that, when a predicate like *orange* is used as an SDP, bearing, as it does an inference of persistence, the coextensiveness requirement for main and secondary predicate is trivially satisfied. Thus, when presented out of context, a sentence like (3b), repeated as (7a), violates the pragmatic requirement that an utterance be informative and relevant, in context. However, because the inference of temporal

persistence is pragmatically grounded, McNally predicts that it can be contextually negated. This prediction seems to be born out. In fact, that putatively individual-level predicates can, in appropriate contexts, serve felicitously as SDPs, is acknowledged by Rapoport (1991, fn 16), and demonstrated here in (7).

(7) a. John bought the carrots orange. (3b)
b. After sitting on his kitchen counter for a week, they turned a leprous brown.

Judgment data tells us little about which level of grammar is implicated in the differential acceptability of sentences like those in (3), or even whether a grammatical explanation is appropriate at all. Only after a grammatical explanation is imposed do we interpret the unacceptability of sentences like (7a), when presented without supporting context, as ungrammaticality.

1.2.3 Island Constraints

Constraints on long distance syntactic movement, manifest as so-called island effects, have been a topic of much interest. A variety of island types have been identified in the literature, but I would like to focus, for the moment, on a particular type, the *wh*-island, which will play a prominent role in chapter 4 of this thesis. A *wh*-island violation is exemplified in (8). To my knowledge, Chomsky (1964) first noted the deviance of sentences like (8), along with a

number of other unacceptable configurations involving *wh*-elements, including both matrix and embedded questions, and relative clauses.[3] The details of Chomsky's (1964) analysis need not concern us, other than to note that it was the first in a line of purely syntactic explanations that attempt to give a unified account of restrictions on *wh*-question and relative clause formation.[4]

(8) *What does John wonder whether Mary likes *t* ?

Refinements to the theory of grammar, and the theory of *wh*-movement in particular, led to the formulation of the subjacency condition on movement (Chomsky, 1973, 1977, 1986). Movement is constrained to proceed by a series of "shortest possible" steps. Each step is determined by the nearest available stopping position (defined in terms of potential adjunction positions). Subjacency constrains the distance which a *wh*-element can move at one time, yet allows for long-distance *wh*-movement to be accomplished through a series of intermediate steps. Chomsky (1986) defines the limits on movement in terms of barriers, (11), and the ancillary concepts of L-marking, (9), and blocking category, (10).[5]

(9) L-marking: a category is L-marked if it is θ-marked by a lexical category.

(10) Blocking category (BC): γ is a BC for β iff γ is not L-marked and γ dominates β.

(11) Barrier: γ is a barrier for β iff (a) or (b):
a. γ immediately dominates δ, δ a BC for β
b. γ is a BC for β and γ is not an IP

Thus, CP of the embedded clause in (8), elaborated in (12), is a barrier. Movement across one or more barriers in a single step results in a subjacency violation with the severity of the resulting anomaly corresponding to the number of barriers crossed. Simplifying somewhat, movement of *what* across CP in (12) results in a 1-subjacent violation, accounting for the modest deviance of (8).

(12) What does John $[_{VP}$ wonder $[_{CP}$ whether $[_{IP}$ Mary $[_{VP}$ likes $t]$]]]

Movement to spec-CP, between CP and IP, can provide an "escape hatch" for *wh*-movement, nullifying a potential barrier. This is illustrated in (13), where *what* moves step-wise from its base position through spec-CP of the embedded clause, to spec-CP of the matrix. If spec-CP is already filled, however, then the escape hatch is unavailable, and the moved element must skip over the Barrier in a single step. This is the case in (8)/(12), where spec-CP is occupied by *whether*.[6]

(13) What does John $[_{VP}$ think $[_{CP}$ t' that $[_{IP}$ Mary $[_{VP}$ likes $t]$]]]

Two additional principles are commonly assumed to constrain syntactic movement. The Condition on Extraction Domains (CED) requires that the phrasal domain from which an element is moved must be properly governed (Huang, 1982), where proper government is defined roughly as in (14).[7] Antecedent concepts of government and c-command are given in (15) and (16), respectively.

(14) α properly governs β iff α governs β and
  a. α θ-marks β, or
  b. α and β are coindexed (by movement)

(15) α governs β iff
  a. α c-commands β, and
  b. there is no γ such that α c-commands γ and γ c-commands β

(16) α c-commands β iff
  a. neither α nor β dominates the other, and
  b. the first branching node that dominates α also dominates β.

The CED accounts for the unacceptability of extraction from within subjects and adjuncts. A CED violation is shown in (17b); the adjunct clause, plausibly adjoined to IP, is not properly governed, and so extraction from it results in a severely degraded sentence. Note that the extraction site itself is properly governed (by the verb *clear*).

(17) a. John washed dishes after Mary cleared the table.
b. *What did John wash the dishes [after Mary cleared *t*]

The Empty Category Principle (ECP), like the CED, appeals to the notion of proper government, but the ECP requires that traces of syntactic movement be properly governed (Chomsky, 1981). The ECP is typically called upon to explain the added unacceptability of adjunct extraction (relative to argument extraction) from within a *wh*-island, as the contrast between (18a) and (8)/(12) (repeated as (18b)) demonstrates. Note that (18a) is to be interpreted as a question expressing John's puzzlement about the reason for Mary's liking beer, not about the reason for John's wondering. Sentence (18a) violates both subjacency and the ECP, while (18b) violates only subjacency.

(18) a. *Why does John wonder [$_{CP}$ whether [$_{IP}$ [$_{IP}$ Mary [$_{VP}$ likes beer]] *t*]]
b. ?What does John [$_{VP}$ wonder [$_{CP}$ whether [$_{IP}$ Mary [$_{VP}$ likes *t*]]]] (12)

In attempting to provide a comprehensive syntactic account of a wide range of island phenomena, the Barriers approach, and derivative work, suffers from some interesting empirical limitations. Three diverse facts about island effects complicate any analysis. First, there is the widely recognized distinction that divides islands into two types, strong and weak. Although the demarcation between these classes is somewhat fuzzy, the basic observation is that strong islands are uniformly inviolable (adjunct-islands are a typical case), while weak islands may allow extraction of a limited class of elements resulting in, at

worst, a mildly anomalous sentence (these include some sub-types of *wh*-islands).

Second, is the fact, first observed by Ross (1969), that for some types of island violation ellipsis of the island has an ameliorating effect. Consider the paradigm in (19).

(19) a. Bill wonders whether Mary likes something in that store.
b. *What does Bill wonder whether Mary likes?
c. Bill wonders whether Mary likes something in that store, and I know what ~~Bill wonders whether Mary likes~~.

Under conventional analyses of sluicing, the syntactic representation of the second conjunct in (19c) does not differ in any relevant way from that of (19b). Yet, there is a pronounced difference in acceptability. Merchant (1999) suggests, given that the syntactic representations do not differ in any relevant way, the unacceptability of (19b) cannot be due to a syntactic constraint. He proposes that at least some island effects are due to a principle or constraint that operates at the level of PF, a post-syntactic level of representation and processing.[8]

Finally, Snyder (2000) and Hiramatsu (2000), discuss the so-called satiation effect. As noted, satiation refers to the fact that repeated exposure to some classes of unacceptable sentences causes them to become increasingly acceptable. An interesting property of judgment satiation is that it occurs with some, but not with all, types of islands. Thus, as noted, satiation occurs with

violations of *wh*-islands, (20a), subject islands, (20b), and violations of the complex noun-phrase constraint, (20c), but not with violations of adjunct islands, (20d). Snyder proposes that the genesis of those island phenomena that are subject to satiation may be the result of a processing based constraint, rather than a syntactic one (also Kluender & Kutas, 1993b; Kluender, 1998). Satiation effects will be discussed in considerably more detail in chapter 4.

(20)	a.	Who does John wonder [$_{CP}$ whether [$_{IP}$ Mary likes *t*]] ?	*wh*-island
	b.	What does John know [$_{CP}$ that [$_{NP}$ a bottle of *t*] fell on the floor] ?	subj-island
	c.	Who does Mary believe [$_{NP}$ the claim [$_{CP}$ that John likes *t*]] ?	CNPC
	d.	Who did [$_{IP}$ John talk with Mary [after seeing *t*] ?	adjunct-island

Boeckx (2001) points out that there is some overlap between the types of unacceptable constructions that improve under sluicing, and the types that exhibit satiation effects. He argues that converging evidence from sluicing and satiation effects supports the idea that certain island phenomena may be attributed to processing constraints, rather than the result of constraints on syntactic movement operations. While the correspondence between amelioration under sluicing and satiation effects is not perfect, there is sufficient overlap to suggest these phenomena may depend on some common underlying condition. I will return to the matter of island phenomena in Chapter 4, exploring the potential contribution to this discussion of data from online sentence processing measures.

1.3 Summary

A review of relevant phenomena demonstrates the sometimes problematic nature of the relationship between sentence judgments and linguistic competence, and points out the insufficiency of sentence judgment data as a monolithic foundation for linguistic theory. Judgment data, like any measure of linguistic performance, is subject to the influence of non-grammatical factors. Especially problematic are cases where non-grammatical factors introduce systematic, though possibly unrecognized, biases into sentence judgments. The three cases reviewed in this chapter illustrate that interpretation of judgment data is not straightforward. A central source of the difficulty lies in the uni-dimensional nature of data garnered from the typical sentence judgment task. A leading idea in this thesis is that drawing upon a variety of data types may lead to greater insight into linguistic competence than can be obtained through use of grammaticality judgments alone. The remaining chapters seek to develop the use of the technique of monitoring eye-movements during reading as a data source to supplement acceptability judgments in shaping and testing linguistic theory. At the same time, novel on-line data with regard to certain island constraints is developed, and a processing-based account of these phenomena is proposed.

Notes

1. In general, I presume pragmatic properties to be those features of sentence meaning that depend on context. At least two senses of the term "pragmatics" can be distinguished. The first refers to the knowledge of general conversational principles that allow a hearer to fill the (sometimes large) gap between what a sentence means, and what the speaker intends to convey. These principles are codified in Grice (1975).

   Of greater relevance in the present context is a second sense of pragmatics, referring to the plausibility of a text. Pragmatic processing, in this sense, involves recruiting world knowledge to evaluate the contextual likelihood of sentence meaning.

2. She specifically restricts her claim to object modifying SDPs (Rapoport, 1990, fn. 25).

3. The unacceptability of sentences like (8) is generally interpreted in the literature as ungrammaticality, but not universally so. Karttunen (1977), for example, considers sentences like (8) to be grammatical. The deviance of such "*whether*-island" violations are typically treated together with "true" embedded *wh*-questions, exemplified in (a), although (8) is, in fact, the embedded correlate of a matrix yes/no question.
   (a) *What does John wonder how Mary made.

   A unified account of these extraction facts would seem to be somewhat problematic. Inarguably, extraction from a *wh*-island, like (a), is much less acceptable than extraction from a *whether*-island, as in (8). Thus, we might question whether the unacceptability of *wh*-extraction from "true" *wh*-islands, and the unacceptability of such extraction from *whether*-islands, are due to violations of the same principle, or, perhaps, have different causes.

4. Ross (1967), discussed a number of constraints on syntactic movement, including *wh*-islands. His was the first detailed formal attempt to give general statements of a variety of island effects.

5. In introducing the theory of barriers, Chomsky (1986) makes some rather stipulative assumptions (not detailed here) about what is, and what is not, a legitimate adjunction site. Rizzi's (1990) advances on the theory dispense with these.

6. The nature of the element occupying spec-CP in indirect questions has been subject to some debate. For example, Lasnik and Saito (1992, p191, n9) place *whether* itself in spec of CP, while Rizzi (1990, p31, 95, 98) places *whether* in C^0 with, presumably, a null *wh*-operator in spec of CP. The particular choice is not important to the present discussion, but, for concreteness, I will follow the Lasnik and Saito treatment.

7. Various versions of the notion proper government have been proposed since Chomsky (1981), details of which are not important here. Salient discussion can be found in Lasnik and Saito (1984; 1992).

8. Stepanov (2001) provides other arguments for typological distinctions among islands, maintaining that some are the result of violations of PF constraints, while others arise as a consequence of the syntactic structure building mechanism.

Chapter 2: Supplementing Sentence Judgments

This chapter reviews evidence that the pattern of eye-movements during reading has significant potential as a marker of theoretically interesting taxonomic distinctions. Anomalous sentences, of various types, have long been used to probe aspects of linguistic competence, chiefly serving as material for the sentence grammaticality/acceptability judgment task. I argue that these types of test materials may be fruitfully employed in studies using the technique of monitoring eye-movements during reading. Further, the juxtaposition of sentence judgment and eye-movement data will certainly provide richer information about the knowledge and processes feeding language comprehension than can be gleaned from sentence judgments alone.

As pointed out in Chapter 1, a variety of factors can render a sentence anomalous, so that it deviates from some standard of "acceptability." Grammatical defects are one source of anomaly, and sentences containing a grammatical defect are properly termed "ungrammatical" (hence, these are technically not sentences at all).[1] Processing difficulty is another ingredient in sentence acceptability. Teasing apart contributions of grammaticality and other factors to sentence acceptability is a challenging task. The decision to categorize a word-string as a grammatical sentence must always be subject to

revision. In fact, the close relationship between grammar and parser, may makes it particularly difficult to distinguish between the contributions of grammatical constraints and processing constraints on any one behavioral indicator.

Previous research has shown that a perception of anomaly may arise where the relationship between the denotation of a sentence and the perceiver's knowledge of the real world is strained. Pragmatic anomaly, of the sort depicted in (21a), clearly lies outside the domain of grammar as such, although it can certainly lead to measurable degradation in acceptability.

(21) a. The cats won't usually bake the food that we leave on the porch.
b. The cats won't usually eat the food that we leave on the porch.

(22) a. The cats won't usually eating the food that we leave on the porch.
b. The cats won't usually eat the food that we leave on the porch.

For example, Ni et al. (1998) found that, in a sentence judgment task, raters considered sentences like (21a) and (22a) both to be less acceptable than control sentences.[2] While there was some tendency for the morpho-syntactic anomalies to be rated as less acceptable than pragmatic anomalies, the trend was only marginally significant. Braze et al. (in press) gathered acceptability judgments (from a larger group of raters) for similar matched sets of sentences and found no difference in the degree of acceptability for syntactic and pragmatic anomalies.[3] Yet, both anomaly types were judged significantly less

acceptable than matched controls. The fact that sentences with such obviously dissimilar anomalies evoke similar ratings underscores a limitation of sentence judgment data.[4]

Sentence judgment data does not lend itself to a data-driven categorization. That is to say, one cannot equate the degree of "unacceptability" of a sentence with the underlying etiology of the violation(s) in question. It is surely possible that two different sources of anomaly, even two different types of grammatical anomaly, can yield the same degree of unacceptability (Schutze, 1996).

2.1 Eye-movements in reading

It would be useful to have a response measure that is differentially sensitive to kinds of anomaly. Recently, the wide availability of techniques more often associated with the exploration of language processing has expanded the set of instruments available for examining, the relationship between grammar, and extra-grammatical processes. These techniques include measures of brain activity (in the course of comprehending various types of linguistic materials), as well as behavioral techniques like recording eye-movements during reading. The latter methodology is exploited in the empirical component of this thesis.

Data from psychophysical techniques, measuring changes in electrical activity of the brain (ERPs), or in cerebral blood flow (fMRI / PET), in response to linguistic stimuli, show some promise in relating linguistic anomaly to cognitive function. But interpretation of these types of data is not without its challenges, largely because the mapping between neural events and cognitive function is not well understood (Coulson, King, & Kutas, 1998; Kutas, Federmeier, & Sereno, 1999).

The mapping problem is not unique to psychophysical methods, but the technique of monitoring eye-movements during reading has shown considerable promise as a probe of specific cognitive function. It is well-established that variations in readers' eye-movements exhibit sensitivity to lexical, structural, semantic, and pragmatic properties of a text (see Rayner, 1998, for an overview). Thus, the sequential path and timing of successive eye fixations during reading can serve as a rich source of information about how sentence features are assimilated in the course of sentence comprehension.

The problem of identifying specific mappings between measured response and cognitive function in the domain of eye-movements during reading has only been partially addressed. The question remains as to whether qualitatively different sentence properties (putatively invoking distinct cognitive functions) evoke qualitatively different patterns of eye-movements. This is one issue that will be addressed in the following chapters.

An advantage to the eye-movement technique is that, unlike the task of making sentence judgments, eye-movements made while reading are largely divorced from conscious control, and so, they offer a privileged position from which to view, if only indirectly, the ordinarily hidden processes associated with linguistic processing. This is a significant advantage, as the intermediate stages of linguistic processing, prior to the final output, are largely automatic and unavailable to consciousness (Fodor, 1983).

2.1.1 The eye-movement signal

Eye-movement patterns are complex, and so it has proven useful to break down the EM signal into various components. Parsing the eye-movement record yields a number of indicators relevant to linguistic processing which provide information that is not readily available to introspection. Among those that have proven most useful in previous research are first-pass reading time (Just & Carpenter, 1980), the incidence of first-pass regressions (Frazier & Rayner, 1982), and regression-path reading-time (Konieczny, Hemforth, & Scheepers, 1997). First-pass reading time is the time spent reading a specific region of text, generally one or two words in length, before looking outside that region, either to the right or to the left. Subsequent refixations of the region do not count toward the first-pass total; these are considered second-pass fixations. By definition, first-pass reading time is

contingent upon there being at least one fixation in the region of interest. In other words, sentence regions that are skipped over, are counted as missing values, rather than being tallied as having "zero" reading time. A region is counted as having a first-pass regression if its final first-pass fixation ends in a backward glance to an earlier region of the text. Hence, first pass regression is a dichotomous variable. Regression-path reading time is defined as the sum of all fixations beginning with the first fixation on a word or region of interest, and ending with the first fixation to the right of it. There is evidence that these indicators, among others, capture non-redundant information about linguistic processes (McConkie, Hogaboam, Wolverton, Zola, & Lucas, 1979; Frazier & Rayner, 1982; Crain, Ni, Shankweiler, Conway, & Braze, 1996; Ni, Crain, & Shankweiler, 1996; Liversedge, Paterson, & Pickering, 1998; Braze et al., in press).

2.1.2 Interpreting eye-movements

Crain et al. (1996), in a study of reduced-relative and prepositional-phrase attachment garden-path sentences, like those in (23), speculate that first-pass reading times reflect the influence of information that is rapidly assimilated by the reader, whereas regressive eye-movements may reflect a processing barrier that cannot immediately be overcome. When disambiguation of a garden-path sentence is structural, as in (23a), the garden-

path effect surfaces in the form of elevated reading times in the disambiguating word, *were* in (23a). When disambiguation relies on plausibility, or pragmatic information, the garden path effect surfaces primarily in the form of an increased incidence of regressive eye-movements from the disambiguating region, *cracks* in (23b). These results suggest that distinct signatures associated with the use of syntactic and pragmatic information may be read off the eye-movement signal (also see Frazier & Rayner, 1982; Ehrlich & Rayner, 1983; Kennedy, 1983).

(23) a. The horses raced past the barn were unable to clear the jump cleanly.
b. The man painted the door with large cracks before the festival.

Ni et al. (1998) explored this possibility in greater detail using sentences similar to those in (21) and (22), and repeated in (24). Thus, the study contrasts sentences that are uncontroversially ungrammatical (morpho-syntactically anomalous) with others that are fully grammatical, yet degraded in acceptability due to their pragmatic content. Ni et al. found no evidence in the eye-movement record of a delay in initial sensitivity to pragmatic constraints, relative to the onset of sensitivity to syntactic anomaly (see also Murray & Rowan, 1998). The eye-movement record reflects the presence of both types of anomaly immediately upon encountering the anomalous verb, *bake* in (24a), and *eating* in (24b). Detection of the two kinds of anomaly was

rapid and simultaneous. Thus there was no support for the prior availability of syntactic information within the parsing mechanism.

(24) a. The cats won't usually bake the food we leave on the porch.
b. The cats won't usually eating the food we leave on the porch.
c. The cats won't usually eat the food we leave on the porch.

Ni et al. (1998) found that, although the onsets of effects for both syntactic and pragmatic anomalies were simultaneous, each kind of anomaly played out subsequently in a distinctive way. This was true both of regional reading times and frequency of regressive eye-movements. For syntactic anomalies the incidence of regressions was immediately elevated at the point of anomaly and just beyond, thereafter returning to the baseline. In contrast, frequency of regressions for pragmatic anomalies increased progressively from the point of anomaly to the end of the sentence. Regional reading times for pragmatically anomalous sentences showed a similar monotonic increase; at no point did reading times for syntactically anomalous sentences rise above baseline. The Ni et al. results support the surmise that the eye-movement record is differentially sensitive to morpho-syntactic and pragmatic anomaly.

Braze et al. (in press) followed up on the Ni et al. study, using similar test sentences. The new materials, however, contained somewhat greater variation in both the pragmatic and syntactic anomalies incorporated. Post-anomaly codas were also longer, and more consistent in length, than was the

case in the Ni et al. materials. See (25). The purpose of incorporating greater variation in the test materials was to explore the limits of the contrast between response to pragmatic and syntactic anomaly reported in Ni et al. (1998). Longer post-anomaly codas allowed for a closer examination of the parser's response to each anomaly type as reading continued beyond the anomalous verb.

(25) a. The roses will soon crack if left out of water for very long.
b. The roses will soon wilted if left out of water for very long.
c. The roses will soon wilt if left out of water for very long.

A key result of this study is that the eye-movement pattern beyond the point of anomaly yields information that is important to understanding the contrasting effects of syntactic and pragmatic incongruities. Braze et al. examined the targets of regressive eye-movements evoked by each anomaly type. They found that, in the case of pragmatic anomaly, eye-movements directed backward (regressions) from sentence final regions (*for very long*) typically targeted sentence regions in which the pragmatic mis-match occurred (*the roses will soon crack*). This was not true of sentences containing syntactic anomalies, where backward eye-movements from sentence end typically targeted the immediately preceding sentence region (*of water*). These results provide even more support for the dissociability of syntactic and pragmatic effects in eye-movement patterns.

2.2 Looking Forward

The empirical component of this thesis comprises three eye-movement experiments designed to probe the human sentence processing mechanism's (HSPM) response to anomalous strings of various types. Experiments 1 and 2 contrast sentences that are incontrovertibly ungrammatical (morpho-syntactically anomalous) with others that are fully grammatical, yet degraded in acceptability due to their pragmatic content. The grammatical status of each type of sentence in experiments 1 and 2 is stipulated as a premiss. These experiments are intended to further vet the eye-movement technique in distinguishing ungrammatical strings from sentences containing anomalies of a non-grammatical nature. Experiment 3 applies the tools refined in experiments 1 and 2 to a presently unsettled issue of current interest in theoretical linguistics: the grammatical status of wh-islands (Berwick & Weinberg, 1986; Kluender & Kutas, 1993b; Merchant, 1999; Snyder, 2000).

Experiment 1 examines strings similar to those in (21) and (22), but with an additional condition containing simultaneous syntactic and pragmatic anomalies. The experiment extends previous work by providing evidence that the qualitatively different patterns of eye-movements evoked by the two categories of anomaly (Ni et al., 1998; Braze et al., in press) are not due merely to quantitative differences in the degree of anomaly represented by the two sentence types. This experiment goes toward establishing the sensitivity and

general effectiveness of the eye-movement technique in discriminating qualitative differences in sentence types. Details can be found in chapter 3.

Experiment 2 is conceptually similar to experiment 1, and to the work reported in Ni et al. (1998) and Braze et al. (in press), but introduces syntactic and pragmatic anomalies that differ in kind from those exploited by earlier studies. Experiment 2 shows that the contrast demonstrated in previous work (op. cit.), and experiment 1, generalizes beyond the subject-verb predication relation in the case of pragmatic anomaly, as exemplified in (25a), and beyond verbal morphology in the case of morpho-syntactic anomaly, exemplified in (25b). The results of this experiment contribute to a demonstration that the levels of processing tapped by eye-movements distinguish between grammatical and non-grammatical causes of unacceptability. Details are in chapter 3.

The purpose of experiment 3 is two-fold. First, it tests for effects that are cognate with satiation effects, in an on-line task. Snyder (2000) shows that satiation effects can be induced experimentally, a result replicated and extended in Hiramatsu (1998). Anomalous sentence types subject to satiation effects include some, but not all, violations of island-constraints on movement. Snyder speculates that sentences showing a satiation effect are not, in fact, ungrammatical, but rather are unacceptable for processing reasons (also Berwick & Weinberg, 1986; Kluender & Kutas, 1993b). The second purpose of

experiment 3 is to explore this possibility by contrasting island violations that show clear satiation effects in grammaticality judgment tasks with those that clearly did not show satiation. Eye-movement responses to the uncertain contrast in grammaticality represented by the sentence types of experiment 3 are evaluated in the context of the qualitatively different patterns observed for the clearly contrasting sentences (ungrammatical versus grammatical-unacceptable) in experiments 1 and 2. Details are in chapter 4.

Notes

1. The typical person can, nonetheless, assign consistent interpretations to a large (and heterogeneous) class of ungrammatical word strings (Katz, 1964). It remains an open question as to whether this ability is due to operation of the human sentence processing mechanism, or, rather, to general inferencing abilities.

2. Ni et al. had subjects rate sentences for acceptability on a 7 point scale. Participant instructions defined deviations from acceptability as any variation from norms perceptible to the participant.

3. In both Braze et al. and Ni et al., the anomaly types were chosen to represent unequivocal representatives of pragmatic and syntactic anomaly. The experiments discussed in the next chapter examine anomalies of a similarly unambiguous nature.

4. We can certainly envision "judgment" tasks where the response measure is something other than a judgment of acceptability along a single dimension. One possibility is a sentence categorization task, in which the subject is required to sort sentences according to the nature some perceived anomaly. The possibility of such tasks does not, however, obviate the truism that judgments used in theoretical argumentation are, virtually without exception, uni-dimensional in nature.

Chapter 3: Syntactic and Pragmatic Anomalies

3.1 Introduction

I assume that comprehension is not a monolithic process. It is marked by a number of sub-stages of processing which, if not fully modular, are at least sufficiently encapsulated that their individual contributions are readily discernable. The strongest form of this hypothesis has it that grammatical and non-grammatical information are handled by distinct processing mechanisms. Evidence supporting this supposition comes from divergent sources. Broca observed over 100 years ago that damage to particular regions of the brain selectively effect different aspects of linguistic ability (Broca, 1988). More recent experimental work with aphasic individuals confirms that some may suffer a selective loss of the ability to process linguistic structure while semantic and pragmatic knowledge remains relatively intact (Caramazza & Zurif, 1976; Lukatela et al., 1995; Swinney, Zurif, Prather, & Love, 1996). Studies of electrical activity of the brain using the event-triggered potential paradigm have shown that distinct waveforms emerge in response to syntactically and pragmatically anomalous sentences (Kutas & Hillyard, 1983; Neville, Nicol, Barss, Forster, & Garrett, 1991). Brain imaging studies have also shown non-overlapping areas of brain activation in response to

syntactically and pragmatically anomalous sentences (Helenius, Salmelin, Service, & Connolly, 1998; Ni et al., 2000). In the realm of overt behavior, it has long been known that even untrained informants can make judgments of grammaticality and of anomaly. Further, informants can easily distinguish anomaly types, as in the pragmatic anomaly in (26a) and the morpho-syntactic anomaly in (26b).

(26) a. Colorless green ideas sleep furiously.
b. He wear socks under his sandals.

One goal of the experiments in this chapter is to provide evidence that the human sentence processing mechanism (HSPM) shows characteristically distinct responses to grammatical and non-grammatical sources of information in the input. As noted in chapter 2, it is well-established that variations in readers' eye-movements show sensitivity to lexical, structural, semantic, and pragmatic properties of a text (e.g. Rayner, 1998). The results of Braze et al. (in press) and Ni et al. (1998), using contrasting morpho-syntactic and pragmatic anomalies as probes, suggest that the human sentence processing mechanism does exhibit qualitatively different responses to morpho-syntactic and pragmatic information.

It is worth clarifying what is meant by "pragmatic properties," in the present context. Essentially, pragmatics refers to the plausibility of a text. Pragmatic processing, in this sense, involves evaluating how well the

denotation of a sentence corresponds to what the hearer/reader knows about how the world works. This necessarily requires recruiting world knowledge to evaluate the contextual likelihood of sentence meaning. By hypothesis, world knowledge is extra-grammatical in nature, and so cannot be called upon in aid of grammatical processing. Extra-grammatical pragmatic properties of a text contrast with fundamentally grammatical cues, like word order and morphology, available in any well-formed utterance.

A related goal is to clarify how the HSPM's response to (a specific class of) grammatical anomaly differs from it's response to (a specific class of) non-grammatical anomaly, as indexed by eye movement behavior during reading. In furtherance of these goals, the two studies presented in this chapter both replicate and incrementally extend previous findings.

3.2 Previous Work

Using the technique of monitoring eye-movements during reading, Ni et al. (1998) and Braze et al. (in press) contrasted sentences containing flaws in verbal morpho-syntax with other, grammatical sentences, depicting pragmatically odd subject-verb relationships, as shown in (27). As noted, pragmatic processing, in the relevant sense, involves recruiting world knowledge to evaluate the contextual likelihood of relationships denoted by a

sentence. A pragmatic anomaly is generated by a sentence that denotes a low-likelihood relationship, like (27c).

(27) a. The wall will surely crack after a few years in this harsh climate.
b. The wall ***will*** surely ***cracking*** after a few years in this harsh climate.
c. The ***wall*** will surely ***bite*** after a few years in this harsh climate.

Figure 1 shows a breakdown of regressive (leftward) eye-movement frequency by sentence region for each sentence type. It is notable that regressions peak at different locations for the two anomaly types. Syntactic anomalies trigger regressions in the region containing the verb, followed by a return to baseline. Pragmatic anomalies, on the other hand, induce a spate of regressions at the ends of the sentences, but show no increase over controls at the region containing the verb.

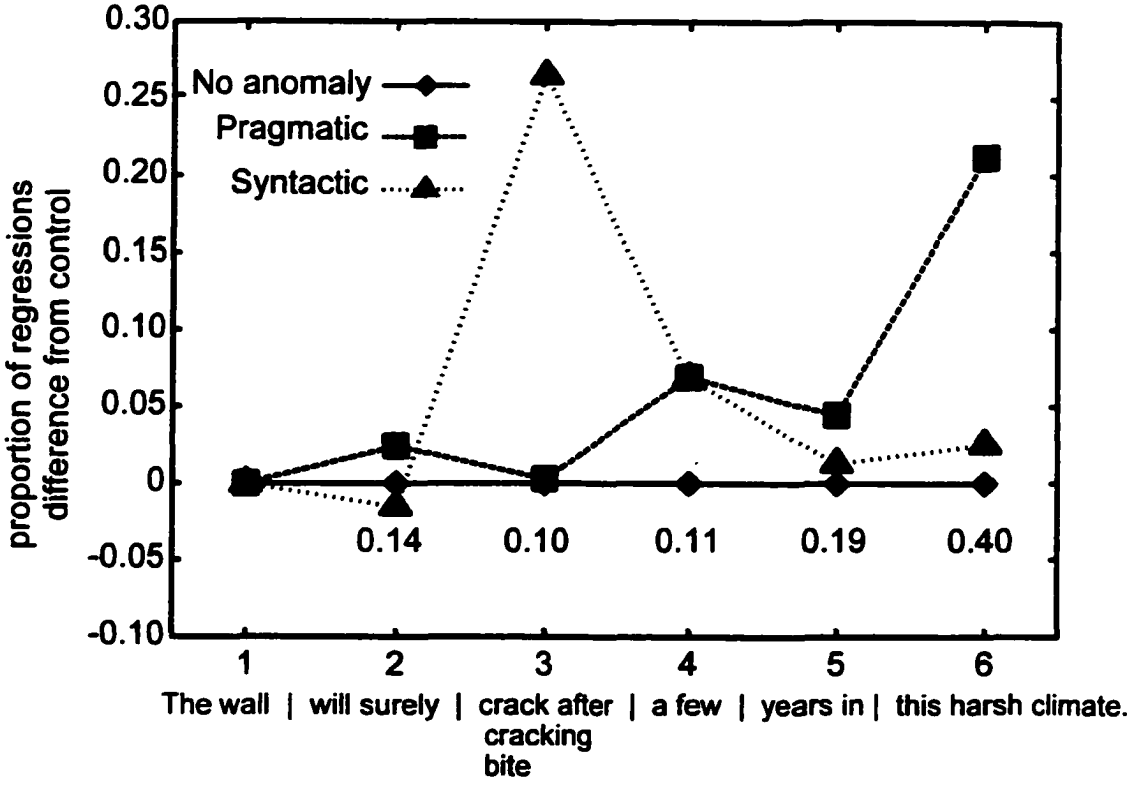

Figure 1: Proportion of regressive eye movements from each sentence region for syntactic and pragmatic anomalies plotted as differences from the non-anomalous control condition (Braze et al., in press).

Regions targeted by regressions launched from the ends of sentences (region 6) were also tabulated, as shown in **Table 1**. Distribution of landing sites for pragmatic anomalies was found to differ from both non-anomalous controls and syntactic anomalies. Most frequent landing sites for regressions associated with pragmatic anomalies were near the beginnings of sentences whereas, for other sentence types, the penultimate region was the most frequent site. Landing site distributions for non-anomalous and syntactically

anomalous sentences do not differ from each other. The split between non-anomalous sentences and syntactic anomalies on the one hand, and pragmatic anomalies on the other, parallels the split seen in the overall frequency of regressive eye-movements from the sentence final region.

Table 1: Distribution of regional landing sites for regressions from sentence region 6 (percent and frequency). Boxes indicate modal landing site for each condition (Braze et al., in press).

Landing	Anomaly Condition					
Site	non-anom.		syntactic		pragmatic	
1	14.1	9	10.2	8	21.7	30
2	17.2	11	25.6	20	29.7	41
3	4.7	3	7.7	6	21.7	30
4	20.3	13	16.7	13	10.1	14
5	43.8	28	39.7	31	16.7	23

Ni et al. (1998) and Braze et al. (in press) give convergent evidence that eye-movement indicators of syntactic and pragmatic processes each display a distinct signature, in keeping with the expectations generated by the ERP and neuro-imaging studies. These findings are consistent with the hypothesis of separate and dissociable processing mechanisms, but, as noted, they do not constitute proof. The present study is intended to address two of these questions. First, are the different regression peaks attributable to differences in degree of anomaly across anomaly types? Second, do the results generalize

beyond the specific types of verb-based anomaly instantiated in the test sentences of Ni et al. (1998) and Braze et al. (in press)?

Experiments 1 and 2 were designed to address these questions. Experiment 1 speaks to the issue of whether degree of anomaly figures into the qualitatively different eye-movement patterns seen in earlier studies. Sentence stimuli are similar to those in (27), but incorporating an additional condition containing simultaneous syntactic and pragmatic anomalies. Experiment 2 extends the paradigm of Ni et al. and Braze et al. to other grammatical and pragmatic relationships using noun-based morpho-syntactic and pragmatic anomalies.

3.3 Method Common to Experiments 1 and 2

The same subjects participated in experiments 1 and 2. Materials for the two experiments were combined in single set of stimulus lists, and therefore materials for one experiment served as filler items for the other. Subjects completed both experiments in a single experimental session that lasted about one hour.

3.3.1 Subjects

Twenty-eight undergraduate students were each paid to participate. All were native speakers of English with vision reported to be normal or corrected to

normal with contact lenses. Participants had no knowledge of the purpose of the experiments and no prior exposure to the test materials.

3.3.2 Apparatus

Eye-movements were recorded with an Iris 6500 eye-tracker from Skalar Medical. The Iris system is a limbus tracker, using infra-red transducers positioned in front of the eye to detect eye-movements. The technique relies on the differential reflectivity of the iris and sclera. The tracker's output is a continuously varying analog signal corresponding to gaze direction. This signal was sampled each millisecond by an Apple computer equipped with a 16 bit analog-to-digital conversion board. The eye-tracker is equipped with a forehead rest and bite-bar mount, necessary to stabilize subjects' head positions in order to minimize artifacts in the eye-gaze signal.

3.3.3 Materials and Design

Materials for experiments 1 and 2 consisted of 28 sentence sets each, for a total of 56 experimental items, described in more detail below. There were four versions of each experimental item. Materials for both experiments were combined in a single set of four stimulus lists. Only one version of each test sentence was included in each list. The purpose of combining materials in this way was to allow presentation of both experiments to each subject in a single

session (each subject sees a single stimulus list). Lists were blocked to ensure that experimental conditions were evenly distributed. Each list contained an equal number of stimulus items of each type. Conditions were counter-balanced within and across lists such that a different version of each experimental item appeared in each list. The complete materials for experiments 1 and 2 are in Appendix A. Stimulus sentences were interspersed quasi-randomly with a like number of foils (fifty-six).

In order to ensure that subjects were attending to the task of reading for comprehension, each critical item was followed by a yes/no comprehension question. After each critical trial, a question appeared on the screen and the subject would respond to the question by using a computer mouse to click on a YES button or a NO button, as appropriate.[1]

3.3.4 Auxiliary Tasks

In addition to the eye-movement task, measures of verbal working memory capacity, reading experience, and reading skill were administered.

Differences in memory capacity have been implicated in the ease with which certain structures are processed (Daneman & Carpenter, 1980; King & Just, 1991; Lukatela et al., 1995; Gibson, 1998; Caplan & Waters, 1999). The test of verbal working memory capacity used here was based on the sentence span task of Daneman and Carpenter (1980). However, unlike most forms of

the sentence span task, an auditory presentation of test materials (rather than print) was used. In this task, subjects listen to increasingly large sets of sentences. After each sentence, they must make a true/false judgment. At the end of each set of sentences, they are prompted to recall as many sentence final words from the set as possible. A subject's working memory score is defined as the total number of sentence final words correctly recounted. Detailed discussion of the task and scoring method can be found in Swainson and Shankweiler (2001).

Differences in linguistic experience have been suggested as one contributor to variation in linguistic knowledge (Stanovich & West, 1989; Stanovich & Cunningham, 1992)[2] and sensitivity to sentence grammaticality in (Ross, 1979). For the population under study, reading experience is plausibly a significant contributor to linguistic experience generally. So, a measure of reading experience, an author recognition task, based on Stanovich and Cunningham (1992) was administered. The test consists of a list of popular writers' names mingled with foils, names of people who are not writers. The subjects' task is to indicate whether they are familiar with the name of a particular writer by putting a check next to it. Test scoring uses a signal detection rational, where the presence of foils guards against the possibility that some subjects might have a low threshold for checking a name as "recognized." A subject's score is determined tallying the number of actual

writers checked and subtracting from that the number of foils (non-writers) checked.

Finally, a standard measure of reading comprehension, the fast reading sub-test of the Stanford Diagnostic Battery (Karlsen & Gardner, 1985) was administered. This is a three minute timed-test in which subjects must read through a short essay, approximately 500 words in length. At thirty places in the story, a word is omitted and the subject must select the word that best fits in that place (from among three predetermined choices). The score is equal to the number of items completed correctly in three minutes.

Scores on each of the auxiliary tasks were consistent with expectations for University of Connecticut undergraduate students. Average scores for the tasks are shown in **Table 2**. There were, however, no consistent relationships between the auxiliary measures and eye movement behavior. For each auxiliary measure, subjects were divided into two groups based on a median split of their scores. These high/low groups were included as factors in analyses of eye movements along with the independent measures of syntactic and pragmatic anomaly. While members of the high memory group were found to be faster readers overall, the data reveal no interaction between memory capacity and sentence type. Reading comprehension and print exposure (author recognition), also failed to show interactions with sentence type. Given

the lack of interesting effects due to the auxiliary measures, they are omitted from subsequent analyses and attendant discussion.

Table 2: Average scores on working memory, author recognition, and reading comprehension tasks for participants in experiments 1 and 2.

Task	mean	sd
Working Memory	34.00	4.61
Author Recognition	9.34	4.66
Reading Comprehension	25.38	4.20

3.3.5 Procedure

Participants were given verbal instructions and a description of the eye-movement method. Because head movements reduce the accuracy of the eye-movement record, individually prepared bite-bars and a forehead rest were used to help stabilize head position. Bite-bar preparation and calibration of the eye-tracker took about 5 minutes for each subject.

Sentences were presented on a computer monitor positioned 64 centimeters from subjects' eyes. Test materials appeared in a 14 point courier font such that each character subtended approximately 12 minutes of visual arc. Viewing was binocular, but eye movements were recorded from the right eye only.

Before the presentation of each sentence, a fixation target appeared at the screen position to be occupied by the first character of the sentence.

Participants were instructed to focus on the target and then to click a mouse button to call up the sentence. Participants were told that they should read the sentence for comprehension, as they normally would, and to click the mouse button when finished. The action of clicking the button served to erase the sentence from the screen. All critical trials were followed by a question which the participant responded to by using the mouse to click YES or NO buttons. The purpose of the question was to ensure that participants were attending to the task of reading for comprehension. Every trial was followed by a brief calibration check. Adjustments were carried out occasionally, as needed.

Each eye-track session lasted about 45 minutes and included two programmed breaks, but subjects were allowed to take additional breaks as often as they wished. Including time to complete the auxiliary tasks, each experimental session lasted about 70 minutes.

3.3.6 Methods of Analysis

To examine the running record of eye movements as a sentence unfolds, the procedure adopted initially divides each target sentence into six regions of approximately two words each, as shown in (28), for the purpose of aggregating measures of reading performance. For each sentence frame, corresponding regions contain the same numbers of words in all six versions.

(28) a. The daisies●were slowly●wilting in●the hot●weather this●afternoon.
b. Jim saw●two black●puppies chasing●a cat●in the●yard this morning.

In experiment 1, (28a), region 1, the subject noun phrase, is one or two words long. Region 2 is a modal verb followed by an adverb. Region 3, the critical region, is the main verb and the word following. Regions 4 and 5 also contain two words each. Region 6, the sentence final region, contains one to three words. In experiment 2, (28b), region 1, a subject and main verb, is two or three words long. Region 2 is made up of either a determiner or quantifier followed by an adjective. Region 3, the critical region, contains a noun and the word following it. Regions 4 and 5 also contain two words each. Region 6, the end of the sentence, is from one to three words long.

In addition, a second analysis focuses on the portion of the sentence that initiates the anomaly. For experiment 1 this is the verb in region 3. For experiment 2 it is the noun in region 3. As the pivot-points of anomaly, these words are of special interest. So, in addition to analyses based on the two-word regions shown in (28), I present separate analyses based on one-word regions consisting of the verb/noun, plus the three succeeding words. In these analyses, measures of reading performance are aggregated for each word individually.

Data were analyzed in terms of first-pass reading times and incidence of first-pass regressive eye-movements, tabulated separately for each sentence region. First-pass reading time is the summed fixation durations within a region, beginning with the first fixation inside the region and ending with, but

not including, the first subsequent fixation outside the region. Reading times were statistically adjusted to compensate for inequalities in verb length across anomaly conditions. The correction was calculated separately for each subject, using a linear regression with region length (number of letters and spaces) as regressor and first-pass reading time as dependent variable. Deviations from predicted reading times were used as the length-corrected reading time measure in subsequent analyses (Trueswell, Tanenhaus, & Garnsey, 1994).

A region is counted as having a first-pass regression if its final first-pass fixation ends in a backward glance to an earlier part of the sentence. Within region leftward eye-movements are not classified as first-pass regressions. The number of times each region is revisited was also tabulated. Since an initial regressive saccade is often followed by additional regressions resulting in a chain of backward eye-movements, or a regression path (Konieczny et al., 1997), the regression target is counted as the left-most sentence region visited during such a chain. Both first-pass reading time and first-pass regression frequency are contingent upon there being at least one fixation of 50 ms or longer within the region (Carpenter & Just, 1983). Regions not meeting this criterion are excluded from the analyses.

3.4 Experiment 1

Experiment 1 contrasts two hypotheses about the timing of peak frequencies of regressive eye-movements in pragmatically and syntactically deviant sentences. Following Fodor (1983), I hypothesize that different mechanisms are involved in processing grammatical and non-grammatical information. However, a plausible alternative explanation is that the observed difference in the placement of regression peaks is due to differences in degree of anomaly for each sentence type.

3.4.1 Materials and Design

The materials used here are similar to those of Braze et al. (in press). The primary difference is the inclusion of a double anomaly condition, permitting a fully crossed design (syntactic anomaly X pragmatic anomaly). The purpose of combining the two types of anomaly within a single condition is to test whether effects due to each anomaly type remain distinguishable. Twenty-eight sentence sets based on the model in (29) were used as stimuli: (29a) is a non-anomalous control sentence, (29b) is syntactically anomalous, (29c) is pragmatically anomalous, and (29d) contains both syntactic and pragmatic anomalies.

(29) a. The daisies were slowly wilting in the hot weather this afternoon. NA
b. The daisies had slowly wilting in the hot weather this afternoon. SA
c. The puddles were slowly wilting in the hot weather this afternoon. PA
d. The puddles had slowly wilting in the hot weather this afternoon. DA

3.4.2 Predictions

Braze et al. and Ni et al. found that readers responded with characteristically distinct patterns of eye-movements to syntactic and pragmatic anomalies in printed sentences. Specifically, an immediate peak in regressive eye-movements was found in response to syntactic anomaly, and a delayed peak in regressions for sentences with pragmatic anomalies, as shown in **Figure 1**. It was argued that this pattern of results follows from the modularity hypothesis (e.g. Fodor, 1983). A dissociation in the modal effects of syntactic and pragmatic manipulations follows from the premise that distinct processing mechanisms are engaged by syntactic and pragmatic information in the linguistic input. The precedence of the syntactic effect follows from the premise that syntactic processing is automatic and reflex-like, whereas pragmatic processing is not.

If separate mechanisms are responsible for handling each type of information, then we should find two peaks in regressions for sentences containing both syntactic and pragmatic anomalies, an early peak corresponding to that of syntactic anomaly, and a late peak corresponding to that of pragmatic anomaly. Alternately, the degree of anomaly hypothesis predicts a single peak in regressive eye-movement frequency for the double anomaly condition, parallel to the syntactic anomaly condition. In this case, we

might also expect that the peak for the double anomaly condition should be of greater magnitude than that for the syntactic anomaly condition.

3.4.3 Results

Figure 2 shows a plot of regional regression frequencies for each sentence type. For the non-anomalous controls, syntactic, and pragmatic anomaly conditions, the pattern replicates that of Braze et al. shown in **Figure 1**). At sentence region 3 syntactic anomalies have far more regressions than control sentences [$F1(1,27) = 47.75$, $p < .0001$; $F2(1,27) = 45.83$, $p < .0001$] or pragmatic anomalies [$F1(1,27) = 41.60$, $p < .0001$; $F2(1,27) = 39.06$, $p < .0001$]. Of greater interest is the fact that double anomalies also show an increase in regressive eye-movements from the region containing the verb. Doubly anomalous sentences induce more regressions than either controls [$F1(1,27) = 30.03$, $p < .0001$; $F2(1,27) = 27.98$, $p < .0001$] or pragmatic anomalies [$F1(1,27) = 24.99$, $p < .0001$; $F2(1,27) = 22.66$, $p < .0001$]. The increase is slightly less pronounced than that for syntactic anomalies.

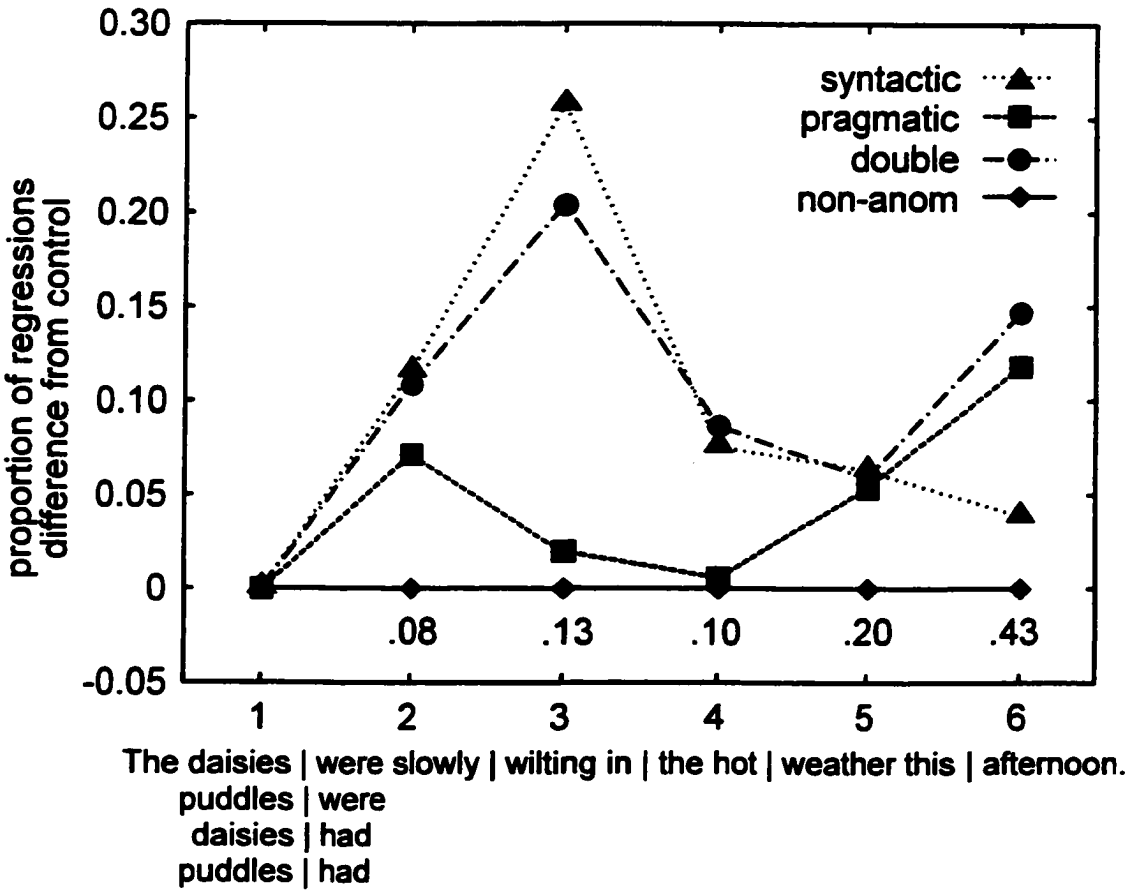

Figure 2: Proportion of regressive eye movements from each sentence region for syntactic, pragmatic, and double anomalies plotted as differences from the non-anomalous control condition (experiment 1, verb-based anomalies).

At sentence region 6 pragmatic anomalies induce a somewhat greater proportion of regressions than controls [$F1(1,27) = 2.86$, $p < .1$], as do doubly anomalous sentences [$F1(1,27) = 5.71$, $p < .05$]. Double anomalies also induce more regressive eye-movements than sentences with syntactic anomalies alone [$F1(1,27) = 4.13$, $p < .05$].

The number of times each region is revisited was also tabulated. No reliable differences in regression targets emerged for regions one through five, but regressions originating in sentence region six show interesting differences in target preference, as can be seen in **Table 3.**

Table 3: Distribution of regional landing sites for regressions from sentence region 6 (percent and frequency). Boxes indicate modal landing site for each condition (experiment 1, verb-based anomalies).

Landing	Anomaly Condition							
Site	non-anom.		syntactic		pragmatic		double	
1	14.3	8	17.0	10	32.9	23	12.7	8
2	21.4	12	30.5	18	18.6	13	41.3	26
3	8.9	5	5.1	3	14.3	10	12.7	8
4	21.4	12	20.3	12	11.4	8	11.1	7
5	33.9	19	27.1	16	22.9	16	22.2	14

For non-anomalous control sentences, sentence region 5 is the modal landing site for regressions originating in region 6. In the case of syntactic anomalies, the modal landing site is sentence region 2. The difference between controls and syntactic anomalies is not, however, statistically significant. On the other hand, landing site distribution for pragmatic anomaly differs from both controls [$\chi^2(4) = 8.57$, $p = .073$] and syntactic anomaly [$\chi^2(4) = 9.63$, $p < .05$]. The target distribution for double anomalies also differs reliably from pragmatic anomalies [$\chi^2(4) = 11.68$, $p < .05$].

3.4.4 Discussion

In experiment 1, the observed patterns of eye-movements for syntactic and pragmatic anomalies confirm the results of earlier work, both in terms of regression frequency and target distribution. In keeping with our prediction, based on the hypothesis that separate mechanisms are responsible for processing grammatical and non-grammatical information, doubly anomalous sentences show two peaks in regressive eye-movements relative to non-anomalous sentences. An early peak, consistent with syntactic anomaly, is followed by another at sentence end, consistent with pragmatic anomaly.

Regressive eye-movements have been interpreted as indicators of the processor's inability to incorporate material on-line (Crain et al., 1996). Analysis of regressions in the present study supports this hypothesis. Regression from the ends of sentences, in response to pragmatic incongruity at the verb, shows the persistence of the processing difficulty imposed by this type of anomaly. Sixty-seven percent of regressions from the sentence final region of pragmatically anomalous items landed in the first three regions (**Table 3**). This is the informative portion of the sentence, where the mismatch occurs between the subject noun (*puddle*) and the verb (*wilt*). Especially note that the frequency of regressions targeting the subject and verb containing regions (regions 1 and 3, respectively) are elevated compared to either non-anomalous controls or syntactic anomalies. Braze et al. (in press) reported a

similar result, and analysis of regression targets in several other studies has shown readers to be adept at consulting that portion of text where comprehension breaks down (Frazier & Rayner, 1982; Kennedy, 1983; Kennedy & Murray, 1987).

3.5 Experiment 2

The purpose of experiment 2 is to test the generality of the patterns seen in previous work, where verb-based anomalies were used. Noun-based anomalies, similar to the verb based anomalies of experiment 1, serve as test items. See the following sub-section for details. Materials incorporate both pragmatic anomalies and syntactic anomalies. The predicate relationship between adjective and noun is exploited to generate pragmatic anomalies, contrasting with the subject/verb pragmatic anomalies of experiment 1. The morphological dependency between determiner/quantifier and noun is manipulated to generate syntactic anomalies, contrasting with the auxiliary verb/main verb morpho-syntactic anomalies of experiment 1.

3.5.1 Materials and Design

Twenty-eight sentence sets on the model depicted in (30) were used as stimuli. The design of experiment 2 differs from that of experiment 1 insofar as syntactic and pragmatic anomaly conditions are not crossed. There is no

double anomaly condition in experiment 2. The three conditions of primary interest are: a non-anomalous control, (30a); a syntactic anomaly, (30b); and a pragmatic anomaly, (30c). The design parallels that of Ni et al. (1998), but with the addition of an additional non-anomalous condition, (30d). The sentence type represented by (30d) differs from the true control condition in that the critical noun and associated determiner are singular, rather than plural. The non-anomalous singular condition was included primarily to simplify blocking of materials across experiments 1 and 2. Thus, this condition serves as a "filler" of sorts. Nonetheless, a potentially interesting comparison between the NA condition with NA_{sg}, particularly at the region containing the critical noun, is made possible by its presence.[3]

(30)
a.	Toni found two antique bottles on a shelf in the back of the shed.	NA
b.	Toni found an antique bottles on a shelf in the back of the shed.	SA
c.	Toni found two anxious bottles on a shelf in the back of the shed.	PA
d.	Toni found an antique bottle on a shelf in the back of the shed.	NAsg

Across sentence sets, adjectives in the pragmatic anomaly condition, e.g. *anxious*, were frequency matched with adjectives not invoking a pragmatic anomaly, e.g. *antique* (Francis & Kucera, 1982). Mean frequencies for each set of adjectives are shown in **Table 4**. Frequency of adjective pairs within each sentence set are also well matched (r = .95). Frequency of each adjective is listed in Appendix A.

Table 4: Mean frequencies for pragmatically anomalous and non-anomalous adjectives in experiment 2 (Francis and Kučera, 1982).

Adjective Set	mean	sd
pragmatic anomaly	22.29	29.94
no pragmatic anomaly	21.57	39.12

3.5.2 Predictions

The most straightforward prediction is that eye-movements evoked in response to the noun-based syntactic and pragmatic anomalies instantiated in this experiment should replicate the general patterns observed in response to the verb-based syntactic and pragmatic anomalies of experiment 1 and of Ni et al. (1998). This is the outcome we expect if the syntactic anomaly derived from a determiner/noun mismatch (*an antique bottles*) elicits a qualitatively similar response to the auxiliary verb/main verb mismatch of experiment 1 (*had slowly wilting*), and similarly, if the processes invoked by pragmatic anomaly derived from an adjective/noun mismatch (*two anxious bottles*) parallel those due to the subject noun/verb mismatch (*puddles were slowly wilting*) of experiment 1.

It should be noted that a potential confound exists in the case of the two types of pragmatic anomalies. The two words that combine to create the verb-based pragmatic anomalies are separated by two words, while the two words

combining to create the noun-based pragmatic anomalies are adjacent to one another.

3.5.3 Results

In sentence region 3, syntactic anomalies generate more regressive eye-movements than either non-anomalous controls [$F1(1,27) = 6.89$, $p = .01$; $F2(1,27) = 8.36$, $p < .01$] or pragmatically anomalous sentences [$F1(1,27) = 5.54$, $p = .02$; $F2(1,27) = 5.67$, $p = .02$]. Counter to expectations, pragmatic anomalies do not show an increase in regressions at sentence end.

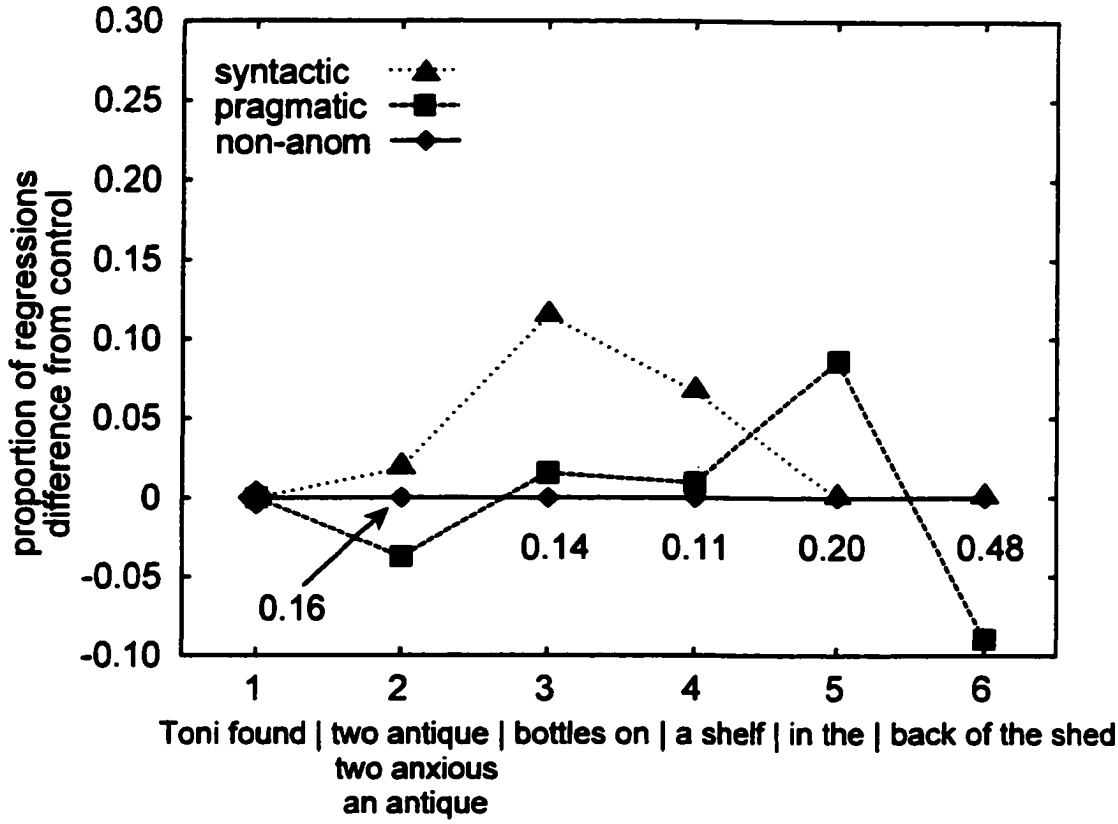

Figure 3: Proportion of regressive eye movements from each sentence region for syntactic and pragmatic anomalies plotted as differences from the non-anomalous control condition (experiment 2, noun-based anomalies).

Table 5: Distribution of regional landing sites for regressions from sentence region 6 (percent and frequency). Boxes indicate modal target region for each condition (experiment 2, noun-based anomalies).

Landing	Anomaly Condition					
Site	non-anom.		syntactic		pragmatic	
1	10.6	7	26.2	16	11.3	6
2	24.2	16	26.2	16	41.5	22
3	18.2	12	19.7	12	20.8	11
4	1.5	1	11.5	7	9.4	5
5	45.4	30	16.4	10	17.0	9

Target distribution for controls differs from that of both syntactic and pragmatic anomalies [$\chi^2(4) = 17.86$, $p = .001$; $\chi^2(4) = 13.79$, $p = .008$], which do not differ from each other.

3.5.4 Discussion

Noun-based morpho-syntactic anomalies show an early peak in regressions that quickly subsides. This pattern is similar to that seen with verb-based syntactic anomalies. On the other hand, the data show no sentence ending peak in regressive eye-movement frequency for noun-based pragmatic anomalies. This contrasts with the situation for verb-based anomalies. However pragmatic anomaly does influence regression target preference, consistent with verb-based pragmatic anomalies.

Experiment 2, using noun-based anomalies, does show a dissociation in the processor's use of syntactic and pragmatic information in the input, although details differ somewhat from the response to verb-based anomalies.

3.6 General Discussion

In confronting the sentence processor with form-based and meaning-based anomalies, I have shown that the eye-movement indicators of syntactic and pragmatic processes each display a distinct signature, in keeping with the expectations generated by ERP and neuroimaging studies, as well as the earlier eye-movement findings of Ni et al. (1998) and Braze et al. (in press).

The dual peak in regressive eye-movement frequency seen in the doubly anomalous sentences of experiment 1 supports the hypothesis of independent processing mechanisms for grammatical and non-grammatical information. The pattern of regressions seen in experiment 2, for noun-based morpho-syntactic anomaly, provisionally supports the generality of the syntactically induced early peak found in previous studies using verb-based anomalies. The similarity in eye-movement patterns in response to syntactic anomaly across experiments 1 and 2 lends support to our supposition that grammatical information is subserved by a specialized processing mechanism.

In summary, this study confirms that eye-movement measures are sensitive to constraints invoked by a range of morpho-syntactic and pragmatic

anomalies. Further, it confirms that the sentence processor's immediate and subsequent responses to the two anomaly categories are distinct. Morpho-syntactic anomalies, whether noun-based or verb-based, invoke similar responses. Pragmatic anomalies seem to place less constraint on the sentence processor's response.

Notes

1. No subject made more than seven errors on the fifty-six comprehension items.

2. Stanovich and Cunningham are concerned particularly with reading experience. Although reading experience is but one facet of linguistic experience, it is arguably a significant component of general linguistic experience, especially in the study population, university students.

3. In fact, no significant difference between the two non-anomalous conditions surfaced in either first pass reading times or incidence of regressive eye movements. Therefore, the NA_{sg} condition was dropped from subsequent analyses, and the results discussed later in this section focus on differences among the true control, syntactic anomaly, and pragmatic anomaly conditions.

Chapter 4: Satiation Effects

4.1 Introduction

Judgment satiation, or syntactic satiation, refers to the fact that the intuitive response to some initially unacceptable types of sentences reflects an increase in acceptability over time (with repeated exposure). The earliest reports of this effect are anecdotal, but Snyder (2000) shows that judgment satiation for certain sentence types can be elicited under experimental control, and with remarkably few exposures. He notes that the class of satiable violations overlaps significantly with classic subjacency violations, and suggests that the satiability of such violations is consistent with a processing-based explanation of the subjacency constraint on long-distance dependencies (for processing accounts see Berwick & Weinberg, 1986; Kluender & Kutas, 1993b).

This chapter describes an experiment designed to test for an on-line counterpart of the judgment satiation effect in the course of reading for comprehension. The technique of monitoring eye-movements during reading is used to determine whether eye-movement patterns while reading unacceptable, yet satiable, sentences change, relative to non-island-violation controls, as subjects read a series of such sentences. Snyder found that *wh*-argument extraction from *whether* complement clauses showed clear satiation

effects, and so this sentence type was selected for the present study. *Whether* complements to *wonder*, and other verbs which take indirect questions as complements, are generally labeled *wh*-islands. However, following Snyder, I will adopt the more specific term *whether*-island. A parallel condition involving a non-satiable violation, argument extraction from adjunct islands, is included for comparison.

4.2 Judgment Satiation Effects

As noted, the term "judgment satiation" refers to the observation that, over time and with repeated exposure, some types of sentences, initially deemed to be degraded in some way, become increasingly acceptable. In a recent study, Snyder (2000) asked subjects to rate *wh*-questions like those in (31) as to whether they were "grammatical" or not -- a binary forced choice. Importantly, context sentences were provided so that each question would be pragmatically sensible, even if not grammatical. For example, (32) provides a felicitous context for (31b) in the sense that it introduces the relevant participants and the relationships among them..

(31)	a.	Who does John want for Mary to meet *t* ?	*want- for*
	b.	Who does John wonder whether Mary likes *t* ?	*whether*-island
	c.	Who does Mary think that *t* likes John?	*that*-trace
	d.	What does John know that a bottle of *t* fell on the floor?	subj-island
	e.	Who does Mary believe the claim that John likes *t* ?	CNPC
	f.	Who did John talk with Mary after seeing *t* ?	adjunct-island
	g.	How many did John buy *t* books?	left-branch

(32) John wonders whether Mary likes Bill.

Snyder's study shows that after exposure to as few as five *whether*-islands, like (31b), or CNPC violations, like (31e), many subjects begin to find these types of sentences increasingly acceptable. Additionally, he reports some indication of satiation for subject island violations, like (31d) -- a statistically marginal effect. It is important to note that judgment satiation does not seem to be a general consequence of the grammaticality judgment task. Other classes of unacceptable sentences that Snyder tested did not become more acceptable with repeated exposure. Neither does it seem that the possibility of satiation is associated in any simple way with the initial degree of acceptability of a sentence type. For example, *that*-trace violations, as in (31c), were comparable in initial acceptability to *whether*-island violations, yet showed no sign of satiation.

Hiramatsu (2000) extends Snyder's results in two ways. In the first of two experiments, she addresses the question of whether satiation effects can be elicited for subject islands, if participants are exposed to a greater number of them than was the case in Snyder's study. She uses materials similar to those of Snyder, including the same variety of sentence types, but exposes subjects to 7 of each type, 2 more than Snyder used. Her results show that judgment satiation does occur with subject island violations, if participants see a sufficiently large number of them.

In a second experiment, Hiramatsu addressed the question of whether adjunct extraction from *whether*-islands would show satiation effects. As is well-known, both argument and adjunct extraction from *wh*-islands yields a sentence that is unacceptable to some degree, although adjunct extraction is generally felt to generate a greater degree of unacceptability than argument extraction. The greater unacceptability of adjunct extraction from the *wh*-island context can be attributed to the fact that, in addition to whatever constraint is responsible for the unacceptability of (31b), (perhaps subjacency, a grammatical constraint on movement, but read on for an alternative explanation), adjunct extraction also violates an additional constraint which requires gap positions to be governed, as per the empty category principle (ECP) (Chomsky, 1981), or the more general condition on extraction domains (CED) (Huang, 1982). However, in both Snyder's study and Hiramatsu's first experiment, the extracted element was always an argument. Materials for Hiramatsu's second experiment included both argument extraction cases, as in (33), and adjunct extraction cases, as in (34). Hiramatsu predicted that the adjunct extraction cases would not induce judgment satiation, on the assumptions that (a) satiation is a property of non-grammatical constraints, and not of grammatical constraints, (b) an explanation is available for the unacceptability of (33) that does not appeal to any grammatical constraint, and (c) that the ECP/CED is a grammatical constraint. The outcome of her

experiment generally support that prediction. Adjunct and argument extraction from *whether*-islands do contrast in the possibility for satiation. Adjunct extraction does not allow for judgment satiation, while argument extraction does.

(33) context: Yesterday, Kelly wondered whether Tina had read "Amistad."
question: What did Kelly wonder whether Tina had read?

(34) context: Gary wondered whether, this year, Jordan would learn French.
question: When did Gary wonder whether Jordan would learn French?

Snyder (2000) speculates that those unacceptable sentence types subject to satiation may, in fact, represent violations of a processing constraint, rather than a constraint on grammatical representations. If this is so, we expect that it may be possible to see correlates of judgment satiation reflected in an online measure with demonstrated sensitivity to processing constraints. The record of eye-movement patterns during reading is such a measure.

4.3 Satiation On-line

Experiment 3 was designed, primarily, to test whether an on-line analog to judgment satiation can be found in the eye-movement record as subjects read *whether*-island violations (involving argument extraction), a sentence type that

evoked satiation effects in the judgment tasks of Snyder (2000) and Hiramatsu (2000). The satiable *whether*-island violations are contrasted with adjunct island violations, which failed to show satiation effects in the sentence judgment studies.[1] Specifically, this study evaluates whether the difficulty of each type of island-constraint violation changes as subjects progress through an experimental session. For each type of island violation, processing difficulty is evaluated relative to appropriate control sentences, which contain no island violation.

The outcome of experiment 3 provides strong evidence for an on-line analog to the judgment satiation effect. On-line satiation surfaces in the form of a change in difficulty for *whether*-island violations, and no corresponding change in difficulty for adjunct island violations, as gauged by reading time and incidence of regressive eye movements (relative to controls, in both cases). A consolidated account of judgment satiation and its on-line analog is provided, in the context of the Referential Model of sentence processing (Crain & Steedman, 1985; Altmann & Steedman, 1988; Ni, 1991), appealing especially to specific properties of verbal working memory (Baddeley, 1986; Crain, Shankweiler, Macaruso, & Bar-Shalom, 1990; Baddeley, 1996).

Secondarily, given that eye movement patterns in response to the two island-constraint violations differ from one-another (independent of satiation-cognate effects), I relate that difference to the contrasting eye movement

patterns evoked in response to the morpho-syntactic and pragmatic anomalies of experiments 1 and 2. Those experiments suggest that the modal response to grammatical anomalies tends to be fast and early, while the modal response to non-grammatical anomalies may be delayed. A delayed modal response for *whether*-island violations relative to adjunct island violations will tend to support the position that the former instantiate a non-grammatical constraint.

As noted in chapter 2, eye-movement measures have a well-established sensitivity to a wide range of linguistic phenomena (Rayner, 1998), and island constraints are among these. For example, Traxler and Pickering (1996, experiment 2) had subjects read temporarily ambiguous sentences like those in (35), in which a potential association between a gap and its associated filler could be ruled out if the parser were sensitive to an island context.

(35) a. We like the book that the author wrote unceasingly and with great dedication about while waiting for a contract.
b. We like the city that the author wrote unceasingly and with great dedication about while waiting for a contract.
c. We like the book that the author who wrote unceasingly and with great dedication saw while waiting for a contract.
d. We like the city that the author who wrote unceasingly and with great dedication saw while waiting for a contract.

Their data show that, in the case of non-island control sentences, (35a-b), subjects initially associate the relative clause head (*book/city*) with a potential gap position following the verb *wrote*, even though this analysis ultimately turns out to be incorrect. On the other hand, their data indicate

that no such association is formed in island contexts like (35c-d), suggesting that such contexts preclude the formation of spurious filler-gap associations. This contrast provides evidence in support of the parser's rapid sensitivity to island constraints, and the registration of such sensitivity in the pattern of eye-movements during reading.[2]

Experiment 3 exploits this sensitivity to explicitly compare patterns of eye-movements elicited in response to different types of island violations. In order to have a baseline against which to compare the eye-movements elicited by each type of island constraint, the experiment uses paradigms like those depicted in (36) and (37) , modeled after the *whether*-island condition, (31a), and the adjunct-island condition, (31d), of Snyder (2000). These two sentence types were selected because, in previous studies, *whether*-island violations showed the most robust satiation effects, while adjunct-island violations proved completely resistant to them, thus providing a clear contrast. Sentences (37a) and (36a) furnish pragmatically felicitous contexts both for the acceptable baselines, (37b) and (36b), and for the island-constraint violations, (37c) and (36c).[3]

(36) a. After the meeting, Valerie wondered whether Don had liked the story.
b. When did Valerie wonder whether Don had liked the story?
c. What did Valerie wonder whether Don had liked?

(37) a. On Saturday, Fraser cleaned the bathroom while Judy mopped the kitchen.
b. When did Fraser clean the bathroom while Judy mopped the kitchen?
c. What did Fraser clean the bathroom while Judy mopped on Saturday?

4.4 Predictions

Of central interest is whether an eye-movement analog can be demonstrated, parallel to the judgment satiation effect reported in Snyder (2000) and Hiramatsu (2000). In the eye-movement task, an analog to judgment satiation should surface as a change in the perturbations of eye-movements attributable to *whether*-island violations, relative to control sentences not containing island-violations, over the course of an experimental session. Conversely, perturbations due to adjunct island violations, relative to their controls, should not change over the course of a session. These predictions are tested by conducting analyses of co-variance at each sentence region, beginning with the earliest point at which an island violation could be detected through sentence end. These analyses incorporate the factor of island-violation (island-violation, control) and the covariate block position (position of item in sequence of presentation).

As a matter of secondary interest, I speculate that the patterns of divergent eye-movements elicited by *whether*-island and adjunct island violations may differ in ways that parallel the dichotomy that emerged for the

grammatically contrasting sentence types in experiments 1 and 2 (also Ni et al., 1998; Braze et al., in press). More precisely, I hypothesize that eye-movement indicators of processing difficulty may show an early peak in the case of adjunct-islands, and a relatively delayed peak in the case of *whether*-islands. In analogy with the experiments contrasting morpho-syntactic and pragmatic violations, a qualitative difference of this sort in eye-movement patterns would provide some support for the non-grammatical basis of the *whether*-island constraint.

4.5 Method

4.5.1 Subjects

Twenty-eight undergraduate students were paid to participate. All were native speakers of English with vision they reported to be normal or corrected to normal with contact lenses. No individual participating in experiments 1 and 2 participated in experiment 3. Participants had no knowledge of the purpose of the experiment and no prior exposure to the test materials.

4.5.2 Apparatus

Please see the description of the eye-tracker in the preceding chapter.

4.5.3 Auxiliary Tasks

In addition to the eye-movement task, measures of reading experience (author recognition), and reading skill (reading comprehension) were administered. Due to time constraints imposed by the length of the eye-tracking materials used in experiment 3, however, the memory span task administered to participants in experiments 1 and 2 was not given to participants in experiment 3. See §3.3.4 for complete descriptions of the author recognition and reading comprehension tasks. Average scores for each task are shown in **Table 6**.[4]

Table 6: Average scores on author recognition, and reading comprehension tasks for participants in experiment 3.

Task	mean	sd
Author Recognition	17.84	8.07
Reading Comprehension	27.17	2.81

The relations between the auxiliary measures and eye movement behavior was tested as described in §3.3.4. For each auxiliary measure, subjects were divided into two groups based on a median split of their scores. High/low groups were included as factors in analyses of eye movements. As was the case in experiments 1 and 2, reading comprehension and print exposure (author recognition), failed to interact with sentence type.

4.5.4 Materials and Design

The materials consist of fifty-six matched pairs of test sentences: twenty-eight *whether*-island violation / control pairs, as shown in (36), and twenty-eight adjunct-island violation / control pairs, as in (37). Hence, there are two versions of each stimulus item, one incorporating an island violation, with the other being a matched control sentence with no island violation. Initially, two stimulus lists were constructed based on these materials. Each stimulus list contained one sentence from each matched pair, for a total of fifty-six test items in each list. Thus, each list contained an equal number (fourteen each) of (a) *whether*-island violations, (b) *whether*-island controls, (c) adjunct-island violations, and (d) adjunct island controls. Each list also included fifty-six foils, one-half of which were ungrammatical in a variety of ways.

Within the lists, items were arranged in fourteen blocks to ensure that item types were evenly distributed throughout. Each block contained: one *whether*-island violation, one *whether*-island grammatical control, one adjunct island violation, one adjunct island grammatical control, and four foils (two of which were ungrammatical).

Finally, two additional stimulus lists were created by reversing the block order of the original lists. In total, four stimulus lists were created.

To ensure that participants were reading for comprehension, one-half of the foils in each list (25% of the total items in each list) were followed by a

possible answer to the question. Participants were asked to respond as to whether the answer was a TRUE answer or a FALSE answer (true and false answers occurred with equal frequency). No participant missed more than four of the twenty-eight comprehension probes.

In order to verify that the materials designed for the eye-movement study were adequate to elicit judgment satiation effects, forty-two subjects judged the grammaticality of these sentences, including foils. Presentation lists were as described above. In the lists used for the judgment task, however, there was an error in counterbalancing such that two presentation lists contained 9 adjunct island violations and 19 *whether*-island violations (rather than 14 of each), while the other two lists contained 19 adjunct island violations and 9 *whether*-island violations. Therefore, only the first 9 of each violation type were included in the judgment task analysis. The counter-balancing error was corrected before carrying out the eye-movement component of the study.

Prior to analysis, any of the forty-two judgment satiation participants who rejected more than one-third of the grammatical control sentences for a given island type was excluded from the analysis of that island type. This resulted in the exclusion of seven subjects from the analysis of *whether*-islands, and twenty-six subjects from the analysis of adjunct islands. Thus, thirty-five

subjects were included in the analysis of *whether*-islands, and sixteen were included in the analysis of adjunct islands.

Judgment satiation effects were evaluated in the same manner as in Snyder (2000), but using nine blocks rather than five. For each island type, the number of subjects with a greater number of YES responses in blocks 6 through 9 than in blocks 1 through 4 was tallied. These subjects were considered to have satiated. The number of subjects with the converse pattern was also tallied (more YES responses in blocks 1 through 4 than in 6 through 9). A sentence type is considered to have satiated if the number of subjects with an increase in YES responses was significantly greater than the number of subjects with a decrease in YES responses for that sentence type.

For *whether*-islands, 21 subjects shifted toward greater acceptance, and 4 shifted toward lesser acceptance (the rest did not change the rate of acceptance). The difference is significant by sign test ($p=0.0036$). A clear satiation effect is in evidence for *whether*-islands. In the case of adjunct islands, only 4 subjects showed any change in their responses, all in the direction of greater acceptance. By sign test the difference is marginal ($p=.0625$), but the low proportion of subjects who showed any change in response dictates that this result must be interpreted with caution. Overall, the results indicate that the materials are adequate to induce selective

judgment satiation parallel to that reported in Snyder (2000) and Hiramatsu (2000).

4.5.5 Methods of Analysis

The general techniques used for data analysis were the same as for experiments 1 and 2, described in chapter 3. Sentences were divided into regions as shown in (38) and (39).

(38) *wh*-island regions
a. When did●Valerie●wonder●whether●Don●had liked●the story?
b. What did●Valerie●wonder●whether●Don●had liked?

(39) adjunct-island regions
a. When did●Fraser●clean●the bathroom●while●Judy mopped●the kitchen?
b. What did●Fraser●clean●the bathroom●while●Judy mopped●on Saturday?

The dependent measures reported in this experiment are the same as in experiments 1 and 2, first-pass regional reading time, and incidence of first-pass regressive eye movements. See chapter 3 for definitions and discussion.

4.6 Results

For each island type, an analysis of covariance incorporating the factor island-violation (island violation, grammatical control) and the covariate block

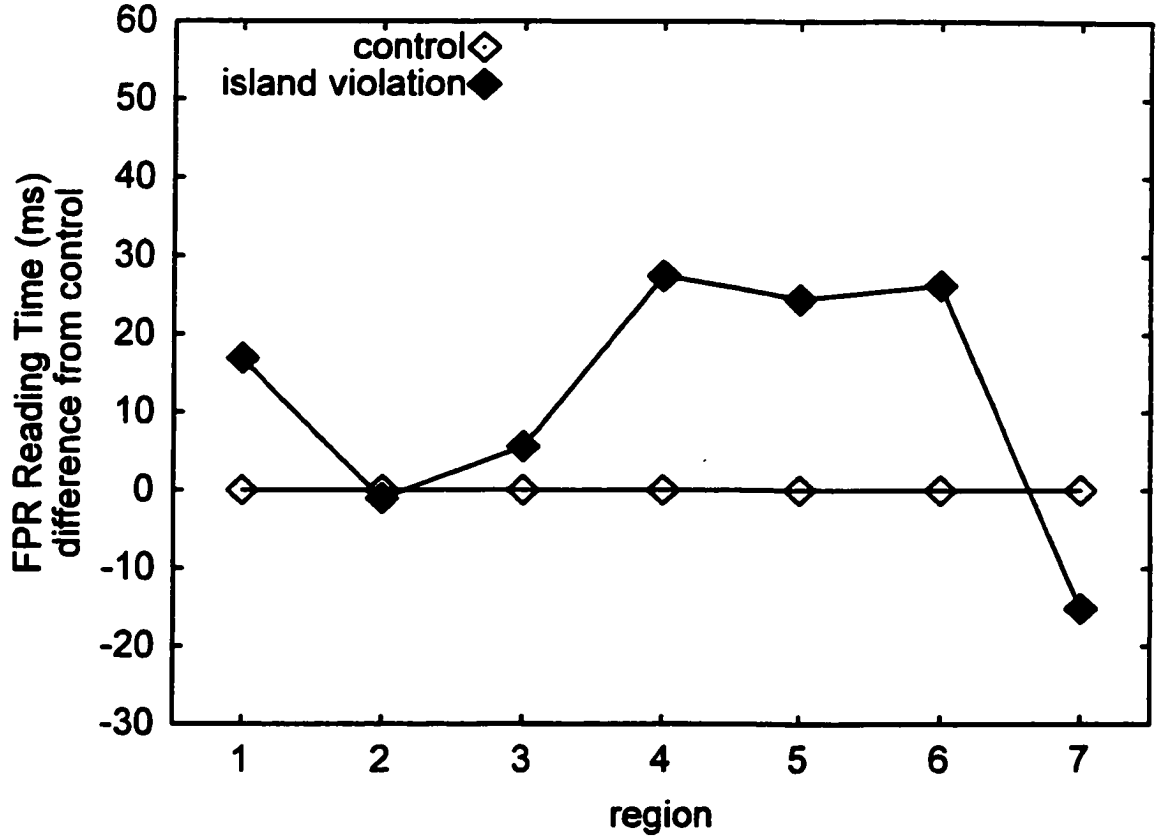

Figure 4: Mean regional length-corrected reading times for adjunct island violations plotted as differences from the non-island-violation control condition (experiment 3).

position (position of item in sequence of presentation) was carried out at each sentence region.

Main effects of island-violation are in evidence for both *whether*-islands and adjunct islands. In the case of adjunct islands, the effect surfaces in both reading times and regressive eye movements. **Figure 4** shows modest reading time differences at region 4 (object of the main clause) [$F1(1,27)=7.53$ $p<.01$; $F2(1,27)=4.46$, $p<.05$] and region 5 (*while*) [$F1(1,27)=8.85$ $p<.01$; $F2(1,27)=11.47$, $p<.01$]. At sentence region 6, reading times for island

violations are numerically greater than reading times for controls, but since variances are larger than in regions 4 and 5, the difference is not statistically reliable. **Table 7** shows length-controlled first-pass reading time for each sentence region.

Table 7: Length-controlled regional first-pass reading times in milliseconds (SD) for adjunct-island violations and non-island violation control sentences (experiment 3).

Sentence	Adjunct-Island Condition			
Region	Violation		Control	
1	8	(118)	-9	(106)
2	-8	(87)	-7	(85)
3	25	(97)	20	(100)
4	-20	(131)	-48	(117)
5	35	(92)	10	(67)
6	70	(264)	47	(213)
7	-42	(217)	-27	(246)

Adjunct island extraction registers in the incidence of regressive eye movements (**Table 8**) as early as the verb of the main clause, region 3 [$F1(1,27)=4.23$ $p<.05$; $F2(1,27)=6.08$, $p<.05$], although the modal point of regression increase due to adjunct island violations occurs at the main clause direct object, region 4 [$F1(1,27)=14.38$ $p<.001$; $F2(1,27)=20.30$, $p<.0001$]. The increase in regressions persists to the complementizer *while* in region 5 [$F1(1,27)=13.09$ $p<.001$; $F2(1,27)=8.65$, $p<.01$], but subsides by region 6. However, the sentence final region shows a resurgence of regressions in the

case of island violations, relative to controls [F1(1,27)=6.09 p<.05; F2(1,27)=11.70, p<.01].

Table 8: Proportion of first-pass regressive eye movements (SD) by region for adjunct-island violations and non-island violation control sentences (experiment 3).

Sentence	Adjunct-Island Condition			
Region	Violation		Control	
1	--	--	--	--
2	.15	(.36)	.14	(.35)
3	.17	(.38)	.09	(.29)
4	.24	(.43)	.12	(.32)
5	.13	(.34)	.06	(.23)
6	.39	(.49)	.37	(.48)
7	.70	(.46)	.52	(.50)

A main effect of grammaticality for *whether*-island violations is also readily apparent in reading times (**Table 9**; **Figure 5**). This first surfaces at the main verb of the matrix clause, *wonder*, in sentence region 3 [F1(1,27)=10.04 p<.005; F2(1,27)=6.14, p<.05], but the modal effect is at the subject of the embedded clause, region 5 [F1(1,27)=12.76 p<.005; F2(1,27)=9.61, p<.005]. The influence of grammaticality on reading times persists to region 6, containing the verb of the embedded clause [F1(1,27)=4.80 p<.05; F2(1,27)=7.65, p<.01]. However, differences between *whether*-island violations and their controls that surface in sentence region 6 must be interpreted with caution due the fact that, while *whether*-island

violation sentences end at region 6, control sentences continue for another two to three words.

Extraction from *whether*-islands also registers in the incidence of regressive eye-movements (**Table 10**). The strongest effect is seen in region 6, [F1(1,27)=41.76 p<.0001; F2(1,27)=109.44, p<.0001], where island violations incur more regressions than control sentences. However, as noted, such effects must be interpreted with caution due to the fact that whether-island violations end at region 6, whereas the control sentences continue for two to three words beyond it. A marginal effect surfaces at region 5 [F1(1,27)=3.93 p=.0576; F2(1,27)=6.79, p<.05]. There is also a small, unexpected, effect, in terms both of timing and direction, at the subject of the main clause, region 2 [F1(1,27)=4.23 p<.05; F2(1,27)=7.90, p<.01], where control sentences provoke more regressions than island violations.

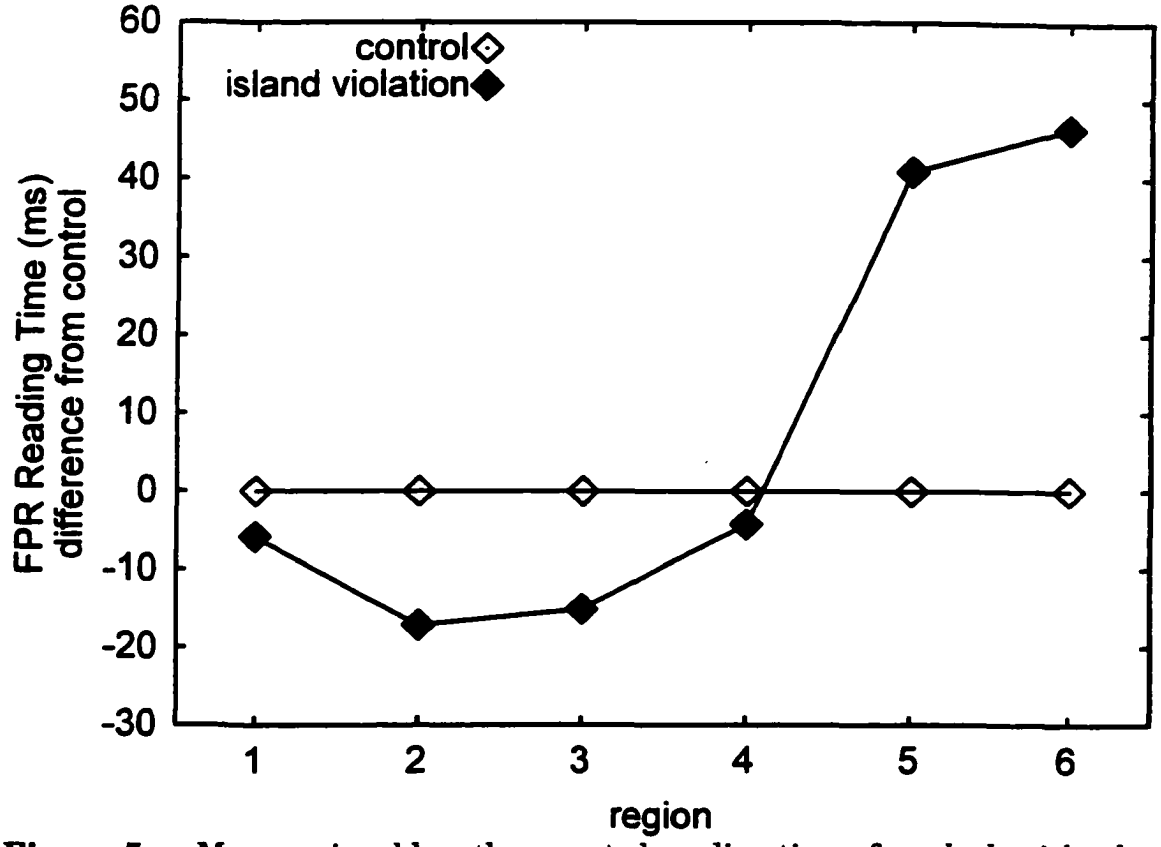

Figure 5: Mean regional length-corrected reading times for *whether*-islands violations plotted as differences from the non-island-violation control condition (experiment 3).

Table 9: Length-controlled regional first-pass reading times in milliseconds (SD) for *whether*-island violations and non-island violation control sentences (experiment 3).

Sentence	*Whether*-Island Condition			
Region	Violation		Control	
1	-11	(102)	-5	(121)
2	-41	(100)	-23	(106)
3	-24	(66)	-9	(71)
4	-22	(82)	-17	(92)
5	0	(134)	-41	(109)
6	72	(211)	26	(180)
7	--	--	-29	(189)

Table 10: Proportion of first-pass regressive eye movements (SD) by region for *whether*-island violations and non-island violation control sentences (experiment 3).

Sentence	*Whether*-Island Condition			
Region	Violation		Control	
1	--	--	--	--
2	.09	(.29)	.14	(.35)
3	.03	(.18)	.07	(.26)
4	.05	(.23)	.09	(.28)
5	.16	(.37)	.10	(.29)
6	.55	(.50)	.23	(.42)
7	--	--	.46	(.50)

Of greatest interest, however, is the clear presence of an analog to judgment satiation in the perturbations of eye-movements attributable to *whether*-island violations, relative to grammatical controls. For each sentence region, ANCOVAs (island-violation by list position) were conducted, using length-corrected reading times as the dependent measure. In the case of *whether*-island sentences, a significant interaction of island condition and list position in sentence region 5 [$F(1,562)=6.10$, $p<.05$] suggests that the effort expended in processing *whether*-island violations, relative to control sentences, increases over the course of the experimental session (**Figure 6**). Sentence region 5 is the point at which the modal reading time effect for *whether*-island violations occurs (see **Table 9** and **Figure 5**).

Conversely, perturbations due to adjunct island violations, relative to their controls, do not change over the course of a session (**Figure 7**). Crucially, separate ANCOVAs were carried out at each sentence region, using both reading times and incidence of regressive eye-movements as dependent measures. In the case of adjunct islands, no sentence region shows an interaction between list position (in test sequence) and island condition (island-violation versus control).

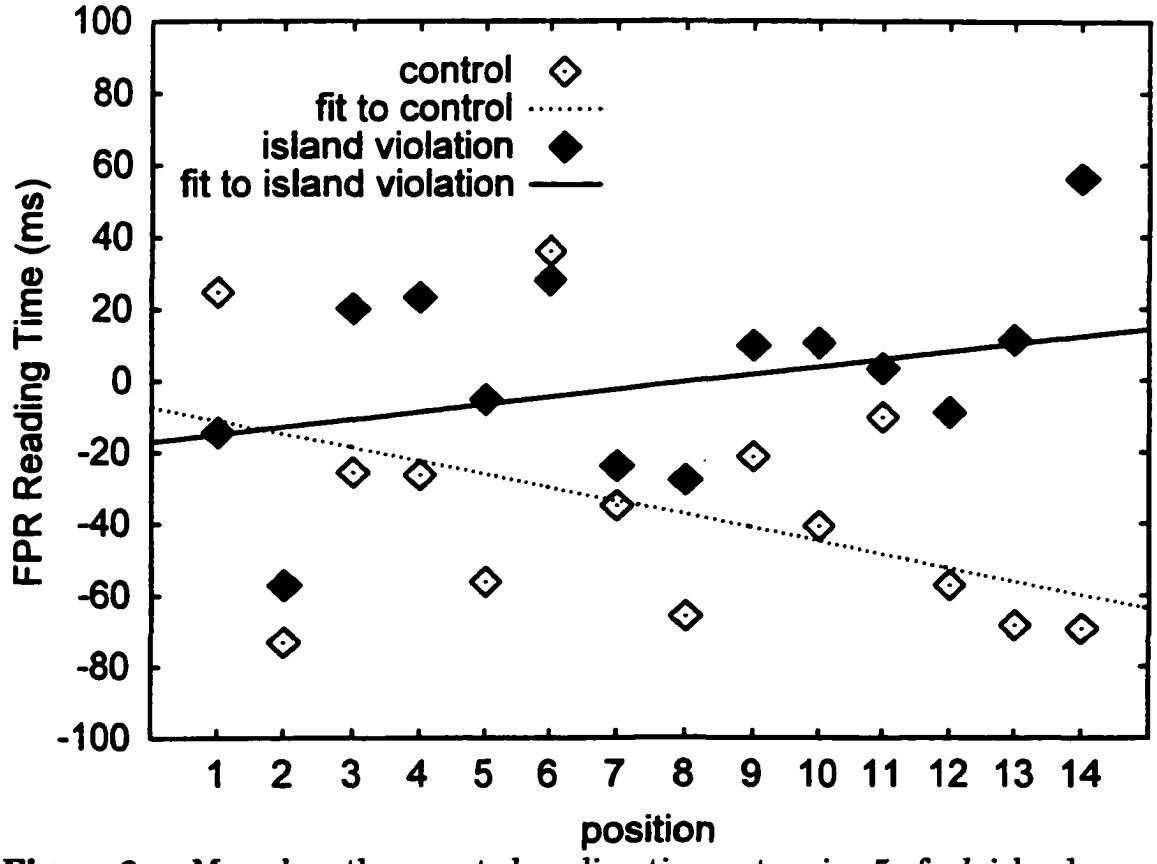

Figure 6: Mean length-corrected reading times at region 5 of *wh*-island violations and non-island-violation controls for each of 14 positions in presentation sequence, with linear fits to each sentence type (experiment 3).

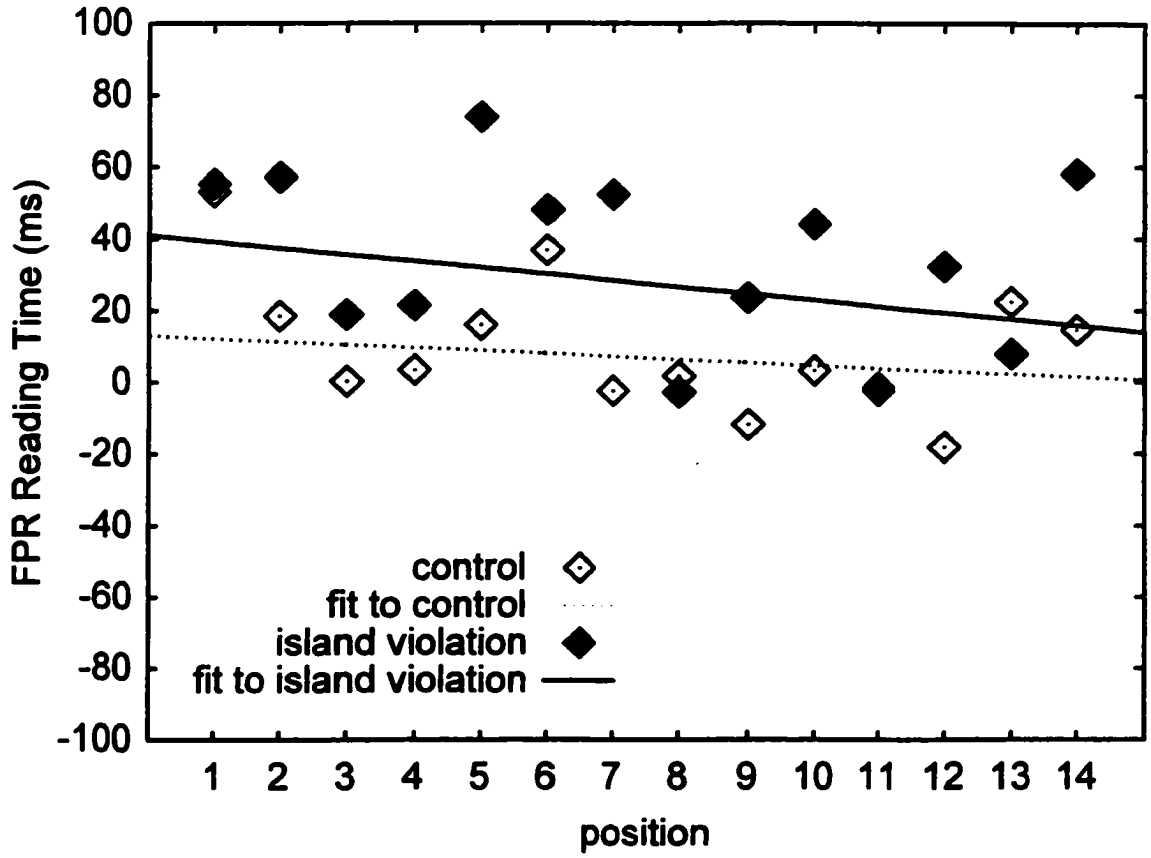

Figure 7: Mean length-corrected reading times at region 5 of adjunct island violations and non-island violation controls for each of 14 positions in presentation sequence, with linear fits to each (experiment 3).

4.7 Discussion

The results of experiment 3 are generally consistent with the predicted outcome. First, main effects of island violation are clearly visible for both adjunct islands and *whether*-islands. Further, these effects are consistent with previous work supporting the early sensitivity of the parsing mechanism to island configurations (Stowe, 1985; McKinnon & Osterhout, 1996; McElree & Griffith, 1998). While a demonstration that the parser is sensitive to island

configurations is hardly surprising, these effects serve to confirm the validity of the present experimental protocols.

Of specific interest is the evidence for an on-line counterpart of the judgment satiation effect (Hiramatsu, 2000; Snyder, 2000). **Figure 6** shows that the difficulty of *whether*-island violations, relative to matched grammatical controls, changes as readers progress through an experimental session. No such change is in evidence for adjunct island violations. At a minimum, the contrast suggests that the intuitive anomaly native English speakers report for these configurations has different etiologies. A more interesting possibility is that the intuitive anomaly due to *whether*-islands derives from a non-grammatical constraint.

Regardless of the correct theoretical account, the empirical facts are clear. This experiment demonstrates that *whether*-island violations evoke an on-line counterpart of judgment satiation in that they induce progressively longer reading times as subjects are exposed to repetitions of these structures, relative to non-island control sentences. Repeated exposure to adjunct island violations, on the other hand, induces no corresponding increase in reading times. Judgment satiation is visible in an on-line non-judgment task -- reading for comprehension.

As noted previously, Snyder (2000) suggests that *whether*-island violations, like (36b), repeated as (40) (and perhaps weak islands, generally),

are perceived as anomalous not because they violate some grammatical principle, but rather because they are difficult to process in some way yet to be made explicit.

(40) What did Valerie wonder whether Don had liked?

Snyder's suggestion follows a proposal of Kluender and Kutas (1993b). In a sentence judgment task, Kluender and Kutas asked subjects to rate the acceptability of sentences like those in (41). The general outcome of the rating experiment was that subjects deemed non-island sentences like (41a) more acceptable than *if*-island violations, (41b), which in turn were more acceptable than *wh*-island violations, (41c). Kluender and Kutas conclude that processing demands associated with crossing a clause boundary, combined with the demands associated with processing words that are in some sense semantically "heavy" (like *if* or *who*), combine to generate a point of exceptional processing load at the embedded clause boundary. It is this processing overload that leads to the intuition of unacceptability in sentences like (41b) and (41c).

(41) a. What did you figure out that you should tell the boss about *e* before the meeting?
b. What can't you figure out if you should tell the boss about *e* before the meeting?
c. What$_j$ can't you figure out who$_i$ e_i should tell the boss about e_j before the meeting?

Other models also make more or less explicit statements about the relationship between processing overload and the intuition of ungrammaticality. Gibson (1998), for example, provides an explicit theory of processing load capable of accounting for the intuition of unacceptability associated with multiply center-embedded clauses. Ferreira and Henderson (1998), in order to account for certain garden-path re-analysis phenomena, assume that the search for a grammatical analysis of a word string terminates after some set period of time (which may vary from one person to the next). Difficult sentences, like those instantiating "hard" garden-path environments, may take so long to analyze, that the processor simply gives up before reaching the correct parse.

Lacking in the story so far, is a specific mechanism to account for the apparent increase in processing effort, demonstrated in the current experiment, that attends the judgment satiation effect. I propose an explanation in the context of the Referential Model of sentence processing (Crain & Steedman, 1985; Altmann & Steedman, 1988; Ni, 1991). The proposal exploits an active memory component of the model, the "central executive," which is based on the verbal working memory conception of Baddeley (1986).

4.7.1 The Referential Model and Some Extensions

The Referential model of sentence processing (Crain & Steedman, 1985; Altmann & Steedman, 1988; Ni, 1991) was developed, in large part, to handle mounting evidence that structural ambiguities are not always resolved through structure-based parsing preferences like, for example, Minimal Attachment (Frazier, 1979). Several features of the Referential Model are noteworthy here.

First, the model adopts a modular view of language processing in which sub-components of the language faculty that operate in distinct domains (phonological, syntactic, and semantic) function in a largely independent manner. These sub-modules are organized in a linear fashion, and information flows from lower levels to higher, but not the reverse. Thus, the Referential Model can be fairly characterized as a serial feed-forward device, although it does allow for limited interaction between modules.

Second, the model allows for limited representational parallelism when ambiguous input is encountered (Crain & Steedman, 1985; Altmann & Steedman, 1988; Ni, 1991; Crain et al., 1996; Ni et al., 1996). In the case of structural ambiguity, all possible structures are constructed by the parser, in parallel. Disambiguation occurs, ideally, as the semantic, or *referential*, component evaluates syntactic structures proposed by the parser on a more or less word-by-word basis. The semantic processor attempts to locate among the candidate structures one that entails fewer extensions to the mental discourse

model than others. Whenever such a structure can be identified, it is selected by the referential component as the preferred parse, and other structures are discarded to free space for subsequent input. Crain and Steedman (1985) formulate the relevant semantic principle as (42).

(42) *principle of parsimony*
If there is a reading that carries fewer unsatisfied but consistent presuppositions or entailments than any other, then, other criteria of plausibility being equal, that reading will be adopted as the most plausible by the hearer, and the presuppositions in question will be incorporated into his or her [mental] model. (p. 333)

The Principle of Parsimony gives a means of determining what readings are least costly in the current context: those that violate the fewest presuppositions or entailments. Application of (42) will not always succeed in identifying one candidate structure as less computationally demanding than others. In these cases, the sentence processor must query the world knowledge base in order to disambiguate a parse. Refer to Crain and Steedman (1985) for further discussion.

Finally, a key feature of the Referential Model is that the flow of information from one sub-module to the next (phonology, syntax, semantics) is regulated by an executive function, which can be thought of as an attentional mechanism, a component of working memory that arbitrates among competing demands of the various sub-processors (Baddeley, 1986; Crain et al., 1990;

Baddeley, 1996). Shankweiler and Crain (1986; Shankweiler, 1989) argue that the executive function is a critical locus of breakdown in failures of sentence processing and reading comprehension. In the standard Referential Model, the primary task of the executive function is to regulate the flow of information from the phonological processor up through the parser and the referential module. The regulatory function is crucial to the smooth operation of the processor, as memory capacity at each level of processing is assumed to be quite limited, and so information must be moved through the system quickly in order to make way for subsequent input.

The Referential Model, as described, is depicted in **Figure 8**. As noted previously, the language processing system consists of phonological, syntactic and semantic/referential sub-processors. Each sub-processor operates largely independently with access to its own memory pool and computational mechanisms. The linguistic signal contacts the system at the phonological component. The phonological component functions to identify words in a linguistic signal by mapping the signal onto entries in the lexicon.[5] The latency with which words are identified depends on a number of factors, including word frequency or morphological frequency (MacDonald, Pearlmutter, & Seidenberg, 1994; Braze, 1999). Consider that, for any given verb, the *-ed* form may more frequently instantiate the past-tense than the past-participle, or vice-versa. MacDonald et al. (1994) suggest that these relative frequencies may

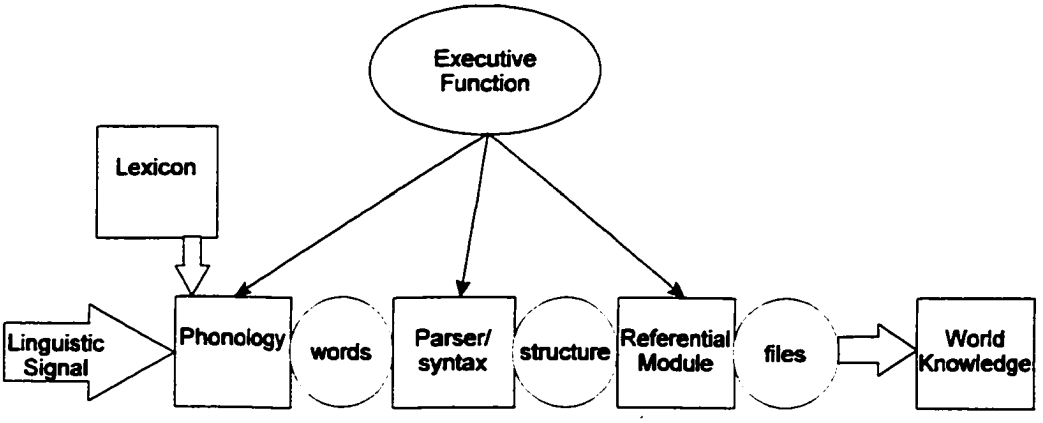

Figure 8: The Referential Model of the language processor.

modulate garden-path effects in ambiguous contexts. On their proposal the strong garden path effect in sentences like (43a) is due, in part, to the fact that *rushed* is used more often as a past-tense form than as a past participle.

Braze (1999) recorded subjects' eye-movements as they read temporarily ambiguous sentences like those in (43). Much research has shown that when reduced relative clause garden-path sentences like (43a) are encountered, readers typically interpret the initial, ambiguous verb (*rushed*) as a past-tense main verb. This analysis is proven incorrect when subsequent verbal material (*couldn't*) cannot be integrated into the existing structure. Evidence of this type of mis-analysis surfaces as elevated reading times at the disambiguating verbal material (*couldn't*) relative to appropriately matched control sentences, like (43b). Braze found that when the ambiguous verb occurs more frequently as past-participle than as a past-tense form, the garden path effect is reduced.

Specifically, shorter reading times are seen at the point of disambiguation when the ambiguous verb is biased toward past-participle use, than when it is biased toward past-tense use. These findings suggest that the MacDonald et al. speculation is on the right track.

(43) a. The felons rushed into the cellblock couldn't see the warden.
b. The felons rushed into the cellblock but couldn't see the warden.

Braze (1999) proposes modifications to the Referential Model to accommodate these findings. First, it is proposed that the phonological component makes available the multiple senses of an ambiguous word to higher level processes, and that the temporal sequence in which these senses become available is modulated by their relative frequencies. The past-tense \ past-participle ambiguity that arises in the *-ed* forms of regular verbs is a special case. As noted, these senses are accessed by the phonology in a sequence determined by their relative frequencies, with the more frequent use being accessed first. Further, the latencies with which syntactic structures are computed by the parser are modulated by the frequencies associated with lexical items being incorporated into those structures. This modification to the Referential Model is in line with the ranked parallel parsing model proposed by Gorrell (1987).[6] See Braze (1999) for details of how verbs with past-tense and past-participial biases proceed under this version of the Referential Model.

The output of the phonological component is a sequence of words, indexed according to the order in which they occur in the linguistic signal. An ambiguous token in the signal will result in the phonological output containing more than one word with the same index in the output. Simplifying somewhat, the output for the first three words of (43) will be indexed as shown in (44). As noted, the latencies with which the two versions of *rushed* are added to the set of words available to the parser may differ, depending on their relative frequencies. Traces of each word remain in the lexical buffer for only a short time. Longer term persistence of a word in the parse depends on its being taken up by higher level processes and incorporated into the product of each processing level (and, ultimately, into the output of the system as a whole).

(44) the_1 $felons_2$ $rushed$[past-tense]$_3$ $rushed$[past-participle]$_3$ . . .

The parser takes words identified by the phonology and combines them into hierarchical structure(s) consistent with their lexico-syntactic properties, and with the previously existing structure. In the case of ambiguous words, the parser will reject alternatives not consistent with existing structure. If multiple alternatives are consistent with existing structure, the parser will "clone" the structure and pursue alternative parses in parallel. Given the input in (44), for example, the parser will build parallel structures, one in which *rushed* is a past-tense main verb, (45a) and another in which it heads a

reduced relative clause, (45b). Maintenance of multiple structures is extremely demanding of the limited resources available to the syntactic component.[7] The functional limitations of the parser demand that any ambiguity be resolved quickly. Structural ambiguity can often (although not always) be resolved by appeal to the referential properties of the candidate structures.

(45) a. $[_{IP}\ [_{DP}$ the felons] $[_{VP}$ rushed] . . .]
b. $[_{IP}\ [_{DP}$ the $[_{NP}\ [_{NP}$ felons] [rushed]]] . . .]

The referential/semantic module appraises the structure(s) offered up by the parser and modifies a mental model of the discourse to maintain consistency with the updated structure. Updating must take place very rapidly in order to keep up with incoming linguistic signal. Hence, semantic processing occurs incrementally, on a more or less word-by-word basis. As a consequence, the semantic processor is, in the general case, faced with the task of evaluating syntactic constituents that are not yet complete. In the case of ambiguity, the referential module attempts to identify, among the alternatives, one that entails fewer modifications to the existing mental model. If it can do so, then that structure is selected and the others are discarded. If no structure can be identified as more parsimonious than another, then world knowledge and extra-linguistic inferencing processes are polled to determine the most plausible candidate. Accessing these extra-linguistic resources places a severe

strain on the storage capacity of the language module, as linguistic processing at lower levels must be deferred until the upstream bottleneck can be resolved.

While the nature of the output of the phonological and syntactic sub-processors is intuitively clear, perhaps more need be said about the output of the syntactic/referential sub-processor. Essentially, the semantic component builds discourse level representations. Ni (1991) adopts the Heim (1982) notion of file change semantics as a model of these representations. Under that assumption (and simplifying somewhat), discourse is represented as a file. A file contains a set of cards; each entity in the discourse is represented by a card. A discourse entity can take the form of either an individual or a set, in either case represented by a single card. Relationships that an entity bears to others in the model are specified on its card. A certain amount of background knowledge, or common ground, is assumed to exist in the file even when the linguistic context is empty (the putative "null" context). New cards are added to the file whenever a new entity is introduced into the conversational context. The introduction of new cards (and the addition of information to existing cards) is assumed to be a costly procedure, and thus is avoided by the semantic processor if possible, as dictated by the Principle of Parsimony.

In the case of the partial syntactic structures shown in (45), the reduced-relative reading requires that the set of felons already in the mental model be partitioned into those that were rushed and those that were not,

while the main-verb reading requires no such partition. Thus, the main verb reading is more parsimonious. The semantic processor will adopt the main-verb reading because doing so is computationally less demanding, at that point in the parse, than adopting the reduced-relative reading.[8]

The Referential Model assumes that operations of the human sentence processor are motivated by the need to minimize effort, or resource consumption, at each operational level. This is motivated by the idea that the processor is subject to rather extreme resource limitations, and so must use those resources efficiently in order to keep pace with the stream of linguistic input. Exceptional demands at any level of processing may create processing bottlenecks that impede the comprehension process, or even derail it completely. The role of the central executive is to allocate attentional resources in such a way as to minimize the occurrence of such bottlenecks. The next section discusses how this mechanism might be used to explain the results of experiment 3.

4.7.2 Incorporating Results of Experiment 3

Previous work has demonstrated that the syntactic, or judgment, satiation effect can be reliably elicited under experimental control (Hiramatsu, 2000; Snyder, 2000). These studies show that repeated exposure to some types of initially unacceptable sentences, including *whether*-island violations like (46a),

leads to such sentences being accepted with increasing frequency. Data from the present study indicates that judgment satiation for *whether*-island violations is attended by an increase in processing effort, relative to non-island violation control sentences like (46b), as indexed by reading times at the subject of the embedded clause. See **Figure 6**.

(46) a. What did Valerie wonder whether Don had liked?
b. When did Valerie wonder whether Don had liked the story?

This finding is consistent with the idea that the relative unacceptability of *whether*-island violations is due to a limitation of the processor, rather than a grammatical constraint. Of course, the challenge for any explanation of judgment satiation and the corresponding eye-movement effect, is to provide an account that fits well with existing presupposition and theory. A grammatical account of these phenomena would seem to require that a relatively limited exposure (five tokens in Snyder's study, seven in Hiramatsu's) to putatively ungrammatical strings of certain types (*whether*-island violations like (46a), among others) have the effect of altering an individual's grammar. While an explanation of satiation that appeals to the malleability of grammar is not outside the realm of possibility, this section will advance a processing-based approach to these phenomena.

Kluender and Kutas (1993b) speculate that the unacceptability of *wh*-extraction from indirect questions is due to the juxtaposition of a semantically

difficult word (the complementizer of an embedded question), and a clause boundary (taken to be a locus of high processing resource demand). Other research suggests that the search process involved in associating a *wh*-word with an appropriate gap places a significant load on the sentence processing system (Fodor, 1978; Gibson, 1991; Traxler & Pickering, 1996). Thus, Kluender and Kutas attribute the difficulty of (46a) to the requirement that the processor execute three computationally demanding operations simultaneously, or nearly so.

Indeed, prior research suggests that clause boundaries are loci of increased processing effort. In the context of eye-movement based research, Rayner, Sereno, Morris, Schmauder and Clifton (1989) demonstrate that a word occurring at the end of a clause attracts a longer gaze duration than when the same word does not end a clause.[9] This is often termed the clause wrap-up effect. Although the genesis of the wrap-up effect is not well understood, many researchers attribute it to the sentence processing mechanism's effort to close off unfinished work associated with the current clause, before moving on (Just & Carpenter, 1980; Balogh, Zurif, Prather, Swinney, & Finkel, 1998; Rayner, Kambe, & Duffy, 2000).[10]

Additional evidence that the clause boundary has some special status with respect to processing effort comes from the "click migration" studies of

the late sixties and early seventies (e.g., Fodor & Bever, 1965; Garrett, Bever, & Fodor, 1966; Holmes & Forster, 1972).

Processing effects of a purely semantic nature have, however, been more difficult to establish.[11] (Recall that Kluender and Kutas (1993) focus on the difference between *that* and *if*.) Cutler (1983) suggests that there are at least two ways in which a semantically complex word might encumber the sentence processing system. First, it may be that lexical access for semantically "heavy" words is somehow delayed, relative to less complex words, although it is not clear just why the semantic properties of a word should affect the process of lexical access in this way. Nonetheless, if such effects were to exist, we would reasonably expect them to show up early in reading time data, probably at the critical word itself. A more plausible possibility is that semantically rich words increase processing load because they entail more complex operations downstream, at the semantic/referential level of processing.

Evidence for semantic effects in on-line tasks is somewhat mixed. Rayner and Duffy (1986), for example, monitored eye-movement patterns while subjects read sentences containing various types of putatively semantically complex verbs (or appropriate controls). Three classes of verbs were tested: decomposable causatives, factives, and negative verbs. For each class, semantically complex verbs were paired with less complex controls in identical sentence frames, as (47) shows for the factive case. Their study failed

to uncover evidence of semantic complexity effects for any of the three classes of verbs.

(47) a. The cook regretted that he had been lying.
b. The cook testified that he had been lying.

For present purposes, we are concerned with the contribution of *whether* to processing load. Like *if* in the Kluender and Kutas study, *whether* introduces an indirect yes/no question, or alternative question. The denotation of an alternative question is typically taken to be a set containing the propositions expressing contextually salient possible answers (or, alternatively, the salient true answers) (Karttunen, 1977). Given a sentence like (48) (compare to (46a)), we can take the reference of the embedded question to be a set containing the two propositions shown in (49). Thus, the object of *Valerie*'s wondering is the veracity of the propositions in that set. I assume that the denotation of an embedded yes/no question is a complex referential entity that must be represented in the discourse model. Specifically, the word *whether* triggers the introduction of a card into the discourse model to represent the embedded question. The precise content of the card is fleshed out by subsequent input.

(48) Valerie wondered whether Don had liked the story.

(49) { "Don had liked the story" "Don had not liked the story" }

It is certain that each source of computational loading (filler-gap association, semantic complexity, or clause recursion) in isolation is within the capacity of the sentence processing mechanism. Kluender and Kutas speculate that, in the context of *wh*-extraction from an indirect question, the combined effect of these loading factors creates a situation in which the sentence processing mechanism is overburdened. It is this overload that leads to the unacceptability of *wh*-extraction from *whether*-islands.

The idea that processing overload can lead to judgments of unacceptability ranging from mild to severe is not new. The well-known case of multiply center embedded clauses is discussed in chapter 1. Gibson (1991; 1998), goes into some detail as to how diverse sources of processing load may contribute to unacceptability in unambiguous contexts. In the context of garden-path sentences, Ferreira and Henderson (1991; 1998) argue that there is a limit to the amount of time that a reader/listener will expend on recovering from a garden-path. If that limit is exceeded before the sentence is fully processed, then the sentence is likely to be perceived as unacceptable (with the unacceptability interpreted as ungrammaticality).

It seems uncontroversial to say that linguistic complexity of various types may overload the sentence processing mechanism, and so lead to sentences being perceived as unacceptable. I adopt this premise. In the context of the Referential Model discussed in the previous section, I propose that there

are limits to the amount of effort each sub-module within the sentence processing mechanism can expend at any particular point in a sentence. If that limit is exceeded then processing breaks down, either partially or wholly. A processing failure of this sort leads to a sentence being perceived as unacceptable. Thus, the apparent ungrammaticality of *whether*-island violations, (46a), is explained without appealing to any grammatical constraint.

How, then, do satiation phenomena fit into this picture? Recall that the executive function of the Referential Model is responsible for arbitrating among competing demands of the various sub-processors in order to facilitate the flow of information through the sentence processing system. The central executive is, by hypothesis, a dynamic system. It is capable of adjusting the allocation of attentional resources to meet the moment-by-moment needs of each sub-processor. When the sentence processing system is confronted with a construction that places heavy demands on more than one sub-processor at the same time (like *wh*-extraction from *whether*-islands), the executive function attempts to balance the allocation of attentional resources in such a way as to maximize the effectiveness of the entire system. I assume that with repeated exposure to identical constructions, this "balancing-act" becomes more refined. Thus, as the executive function sees more *whether*-island violations, it learns to cope with them more efficiently, within the limits of its capacity. This is why the acceptability of *whether*-island violations increases with repeated exposure.

The next question is that of how the increase in reading times for *whether*-island violations might follow from this proposal. I assume that each sub-processor is allocated some default quantity of the executive function's capacity, and that the distribution across all sub-processors is sub-optimal for the purpose of comprehending *whether*-island violations. Degraded acceptability in the judgment task occurs when one sub-processor (by hypothesis, the referential sub-processor, in the case of *whether*-island violations) is unable to fully carry out its function due to a resource limitation. In other words, the sub-processor runs out of time. As noted earlier in this chapter, Ferreira and Henderson (1998) argue that certain operations of the human sentence processing mechanism are temporally bounded. The increasing reading times subjects exhibit in the course of comprehending *whether*-island violations arise when the executive function shifts resources to the initially deprived referential sub-processor. This shift in resources allows the referential component to (more) fully carry out its function, taking more time in the process. Thus, the increases in acceptability and in reading time are accounted for by the same mechanism.

This hypothesis leads to an interesting set of predictions. First, other constructions whose difficulty is conventionally attributed to working memory limitations may also show judgment satiation effects. One such case is that of center-embedded clauses like (2a) and (2b) repeated in (50). Sentences like

(50a), and even the singly embedded (50b), are well-known for their processing difficulty and corresponding lack of acceptability (see chapter 1 for discussion). The present hypothesis about the role of working memory in sentence comprehension suggests that sentences like these may become increasingly acceptable if they are encountered frequently enough for the executive function to acclimate to them.

(50) a. The rat that the cat that the dog chased killed had eaten the cheese.
b. The rat that the cat killed had eaten the cheese.
c. The dog chased the cat that killed the rat that had eaten the cheese.

A second prediction is that violations of other types of "weak" islands may show both judgment satiation effects and processing satiation effects. One plausible candidate is the factive island. Ross (1967) observed that *wh*-extraction from the complement of a factive verb leads to a modestly unacceptable sentence as shown in (51a).[12] If this prediction turns out to be accurate, then we will have found oblique evidence for an influence of factivity on processing. This will be interesting in light of Rayner and Duffy's (1986) failure to detect an effect of factivity in eye-movements.

(51) a. ?? What did Bill confirm that Roger had eaten?
b. What did Bill allege that Roger had eaten?

Another interesting test case is extraction from the scope of negation or other affective operator (Ross, 1984; Rizzi, 1990), as shown below. Here too we might expect to see satiation effects in both judgments and reading times. Note that (52a) is of the same class as the negative verbs examined in Rayner and Duffy (1986), although it incorporates *wh*-movement as well. As in the case of factives, compounding multiple sources of processing load may give us a means to detect effects that are otherwise below the threshold of sensitivity of standard methods.

(52) a. * How well did John deny that he performed?
b. * How well did few critics think that he performed?

This chapter develops some novel data with respect to the processing of *whether*-islands, and proposes an account of these data in terms of working memory and processing load. It is suggested that a range of additional effects may be amenable to analysis under the proposed model. Tentative predictions are made as to how processing and judgment satiation data could confirm or disconfirm that possibility.

Notes

1. As should be clear from discussion in the previous chapter, processing difficulty for a given sentence type can be deduced from elevated reading times, increased incidence of regressive eye movements, and other eye-movement indicators, relative to an appropriate baseline condition.

2. Evidence of the parser's rapid sensitivity to syntactic islands can be found in studies exploiting a variety of behavioral and neural measures (Crain & Fodor, 1985; Stowe, 1985; Kluender & Kutas, 1993b, 1993a; McKinnon & Osterhout, 1996; McElree & Griffith, 1998; Phillips & Wong, 2000).

3. In order to balance the fact that all control items are *when* questions, whereas all island violations are *what* questions, one-half of the fifty-six filler items are grammatical *what* questions, while the other half are ungrammatical *when* questions.

4. Author recognition scores for participants in experiment 3 are considerably higher than those for participants in experiments 1 and 2. This is likely due to the fact that the participants in the first two experiments were almost entirely underclassmen (freshmen and sophomores), while experiment 3 included students with a more variable number of years of post-high-school education, including underclassmen, juniors, seniors, and graduate students.

5. The linguistic signal may be acoustic, orthographic, or even manual. In order to accommodate each type of input, phonological processing must operate on highly abstract representations.

6. There are substantial differences between the Referential Model proposed here, and Gorrell's ranked parallel model. In Gorrell's system, timing differences in the computation of parallel syntactic structures are due solely to the relative complexity of those structures, while Braze's proposal allows for the timing of lexical access to influence the time course of structure building operations. Further, in Gorrell's system, timing differences are ultimately converted to ranks, with the least complex (and, therefore, the most rapidly computed) structure being ranked highest. The highest ranked structure is most accessible to higher level processes. While Braze also allows for timing to influence the accessibility of competing structures, there is no explicit ranking procedure. Further, when multiple syntactic structures are available to higher level processes simultaneously (or nearly so), it is computational

complexity at that higher (semantic/referential) level that determines which structure is adopted.

7. If the incoming word is incompatible with existing structure, a process of syntactic re-analysis is initiated. The precise nature of this process is beyond the scope of this thesis, although a number of proposals have recently emerged in the literature. See, for example, Fodor and Inoue (1994; 2000), or the various papers in Fodor and Ferreira (1998).

8. The model's structural preference can be switched by manipulating the referential/semantic properties of the sentence, as by substituting the focus particle *only* for the determiner, *the*, in the subject NP. See Ni, Crain, and Shankweiler (1996) for details.

9. The clauses in question were preposed adjunct clauses (variously temporal or purposive).

10. The fact that the wrap-up effect occurs at a syntactically defined juncture does not entail that the associated processing is syntactic in nature. Rayner, Kambe and Duffy (2000), for example, show that the process of anaphor resolution impacts on the strength of the wrap-up effect.

11. Semantic effects discussed here stand in contrast to pragmatic effects of the sort discussed in chapter 3 (also see chapter 1, note1).

12. The factive island case is borrowed from Ross (1967, example 6.191), judgments intact.

Chapter 5: Future Directions

This thesis reports on the results of three experiments designed to further the goal of using controlled behavioral methods to evaluate grammatical and non-grammatical influences on language comprehension processes. The outcome of these experiments is argued to be consistent with the Referential Model of the human sentence processing mechanism (HSPM). Experiments 1 and 2, reported in chapter 3, serve two related purposes. The first is to demonstrate that patterns of eye movements during reading are characteristically distinct in response to specific types of grammatical and non-grammatical anomaly in the input. The second purpose is to clarify how the HSPM's response to (a specific class of) grammatical anomaly differs from it's response to (a specific class of) non-grammatical anomaly, as indexed by eye movement behavior. Experiment 3, described in chapter 4, uncovers an on-line counter-part to the judgment satiation effect. The results tell of a somewhat surprising correspondence between perceived acceptability and on-line reading times that may have far-reaching theoretical implications.

Experiment 1 provides evidence supporting the dissociability of sentence processing effects due to subject-verb pragmatic anomaly in the first case, or to defects of verbal morpho-syntax in the second. The outcome of this experiment

successfully replicates previous work (Ni et al., 1998; Braze et al., in press) and also yields strong new evidence of dissociability based on the eye movement signature elicited in response to sentences containing both pragmatic and morpho-syntactic anomalies.

Experiment 2 moves to extend this result into the domain of nominal morph-syntax, and adjective-noun pragmatic relationships. This move is partially successful, in that eye movements in response to anomalies of nominal morpho-syntax closely parallel those elicited in response to anomalies of verbal morpho-syntax. One issue that will bear revisiting in a future study, however, is the lack of parallelism between the noun-based pragmatic anomalies of experiment 2 and the verb-based pragmatic anomalies of experiment 1. As noted in chapter 3, the two words involved in verb-based pragmatic anomalies (subject-noun and main verb) are separated by two words (auxiliary verb and adverb), while the two words involved noun-based pragmatic anomalies (adjective and noun) are adjacent. Because of this potential confound, it is not clear whether the partially discordant results of experiments 1 and 2 indicate a real difference in the way that subject/verb and adjective/noun pragmatic relationships are evaluated. Certainly, it has been noted in the literature that the distance between the trigger and locus of processing difficulty may play a role in determining the associated degree of anomaly that is registered by the sentence processing mechanism (e.g. Frazier

& Rayner, 1982; Ferreira & Henderson, 1991). The question of whether trigger/locus distance, or some other factor, is primarily responsible for the lack of correspondence between experiments 1 and 2 requires further study.

The question of replicability may, additionally, pose a challenge for the novel results of experiment 3. There, reading times were seen to increase with repeated exposure to *whether*-island violations, a structure known to induce satiation in a judgment task. Conversely, no such increase appeared in the case of adjunct-island violations, a structure that does not induce judgment satiation. The outcome of experiment 3 leads to a rather counter-intuitive surmise, that judgment satiation corresponds to an increase in processing load for at least some sub-components of the sentence processing mechanism. The theoretical mechanisms necessary to account for this conundrum are argued to already be in place in the Referential Model of sentence processing (once suitably sharpened). However, the hypothesis can only maintain relevance if its empirical basis holds up to scrutiny. A replication of experiment 3 will add considerably to our confidence in the nature of on-line correlates of judgment satiation effects. Additionally, the proposed account of experiment 3's results predicts that a range of constructions, from other types of syntactic islands (e.g., inner islands, factive islands) to structurally complex constructions (e.g., center-embedded clauses, object-gap relative clauses)

should be subject to satiation effects in acceptability judgments and in on-line measures.

To the extent that the results of experiment 3 can be sustained and extended, they plausibly support a non-grammatical explanation of *whether*-island effects (and other weak island effects) over more traditional grammatical accounts. Maintaining such an account will necessitate reinterpretation of many generally accepted findings.

Clearly, there is a lot of work to be done. The phenomena explored in this thesis will need to be studied from a variety of perspectives. On the methodological front, the eye-movement technique used here is argued to be an efficient means of collecting behavioral data unlikely to be contaminated by meta-linguistic processes. Yet, the facts exposed here will make interesting fodder for other behavioral methods, or even ERP or brain-imaging techniques. Additionally, further work can be done to sharpen theoretical perspective adopted here, and to test the ability of other perspectives to handle the data developed in this study.

Appendix A: Experimental Materials

Experiment 1: verb-based anomalies

1
The daisies were slowly wilting in the hot weather this afternoon.
The puddles were slowly wilting in the hot weather this afternoon.
The daisies had slowly wilting in the hot weather this afternoon.
The puddles had slowly wilting in the hot weather this afternoon.
Is it hot this afternoon? Y

2
The crops have even sprouted early this spring in the mild weather.
The lambs have even sprouted early this spring in the mild weather.
The crops may even sprouted early this spring in the mild weather.
The lambs may even sprouted early this spring in the mild weather.
Is the weather fierce? N

3
The otters will often swim ten miles a day just for the fun of it.
The ducks will often swim ten miles a day just for the fun of it.
The otters had often swim ten miles a day just for the fun of it.
The ducks had often swim ten miles a day just for the fun of it.
Do they swim for fun? Y

4
Sparrows are always flitting around in search of something good to eat.
Minnows are always flitting around in search of something good to eat.
Sparrows must always flitting around in search of something good to eat.
Minnows must always flitting around in search of something good to eat.
Do they lumber around? N

5
The lawyer had swiftly attacked the company for its new hiring policy.
The priest had swiftly attacked the company for its new hiring policy.
The lawyer will swiftly attacked the company for its new hiring policy.
The priest will swiftly attacked the company for its new hiring policy.
Has the company changed its policy? Y

6
The author must quickly type the last few pages before the due date.
The child must quickly type the last few pages before the due date.
The author had quickly type the last few pages before the due date.
The child had quickly type the last few pages before the due date.
Must the typing be done quickly? Y

7
The pliers have certainly rusted from being left out in the rain.
The towels have certainly rusted from being left out in the rain.
The pliers will certainly rusted from being left out in the rain.
The towels will certainly rusted from being left out in the rain.
Should we leave things out in the rain? N

8
The disaster will greatly depress the entire crew when news gets out.
The festival will greatly depress the entire crew when news gets out.
The disaster has greatly depress the entire crew when news gets out.
The festival has greatly depress the entire crew when news gets out.
Has the news gotten out yet? N

9
Cougars are often prowling all through the day in search of food.
Donkeys are often prowling all through the day in search of food.
Cougars will often prowling all through the day in search of food.
Donkeys will often prowling all through the day in search of food.
Are the animals hungry? Y

10
Wood will readily combust when heated to a high enough temperature.
Iron will readily combust when heated to a high enough temperature.
Wood is readily combust when heated to a high enough temperature.
Iron is readily combust when heated to a high enough temperature.
Can high temperatures induce combustion? Y

11
The river was certainly flooded again from the thunder storms this week.
The forest was certainly flooded again from the thunder storms this week.
The river did certainly flooded again from the thunder storms this week.
The forest did certainly flooded again from the thunder storms this week.
Was the weather peaceful this week? N

12
Some restaurants have happily served for longer hours during the holidays.
Some retailers have happily served for longer hours during the holidays.
Some restaurants will happily served for longer hours during the holidays.
Some retailers will happily served for longer hours during the holidays.
Are holiday hours longer? Y

13
Youngsters will sometimes cry on their first day away from home.
Soldiers will sometimes cry on their first day away from home.
Youngsters have sometimes cry on their first day away from home.
Soldiers have sometimes cry on their first day away from home.
Is leaving home always pleasant? N

14
A razor can easily gash you if not handled with proper care.
A hammer can easily gash you if not handled with proper care.
A razor is easily gash you if not handled with proper care.
A hammer is easily gash you if not handled with proper care.
Should tools be handled carefully? Y

15
The carpenter had fully sanded the cabinet before leaving this morning.
The plumber had fully sanded the cabinet before leaving this morning.
The carpenter should fully sanded the cabinet before leaving this morning.
The plumber should fully sanded the cabinet before leaving this morning.
Did someone leave this morning? Y

16
The butter should quickly ooze into the fresh baked oat bran muffins.
The knife should quickly ooze into the fresh baked oat bran muffins.
The butter had quickly ooze into the fresh baked oat bran muffins.
The knife had quickly ooze into the fresh baked oat bran muffins.
Are the muffins stale? N

17
The paint will surely flake after a few years in this harsh climate.
The color will surely flake after a few years in this harsh climate.
The paint has surely flake after a few years in this harsh climate.
The color has surely flake after a few years in this harsh climate.
Is the climate mild? N

18
Those beams will probably crack under the excessive load placed on them.
Those tires will probably crack under the excessive load placed on them.
Those beams were probably crack under the excessive load placed on them.
Those tires were probably crack under the excessive load placed on them.
Is the load excessive? Y

19
The rabbits can easily scamper from the poorly built enclosure at the zoo.
The snakes can easily scamper from the poorly built enclosure at the zoo.
The rabbits are easily scamper from the poorly built enclosure at the zoo.
The snakes are easily scamper from the poorly built enclosure at the zoo.
Is the enclosure badly made? Y

20
The hawks were gracefully gliding through the warm days of early spring.
The deer were gracefully gliding through the warm days of early spring.
The hawks will gracefully gliding through the warm days of early spring.
The deer will gracefully gliding through the warm days of early spring.
Are the days cold? N

21
The dog was angrily yammering at the squirrel in the old oak tree.
The cat was angrily yammering at the squirrel in the old oak tree.
The dog had angrily yammering at the squirrel in the old oak tree.
The cat had angrily yammering at the squirrel in the old oak tree.
Is the squirrel in a maple tree? N

22
The jeep will probably turn past the hill and continue down the valley.
The road will probably turn past the hill and continue down the valley.
The jeep is probably turn past the hill and continue down the valley.
The road is probably turn past the hill and continue down the valley.
Is there a mountain beyond the hill? N

23
The lamps had festively glittered for the entire week of the festival.
The bells had festively glittered for the entire week of the festival.
The lamps will festively glittered for the entire week of the festival.
The bells will festively glittered for the entire week of the festival.
Does the festival last for two weeks? N

24
The wine will surely stain the carpet after soaking in all night.
The water will surely stain the carpet after soaking in all night.
The wine has surely stain the carpet after soaking in all night.
The water has surely stain the carpet after soaking in all night.
Was the mess cleaned up immediately? N

25
The oranges are quickly ripening in the mild South African climate.
The carrots are quickly ripening in the mild South African climate.
The oranges had quickly ripening in the mild South African climate.
The carrots had quickly ripening in the mild South African climate.
Is the South African climate mild? Y

26
Wolves may possibly hunt in this forest again after a long absence.
Moose may possibly hunt in this forest again after a long absence.
Wolves are possibly hunt in this forest again after a long absence.
Moose are possibly hunt in this forest again after a long absence.
Have some animals been missing? Y

27
The bowl will surely shatter from being left on the hot stove.
The pan will surely shatter from being left on the hot stove.
The bowl had surely shatter from being left on the hot stove.
The pan had surely shatter from being left on the hot stove.
Was the stove off? N

28
The farmers will easily raise a good crop due to the plentiful rain.
The fields will easily raise a good crop due to the plentiful rain.
The farmers have easily raise a good crop due to the plentiful rain.
The fields have easily raise a good crop due to the plentiful rain.
Is the rain plentiful? Y

Experiment 2: noun-based anomalies

Numbers indicate lexical frequencies of alternating adjectives (Francis & Kucera, 1982).

1 f=203/143
Jim saw two black puppies chasing a cat in the yard this morning.
Jim saw two blue puppies chasing a cat in the yard this morning.
Jim saw one black puppies chasing a cat in the yard this morning.
Jim saw one black puppy chasing a cat in the yard this morning.
Was the cat being chased? Y

2 f=2/2
Gary placed three aromatic lilies in the vase on the dinner table.
Gary placed three arrogant lilies in the vase on the dinner table.
Gary placed one aromatic lilies in the vase on the dinner table.
Gary placed one aromatic lily in the vase on the dinner table.
Was the vase on the table? Y

3 f=5/9
Trevor views many sparkling wines as good to serve with hearty foods.
Trevor views many spacious wines as good to serve with hearty foods.
Trevor views every sparkling wines as good to serve with hearty foods.
Trevor views every sparkling wine as good to serve with hearty foods.
Is the wine of poor quality? N

4 f=11/11
The TA put some trivial equations in the test's extra-credit section.
The TA put some tropical equations in the test's extra-credit section.
The TA put one trivial equations in the test's extra-credit section.
The TA put one trivial equation in the test's extra-credit section.
Is extra-credit available? Y

5 f=1/1
She studied the fluffy clouds drifting slowly across the clear blue sky.
She studied the floppy clouds drifting slowly across the clear blue sky.
She studied each fluffy clouds drifting slowly across the clear blue sky.
She studied each fluffy cloud drifting slowly across the clear blue sky.
Is the sky cloudless? N

6 f=5/1
I believe most greedy gamblers will quickly lose all of their money.
I believe most gritty gamblers will quickly lose all of their money.
I believe every greedy gamblers will quickly lose all of their money.
I believe every greedy gambler will quickly lose all of their money.
Do gamblers usually win? N

7 f=33/29
Jill wants all violent felons to get the harshest possible penalty.
Jill wants all vigorous felons to get the harshest possible penalty.
Jill wants each violent felons to get the harshest possible penalty.
Jill wants each violent felon to get the harshest possible penalty.
Should the penalties be stiff? Y

8 f=31/33
He says any mature cats caught chasing mice must be promptly neutered.
He says any massive cats caught chasing mice must be promptly neutered.
He says every mature cats caught chasing mice must be promptly neutered.
He says every mature cat caught chasing mice must be promptly neutered.
Is mouse chasing tolerated? N

9 f=0/0
She decided two flawed diamonds would be better than none at all.
She decided two flabby diamonds would be better than none at all.
She decided a flawed diamonds would be better than none at all.
She decided a flawed diamond would be better than none at all.
Does she like diamonds? Y

10 f=0/0
Al likes several robust beers from the new micro-brewery in Hartford.
Al likes several robotic beers from the new micro-brewery in Hartford.
Al likes every robust beers from the new micro-brewery in Hartford.
Al likes every robust beer from the new micro-brewery in Hartford.
Is the brewery in New Haven? N

11 F=23/44
Tom noticed many monthly reports had been filed under the wrong date.
Tom noticed many mobile reports had been filed under the wrong date.
Tom noticed every monthly reports had been filed under the wrong date.
Tom noticed every monthly report had been filed under the wrong date.
Were the reports properly filed? N

12 f=12/29
Toni found two antique bottles on a shelf in the back of the shed.
Toni found two anxious bottles on a shelf in the back of the shed.
Toni found an antique bottles on a shelf in the back of the shed.
Toni found an antique bottle on a shelf in the back of the shed.
Was Toni in the shed? Y

13 f=2/8
Fred threw several rotten potatoes into the trash and cooked the rest.
Fred threw several rusty potatoes into the trash and cooked the rest.
Fred threw every rotten potatoes into the trash and cooked the rest.
Fred threw every rotten potato into the trash and cooked the rest.
Did he cook all of the potatoes? N

14 f=22/45
Debbie caught two unfortunate rats and five mice in her house last week.
Debbie caught two universal rats and five mice in her house last week.
Debbie caught one unfortunate rats and five mice in her house last week.
Debbie caught one unfortunate rat and five mice in her house last week.
Was the house infested? Y

15 f=82/66
Sue has some interesting sculptures on display in her new home.
Sue has some intellectual sculptures on display in her new home.
Sue has an interesting sculptures on display in her new home.
Sue has an interesting sculpture on display in her new home.
Does she have a new home? Y

16 f=15/11
The envoy spent two tense hours on last minute details of the treaty.
The envoy spent two tender hours on last minute details of the treaty.
The envoy spent a tense hours on last minute details of the treaty.
The envoy spent a tense hour on last minute details of the treaty.
Did he work on the treaty? Y

17 f=2/1
The article rebuts some untrue claims about the senator's past behavior.
The article rebuts some untidy claims about the senator's past behavior.
The article rebuts each untrue claims about the senator's past behavior.
The article rebuts each untrue claim about the senator's past behavior.
Does the article prove the claims? N

18 f=30/24
A lawyer argues all crucial details of a case with great care.
A lawyer argues all criminal details of a case with great care.
A lawyer argues every crucial details of a case with great care.
A lawyer argues every crucial detail of a case with great care.
Can details be ignored? N

19 f=4/4
Ralph ate two juicy melons and a bushel of apples to win the contest.
Ralph ate two jerky melons and a bushel of apples to win the contest.
Ralph ate one juicy melons and a bushel of apples to win the contest.
Ralph ate one juicy melon and a bushel of apples to win the contest.
Did John win? Y

20 f=3/2
Mary stole many flashy trinkets from the store that she broke into.
Mary stole many flaky trinkets from the store that she broke into.
Mary stole every flashy trinkets from the store that she broke into.
Mary stole every flashy trinket from the store that she broke into.
Is Mary a thief? Y

21 f=5/3
Lee took two striped kittens to school for show-and-tell today.
Lee took two stringy kittens to school for show-and-tell today.
Lee took a striped kittens to school for show-and-tell today.
Lee took a striped kitten to school for show-and-tell today.
Is show-and-tell tomorrow? N

22 f=1/3
The workers carried several rickety ladders from the barn to the trailer.
The workers carried several rustic ladders from the barn to the trailer.
The workers carried every rickety ladders from the barn to the trailer.
The workers carried every rickety ladder from the barn to the trailer.
Are the ladders on the trailer? Y

23 f=1/2
Sandra mixed the leafy vegetables with the tomatoes for a dinner salad.
Sandra mixed the leaky vegetables with the tomatoes for a dinner salad.
Sandra mixed a leafy vegetables with the tomatoes for a dinner salad.
Sandra mixed a leafy vegetable with the tomatoes for a dinner salad.
Are there peppers in the salad? N

24 f=23/34
Bill discussed many relevant topics for the essay with his class.
Bill discussed many regional topics for the essay with his class.
Bill discussed every relevant topics for the essay with his class.
Bill discussed every relevant topic for the essay with his class.
Was there a choice of topics? Y

25 f=5/5
Jack wrote two mediocre essays instead of the three that were assigned.
Jack wrote two melodic essays instead of the three that were assigned.
Jack wrote one mediocre essays instead of the three that were assigned.
Jack wrote one mediocre essay instead of the three that were assigned.
Did he complete the assignment? N

26 f=46/63
Studies find many native birds in this ecosystem to be slowly dying.
Studies find many narrow birds in this ecosystem to be slowly dying.
Studies find every native birds in this ecosystem to be slowly dying.
Studies find every native bird in this ecosystem to be slowly dying.
Is the ecosystem healthy? N

27 f=14/24
Jody picked several ripe oranges for the fruit salad she was making.
Jody picked several rigid oranges for the fruit salad she was making.
Jody picked every ripe oranges for the fruit salad she was making.
Jody picked every ripe orange for the fruit salad she was making.
Did she use oranges? Y

28 f=23/27
They saw few giant ferns growing in the remnants of the rain forest.
They saw few gentle ferns growing in the remnants of the rain forest.
They saw one giant ferns growing in the remnants of the rain forest.
They saw one giant fern growing in the remnants of the rain forest.
Are the ferns common? N

Experiment 3: on-line satiation

***Whether*-island materials**

1.
Earlier today, Sandra wondered whether Howard had reviewed the article.
When did Sandra wonder whether Howard had reviewed the article?
What did Sandra wonder whether Howard had reviewed?

2.
After the meeting, Valerie wondered whether Don had liked the story.
When did Valerie wonder whether Don had liked the story?
What did Valerie wonder whether Don had liked?

3.
Last night, the detective wondered whether the robber had hidden the loot.
When did the detective wonder whether the robber had hidden the loot?
What did the detective wonder whether the robber had hidden?

4.
This morning, the trainer wondered whether boxer had eaten a good meal.
When did the trainer wonder whether the boxer had eaten a good meal?
What did the trainer wonder whether the boxer had eaten?

5.
After breakfast, Larry wondered whether Sally had read the morning paper.
When did Larry wonder whether Sally had read the morning paper?
What did Larry wonder whether Sally had read?

6.
Two days ago, Douglas wondered whether Susan had written another poem.
When did Douglas wonder whether Susan had written another poem?
What did Douglas wonder whether Susan had written?

7.
Last week, Morris wondered whether Janet had bought a new coat.
When did Morris wonder whether Janet had bought a new coat?
What did Morris wonder whether Janet had bought?

8.
Earlier today, the boss wondered whether Thomas had finished the report.
When did the boss wonder whether Thomas had finished the report?
What did the boss wonder whether Thomas had finished?

9.
Yesterday, Chris wondered whether the embassy had received the document.
When did Chris wonder whether the embassy had received the document?
What did Chris wonder whether the embassy had received?

10.
At lunch, Laura wondered whether Kevin would finish the crossword puzzle.
When did Laura wonder whether Kevin would finish the crossword puzzle?
What did Laura wonder whether Kevin would finish?

11.
This morning, Billy wondered whether Kathy would believe the nasty rumor.

When did Billy wonder whether Kathy would believe the nasty rumor?
What did Billy wonder whether Kathy would believe?

12.
Over lunch, the doctor wondered whether the NIH would approve his proposal.
When did the doctor wonder whether the NIH would approve his proposal?
What did the doctor wonder whether the NIH would approve?

13.
After reading the report, Rick wondered whether the mayor had taken a bribe.
When did the Rick wonder whether the mayor had taken a bribe?
What did the Rick wonder whether the mayor had taken?

14.
At the summit, Bush wondered whether Barak would agree to a new time-table.
When did Bush wonder whether Barak would agree to the new time-table?
What did Bush wonder whether Barak would agree to?

15.
During the gale, the crew wondered whether the ship could survive the trip.
When did the crew wonder whether the ship could survive the trip?
What did the crew wonder whether the ship could survive?

16.
Hijacked, the crew wondered whether they would survive the ordeal.
When did the crew wonder whether they would survive the ordeal?
What did the crew wonder whether they would survive?

17.
As the market fell, the broker wondered whether he could recover his losses.
When did the broker wonder whether he could recover his losses?
What did the broker wonder whether he could recover?

18.
As the rider fell, the clown wondered whether he could distract the bull.
When did the clown wonder whether he could distract the bull?
What did the clown wonder whether he could distract?

19.
Before the opening, the soprano wondered whether she would get good reviews.
When did the soprano wonder whether she would get good reviews?
What did the soprano wonder whether she would get?

20.
After the cadenza, Yoyo wondered whether the audience enjoyed his effort.
When did Yoyo wonder whether the audience enjoyed his effort?
What did Yoyo wonder whether the audience enjoyed?

21.
Writing his story, Bernstein wondered whether the witness had any new facts.
When did Bernstein wonder whether the witness had any new facts?
What did Bernstein wonder whether the witness had?

22.
Reading the script, Crowe wondered whether the writer would revise scene 3.
When did Crowe wonder whether the writer would revise scene 3?
What did Crowe wonder whether the writer would revise?

23.
Before the race, Harold wondered whether the press would cover the event.
When did Harold wonder whether the press would cover the event?
What did Harold wonder whether the press would cover?

24.
Before upgrading, John wondered whether his PC could run the new program.
When did John wonder whether his PC could run the new program?
What did John wonder whether his PC could run?

25.
After the storm, Jack wondered whether his insurance would cover the damage.
When did Jack wonder whether his insurance would cover the damage?
What did Jack wonder whether his insurance would cover?

26.
During the riot, the warden if whether the felons would damage the prison.
When did the warden wonder if the felons would damage the prison?
What did the warden wonder if the felons would damage?

27.
Driving home, Peter wondered whether his team had beat the point spread.
When did Peter wonder whether his team had beat the point spread?
What did Peter wonder whether his team had beat?

28.
During dinner, the chef wondered whether the critic would enjoy the desert.
When did the chef wonder whether the critic would enjoy the desert?
What did the chef wonder whether the critic would enjoy?

Adjunct-island materials

1.
On Saturday, Fraser cleaned the bathroom while Judy mopped the kitchen.
When did Fraser clean the bathroom while Judy mopped the kitchen?
What did Fraser clean the bathroom while Judy mopped on Saturday?

2.
After lunch, Harry called the travel agent, while Marge studied the map.
When did Harry call the travel agent while Marge studied the map?
What did Harry call the travel agent while Marge studied after lunch?

3.
Last night, Lois worked on the report, while Jamie reviewed the lab notes.
When did Lois work on the report while Jamie reviewed the lab notes?
What did Lois work on the report while Jamie reviewed last night?

4.
Yesterday, Tom looked for receipts, while Julie prepared the tax forms.
When did Tom look for receipts while Julie prepared the tax forms?
What did Tom look for receipts while Julie prepared yesterday?

5.
This morning, Lilly beat the eggs, while Joe greased the cake pan.
When did Lilly beat the eggs while Joe greased the cake pan?
What did Lilly beat the eggs while Joe greased this morning?

6.
This evening, Cedric watched tv while Peter fixed dinner.
When did Cedric watch tv while Peter fixed dinner?
What did Cedric watch tv while Peter fixed this evening?

7.
Last night, Joseph made popcorn while Laurel started the VCR.
When did Joseph make popcorn while Laurel started the VCR?
What did Joseph make popcorn while Laurel started last night?

8.
This afternoon, Nancy baked a cake, while Jason decorated the patio.
When did Nancy bake a cake while Jason decorated the patio?
What did Nancy bake a cake while Jason decorated this afternoon?

9.
This morning, the professor napped, while the TA graded essays.
When did the professor nap while the TA graded essays?
What did the professor nap while the TA graded this morning?

10.
After dinner, Jacob washed dishes, while Saul prepared desert.
When did Jacob wash dishes, while Saul prepared desert?
What did Jacob wash dishes, while Saul prepared after dinner?

11.
This morning, Karin shoveled snow, while Louis scraped the windshield.
When did Karin shovel snow while Louis scraped the windshield?
What did Karin shovel snow while Louis scraped this morning?

12.
Before dinner, Betty built a fire, while Dolores cleaned the fish.
When did Betty build a fire, while Dolores cleaned the fish?
What did Betty build a fire, while Dolores cleaned before dinner?

13.
Yesterday, Ian sanded the counter, while Phillip installed a new sink.
When did Ian sand the counter, while Phillip installed the sink?
What did Ian sand the counter, while Phillip installed yesterday?

14.
Before class, Harry searched the catalog, while Bill scanned the bibliography.
When did Harry search the catalog while Bill scanned the bibliography?
What did Harry search the catalog while Bill scanned before class?

15.
Before surgery, The doctor scrubbed up, while the nurse checked the chart.
When did the doctor scrub up while the nurse checked the chart?
What did the doctor scrub up while the nurse checked before surgery?

16.
Before take-off, Jill checked the plane, while the Laurel studied a map.
When did Jill check the plane while the Laurel studied a map?
What did Jill check the plane while the Laurel studied before take-off?

17.
Yesterday, Mary studied the thesis, while Jonathan read a research paper.
When did Mary study the thesis while Jonathon read a research paper?
What did Mary study the thesis while Jonathon read yesterday?

18.
This morning, The cat hid in the grass, while the mole peeked from his hole.
When did the cat hide in the grass while the mole peeked from his hole?
What did the cat hide in the grass while the mole peeked from this morning?

19.
Yesterday, Allen read the review, while Randy revised the essay.
When did Allen read the review, while Randy revised the essay?
What did Allen read the review, while Randy revised yesterday?

20.
After closing, Fred counted the money, while Helen tallied receipts.
When did Fred count the money while Helen tallied receipts?
What did the Fred count the money while Helen tallied after closing?

21.
This morning, Chris addressed an envelope, while Mary wrote the check.
When did Chris address an envelope while Mary wrote the check?
What did Chris address an envelope while Mary wrote this morning?

22.
After lunch, Saul read the paper, while Ellen finished a crossword puzzle.
When did Saul read the paper while Ellen finished a crossword puzzle?
What did Saul read the paper while Ellen finished after lunch?

23.
This afternoon, John studied the report, while Sara reviewed her notes.

When did John study the report while Sara reviewed her notes?
What did John study the report while Sara reviewed this afternoon?

24.
This morning, Cheney read the report, while his advisors discussed the poll.
When did Cheney read the report while his advisors discussed the poll?
What did Cheney read the report while his advisors discussed after the speech?

25.
Earlier, Edith chatted with Matthew, while Bill fixed the broken cabinet.
When did Edith chat with Matthew while Bill fixed the broken cabinet?
What did Edith chat with Matthew while Bill fixed today?

26.
Last night, Sandra played the piano, while Jennifer sang an aria.
When did Sandra play the piano while Jennifer sang aria?
What did Sandra play the piano while Jennifer sang?

27.
After the blowout, Jim loosened the lug nuts, while Mona checked the spare.
When did Jim loosen the lug nuts while Mona checked the spare?
What did Jim loosen the lug nuts while Mona checked?

28.
After dinner, Sally read the paper, while Bill washed dishes.
When did Sally read the paper while Bill washed dishes?
What did Sally read the paper while Bill washed?

References

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191-238.

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Clarendon Press/Oxford University Press.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49a*(1), 5-28.

Balogh, J., Zurif, E. B., Prather, P., Swinney, D. A., & Finkel, L. (1998). Gap-filling and end-of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language, 61*(2), 169-182.

Berwick, R. C., & Weinberg, A. S. (1986). *The grammatical basis of linguistic performance: Language use and acquisition* (Vol. 11). Cambridge, MA: MIT Press.

Blumenthal, A. L. (1966). Observations with self-embedded sentences. *Psychonomic Science, 6*(10), 453-454.

Boeckx, C. (2001). *The contribution of satiation effects to a typology of islands.* Poster presented at the The 14th Annual CUNY Conference on Human Sentence Processing, Philadelphia, PA, March 15-17.

Braze, D. (1999). Garden path sentences and morphological frequency biases. *University of Connecticut Working Papers in Linguistics, 10*, 13-34.

Braze, D., Shankweiler, D. P., Ni, W., & Palumbo, L. C. (in press). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*.

Broca, P. (1988). On the speech center. In L. T. Benjamin, Jr. (Ed.), *A history of psychology: Original sources and contemporary research* (pp. 94-97). New York, NY, US: Mcgraw-Hill Book Company.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(1), 77-126.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language, 3*(4), 572-582.

Carlson, G. N. (1977). *Reference to kinds in english.* Unpublished Ph.D. dissertation, University of Massachusetts, Amherst.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (Perspectives in neurolinguistics, neuropschology, and psycholinguistics, pp. 275-307). New York: Academic Press.

Chomsky, N. (1964). Current issues in linguistic theory. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language* (pp. xx-xx). Englewood Cliffs, NJ: Prentice Hall.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (Eds.), *A festschrift for morris halle* (pp. 232-286). New York: Holt, Rinehart, and Winston.

Chomsky, N. (1975). *The logical structure of linguistic theory.* New York, NY: Plenum Press.

Chomsky, N. (1977). On wh-movement. In P. W. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax* (pp. xx-xx). New York: Academic Press.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Barriers.* Cambridge, MA: MIT Press.

Chomsky, N., & Miller, G. A. (1958). Finite state languages. *Information and Control, 1*, 91-112.

Coulson, S., King, J. W., & Kutas, M. (1998). Erps and domain specificity: Beating a straw horse. *Language and Cognitive Processes, 13*(6), 653-672.

Cowart, W. (1996). *Experimental syntax: Applying objective methods to sentence judgments.* Thousand Oaks, CA: Sage Publications.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 94-128). Cambridge UK: Cambridge University Press.

Crain, S., Ni, W., Shankweiler, D. P., Conway, L., & Braze, D. (1996). Meaning, memory and modularity, *Proceedings of the nels 26 sentence processing workshop* (Mit occasional papers in linguistics, Vol. 9, pp. 27-44). Cambridge, MA: MIT Working Papers in Linguistics.

Crain, S., Shankweiler, D. P., Macaruso, P., & Bar-Shalom, E. (1990). Working memory and comprehension of spoken sentences: Investigations of children with reading disorder. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short term memory* (pp. 477-508). New York, NY: Cambridge University Press.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 320-358). Cambridge: Cambridge University Press.

Cutler, A. (1983). Lexical complexity and sentence processing. In G. B. Flores D'Arcais & R. J. Jarvella (Eds.), *The processes of language understanding* (pp. 43-79). New York: Wiley.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22*(1), 75-87.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden path sentences. *Journal of Memory and Language, 30*(6), 725-745.

Ferreira, F., & Henderson, J. M. (1998). Syntactic reanalysis, thematic processing, and sentence comprehension. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 73-100). Dordrecht: Kluwer Academic Publishers.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. A., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*(5), 414-420.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry, 9*, 427-473.

Fodor, J. D., & Ferreira, F. (1998). *Reanalysis in sentence processing*. Dordrecht: Kluwer Academic Publishers.

Fodor, J. D., & Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research, 23*(5), 407-434.

Fodor, J. D., & Inoue, A. (2000). Syntactic features in reanalysis: Positive and negative symptoms. *Journal of Psycholinguistic Research, 29*(1), 25-36.

Francis, N. W., & Kucera, H. (1982). *Frequency analysis of english usage: Lexicon and grammar*. Boston: Houghton-Mifflin.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*(2), 178-210.

Garrett, M. F., Bever, T. G., & Fodor, J. A. (1966). The active use of grammar in speech perception. *Perception and Psychophysics, 1*(1), 30-32.

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

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

Gibson, E., & Thomas, J. (1996). The processing complexity of english center-embedded and self-embedded structures. *MIT Occasional Papers in Linguistics, 9*, 45-71.

Gorrell, P. G. (1987). *Studies of human syntactic processing: Ranked-parallel versus serial models.* Unpublished Ph.D. dissertation, The University of Connecticut, Storrs, CT.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Speech acts* (Syntax and semantics, Vol. 3, pp. 41-58). New York: Academic Press.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases.* Unpublished Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain, 121*(6), 1133-1142.

Hiramatsu, K. (1998). *Natural classes of subjacency violations: Evidence from syntactic satiation.*Unpublished manuscript, University of Connecticut.

Hiramatsu, K. (2000). *Competence vs. Performance: Evidence from children's and adults' grammaticality judgments.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs.

Holmes, V. M., & Forster, K. I. (1972). Click location and syntactic structure. *Perception and Psychophysics, 12*(1-a), 9-15.

Huang, C. T. J. (1982). *Logical relations in chinese and the theory of grammar.* Unpublished Ph.D. dissertation, MIT, Cambridge, Mass.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Karlsen, B., & Gardner, E. F. (1985). *Stanford diagnostic reading test* (3rd ed.). San Antonio, Texas: Harcourt Brace Educational Measurement.

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy, 1*(1), 3-44.

Katz, J. J. (1964). Semi-sentences. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language* (pp. 400-416). Englewood Cliffs, NJ: Prentice Hall.

Kennedy, A. (1983). On looking into space. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (Perspectives in neurolinguistics, neuropsychology, and psycholinguistics, pp. 237-251). New York: Academic Press.

Kennedy, A., & Murray, W. S. (1987). The components of reading time: Eye movement patterns of good and poor readers. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 509-520). Amsterdam, Netherlands: Elsevier Science.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580-602.

Kluender, R. E. (1998). On the distinction between strong and weak islands: A processing perspective. In P. W. Culicover & L. McNally (Eds.), *The limits of syntax* (Syntax and semantics, Vol. 29, pp. 241-280). San Diego, CA: Academic Press.

Kluender, R. E., & Kutas, M. (1993a). Bridging the gap: Evidence from erps on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*(2), 196-214.

Kluender, R. E., & Kutas, M. (1993b). Subjacency as a processing phenomenon. *Language and Cognitive Processes, 8*(4), 573-633.

Konieczny, L., Hemforth, B., & Scheepers, C. (1997). The role of lexical heads in parsing. *Language and Cognitive Processes, 12*(2-3), 307-348.

Kratzer, A. (1995). Stage-level and individual-level predicates. In G. N. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 125-175). Chicago, IL: University of Chicago Press.

Kurtzman, H. S., & Crawford, L. F. (1990). Processing parasitic gaps, *Proceedings of nels 21* (pp. 217-231). Montreal, Canada.

Kutas, M., Federmeier, K. D., & Sereno, M. I. (1999). Current approaches to mapping language in electromagnetic space. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 359-392). Oxford: Oxford Univerisity Press.

Kutas, M., & Hillyard, S. A. (1983). Event related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition, 11*(5), 539-550.

Lasnik, H., & Saito, M. (1984). On the nature of proper government. *Linguistic Inquiry, 15*, 235-289.

Lasnik, H., & Saito, M. (1992). *Move /alpha.* Cambridge, MA: MIT Press.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55-75). Oxford, England UK: Anonima Romana.

Lukatela, K., Shankweiler, D. P., & Crain, S. (1995). Syntactic processing in agrammatic aphasia by speakers of a slavic language. *Brain and Language, 49*(1), 50-76.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676-703.

Marks, L. E. (1968). Scaling of grammaticalness of self-embedded english sentences. *Journal of Verbal Learning and Verbal Behavior, 7*(5), 965-967.

McConkie, G. W., Hogaboam, T. W., Wolverton, G. S., Zola, D., & Lucas, P. A. (1979). Toward the use of eye movements in the study of language processing. *Discourse Processes, 2*(3), 157-177.

McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 432-460.

McKinnon, R., & Osterhout, L. E. (1996). Constraints on movement phenomena in sentence processing: Evidence from event related brain potentials. *Language and Cognitive Processes., 11*(5), 495-523.

McNally, L. (1994). Adjunct predicates and the individual/stage distinction. In E. Duncan, D. Farkas, & P. Spaelti (Eds.), *The proceedings of the twelfth west coast conference on formal linguistics* (pp. 561-576). Stanford, CA: Center for the Study of Language and Information.

McNulty, E. M. (1988). *The syntax of adjunct predicates.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT.

Merchant, J. (1999). *The syntax of silence: Sluicing, islands, and identity in ellipsis.* Unpublished Ph.D. dissertation, University of California Santa Cruz.

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. R. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 419-491). New York: John Wiley and Sons.

Milsark, G. L. (1974). *Existential sentences in english.* Unpublished Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Miyamoto, Y. (1994). *Secondary predicates and tense.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT.

Murray, W. S., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research, 27*(1), 1-22.

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*(2), 151-165.

Ni, W. (1991). *Referentiality and garden path phenomena.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT.

Ni, W., Constable, T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., Shaywitz, B. A., & Gore, J. C. (2000). An event-related neuroimaging study: Distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience, 12*(1), 120-133.

Ni, W., Crain, S., & Shankweiler, D. P. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes, 11*(3), 283-334.

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. P. (1998). Anomalous strings: Eye movement patterns. *Journal of Psycholinguistic Research, 27*(5), 515-539.

Phillips, C., & Wong, K. (2000). *Island constraints in parsing: How the parser solves a look-ahead problem.* Poster presented at the 13th CUNY Conference on Human Sentence Processing, La Jolla, CA, March 30 - April 1.

Rapoport, T. R. (1990). Secondary predication and the lexical representation of verbs. *Machine Translation, 5*(1), 31-55.

Rapoport, T. R. (1991). Adjunct-predicate licensing and d-structure. In S. D. Rothstein (Ed.), *Perspectives on phrase structure: Heads and licensing* (Syntax and semantics, Vol. 25, pp. 159-187). San Diego, CA: Academic Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14*(3), 191-201.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53a*(4), 1061-1080.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*(3-4), 21-49.

Rizzi, L. (1990). *Relativized minimality*. Cambridge, MA: MIT Press.

Ross, J. R. (1967). *Constraints on variables in syntax*. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology.

Ross, J. R. (1969). Guess who? In R. Binnick, A. Davison, G. Green, & J. Morgan (Eds.), *Proceedings of cls 5* (pp. 252-286).

Ross, J. R. (1979). Where's english? In C. J. Fillmore, D. Kempler, & W. S.-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 127-163). New York: Academic Press.

Ross, J. R. (1984). Inner islands, *Proceedings of bls 10*. University of California, Berkeley: Berkeley Linguistic Society.

Schutze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Shankweiler, D. P. (1989). How problems of comprehension are related to difficulties in decoding. In D. P. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (International academy for reseach in learning disabilities monograph series, Vol. 6, pp. 35-67). Ann Arbor, MI: University of Michigan Press.

Shankweiler, D. P., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition, 24*(1-2), 139-168.

Snyder, W. (2000). An experimental investigation of syntactic satiation effects. *Linguistic Inquiry, 31*, 575-582.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory and Cognition, 20*(1), 51-68.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*(4), 402-433.

Stepanov, A. (2001). *Cyclic domains in syntactic theory.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT.

Stowe, L. A. (1985). Parsing wh constructions: Evidence for on-line gap location. *Language and Cognitive Processes., 1*(3), 227-245.

Swainson, B., & Shankweiler, D. P. (2001). *Working memory and reading comprehension: An examination of the reliability and validity of auditory sentence span.*Unpublished manuscript, Haskins Laboratories, New Haven, Connecticut.

Swinney, D. A., Zurif, E. B., Prather, P., & Love, T. (1996). Neurological distribution of processing resources underlying language comprehension. *Journal of Cognitive Neuroscience, 8*(2), 174-184.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye tracking study. *Journal of Memory and Language, 35*(3), 454-475.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*(3), 285-318.