NORTHWESTERN UNIVERSITY

The Structure and Real-Time Comprehension of
Quantifier Scope Ambiguity

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

CATHERINE ANDERSON

EVANSTON, ILLINOIS

December 2004

Volume I
ABSTRACT

The Structure and Real-Time Comprehension of Quantifier Scope Ambiguity

Catherine Anderson

This dissertation argues that the real-time comprehension of sentences with quantifier scope ambiguities is governed by the properties of the human sentence processing mechanism, which is driven by structural principles. Chapter 1 presents an overview of existing theories of quantifier representation and models of sentence comprehension. The chapter concludes by proposing a principle of Processing Scope Economy. According to the principle, quantifier scope ambiguity is a syntactic phenomenon and computing quantifier scope relations in real time involves computing syntactic representations. Because inverse-scope representations are more complex than surface-scope representations, computing inverse scope incurs a processing cost.

Chapter 2 provides evidence from three questionnaires and five self-paced reading experiments that inverse-scope interpretations are dispreferred by the processor and that assigning inverse scope consumes more processing resources than assigning surface scope does. This processing cost is incurred not only when a doubly quantified sentence is presented in isolation, but also when it is presented in a context that supports the inverse-scope interpretation, and even when the sentence is unambiguously inverse scope. These results indicate that abstract linguistic structure plays a central role in the comprehension of scope-ambiguous sentences and support the principle of Processing Scope Economy.

Chapter 3 examines Fox’s (2000) grammatical principle of Scope Economy, which proposes that a syntactic scope-shifting operation like Quantifier Raising is prohibited if it is semantically vacuous. The results of two questionnaires and three self-paced reading experiments suggest, however, that inverse-scope configurations are indeed permitted in sentences that are scopally commutative. Rather than a principle of Grammatical Scope Economy, whereby the well-formedness of a syntactic representation is governed by its semantics, the results offer support for the principle of Processing Scope Economy, which proposes that inverse-scope configurations are fully grammatical but dispreferred.

Taken together, the data presented in Chapters 2 and 3 offer a coherent picture of a model of on-line sentence comprehension that is crucially informed by syntactic structure. The properties of this model then allow us to locate the effects of Scope Economy in the processor rather than in the grammar.
Acknowledgements

In the Spring of 2000, Michael Walsh Dickey taught a course at Northwestern on Sentence Processing. It was in that course that I became excited about using experimental methods to investigate syntactic and semantic problems. In the Fall of that year, Chris Kennedy and Lance Rips offered a seminar on Quantifier Scope, which provided a lovely, juicy problem in syntax-semantics to investigate. It is no surprise, then, that these three people formed my dissertation committee. Lance has provided quiet and patient support throughout the project, listening to me ramble about my latest experiments in the corridors of Swift-Cresap and in the Kellogg cafeteria. I am thankful for his helpful questions that pointed out blind spots in the research. Chris has been an advisor to me since the very beginning of my graduate program, through many courses, a qualifying paper, and finally the dissertation. I was sometimes daunted when he returned a manuscript covered with comments in pink or orange or turquoise ink, but those comments invariably led to stronger arguments and sharper prose. And I am more grateful to Mike than words can express for his support in every way. Mike has read, commented on and revised innumerable drafts, papers, chapters, funding applications and abstracts; instructed me in the mysteries of experiment-running software; given realistic and supportive career advice; required me to be rigorous in thinking through the empirical predictions of my proposals and the theoretical consequences of my data; and more than once has talked me through being discouraged and overwhelmed so that I could get back to work. He has been an invaluable advisor and a true friend.

Of course, there are many others at Northwestern to whom I owe thanks, particularly Gregory Ward who, as chair, has been a tireless advocate for his students and faculty. I am also grateful to Ann Bradlow, Jeff Lidz and Janet Pierrehumbert for their teaching and advice. The students in the Department of Linguistics are a remarkable group, and I thank Tessa Bent, Elisa Sneed, Kristen Syrett and Saundra Wright for their friendship, and especially Ann Bunger for being a kindred spirit. Thanks also to Katy Carlson for her friendship and collaboration.

This work has also benefitted from helpful conversations with Danny Fox and Mark Steedman, and from the questions and comments of linguists at the 2004 LSA Annual Meeting and the 2004 CUNY Conference on Sentence Processing.

For extra-linguistic support, generous friendship, and spirited recreation I am grateful to the members of the choir at St. Mark’s Episcopal Church in Evanston, especially Mary Lew and Bob Finster, Sandra and Andy Howard, and Carol Martin. Above all, I am
thankful for the society, help and comfort in both prosperity and adversity of my beloved husband, Derek.
Table of Contents

List of Figures .................................................................................................................................................. ix
List of Tables .................................................................................................................................................... xi
CHAPTER 1:  UNDERSTANDING SCOPE AMBIGUITY .......................................................... 1
  Introduction .................................................................................................................................................... 2
    The Facts about Scope Ambiguity ............................................................................................................ 3
    Technical and Notational Matters ........................................................................................................ 4
      Type Composition ................................................................................................................................ 4
      Lambda Notation ................................................................................................................................. 6
  The Semantics of Quantifier Scope ........................................................................................................ 6
    Generalized Quantifiers ......................................................................................................................... 6
    Interpreting Quantifiers in Object Position ......................................................................................... 9
      Quantifier Raising at LF ..................................................................................................................... 9
    Flexible Types ....................................................................................................................................... 13
    Combinatory Categorial Grammar .................................................................................................... 16
  Summary ....................................................................................................................................................... 18
  Quantifier Scope and Sentence Processing ....................................................................................... 19
    Serial vs. Parallel Processing .............................................................................................................. 20
    Linguistic Structure and Conceptual Structure ............................................................................... 21
    Sentence Processing and Linguistic Theory ..................................................................................... 22
  A Proposal for Understanding Scope Ambiguity in Real Time ....................................................... 23
CHAPTER 2:  COMPUTING SCOPE RELATIONS .............................................................. 24
  Introduction .................................................................................................................................................. 25
    Background: Theories of Quantifier Scope Comprehension .......................................................... 26
      Hierarchical Principles Interact to Determine Scope Preferences .............................................. 26
      Non-Linguistic Information Determines Scope Preferences ....................................................... 27
      Multiple Constraints Interact to Determine Scope Preferences ................................................. 29
      LF Structure Determines Scope Preferences .............................................................................. 30
    Chapter Outline .................................................................................................................................... 31
  Experiments ............................................................................................................................................... 32
    Off-line Questionnaires ....................................................................................................................... 32
    Experiment 1 ............................................................................................................................................. 32
      Experiment 1 Method ......................................................................................................................... 32
      Experiment 1 Results ....................................................................................................................... 34
      Experiment 1 Discussion ................................................................................................................... 34
    Experiment 2 ............................................................................................................................................. 34
      Experiment 2 Method ....................................................................................................................... 35
Experiment 2 Results ................................................................. 39
Experiment 2 Discussion ......................................................... 41
Experiment 3 Discussion .......................................................... 42
Experiment 3 Method ............................................................... 42
Experiment 3 Results .............................................................. 43
Experiment 3 Discussion .......................................................... 44
Self-Paced Reading Experiments .............................................. 44
Experiment 4 Discussion .......................................................... 44
Experiment 4 Method ............................................................... 45
Experiment 4 Results .............................................................. 47
Experiment 4 Discussion .......................................................... 54
Experiment 5 Discussion .......................................................... 56
Experiment 5 Method ............................................................... 57
Experiment 5 Results .............................................................. 60
Experiment 5 Discussion .......................................................... 67
Experiment 6 Discussion .......................................................... 68
Experiment 6 Method ............................................................... 69
Experiment 6 Results .............................................................. 70
Experiment 6 Discussion .......................................................... 73
Experiment 7 Discussion .......................................................... 73
Experiment 7 Method ............................................................... 74
Experiment 7 Results .............................................................. 75
Experiment 7 Discussion .......................................................... 79
Experiment 8 Discussion .......................................................... 80
Experiment 8 Method ............................................................... 80
Experiment 8 Results .............................................................. 83
Experiment 8 Discussion .......................................................... 86
General Discussion .................................................................... 86
CHAPTER 3: SCOPE ECONOMY IN REAL TIME .......................... 88
Theoretical Background ......................................................... 90
Experiments .......................................................................... 97
Stimuli .............................................................................. 97
Predictions .......................................................................... 100
Experiment 9 Discussion .......................................................... 102
Experiment 9 Method ............................................................... 102
Experiment 9 Results .............................................................. 104
Experiment 9 Discussion .......................................................... 105
Self-paced Reading Experiment Design ................................. 106
Experiment 7 ................................................................. 110
  Experiment 7 Results .................................................. 111
  Experiment 7 Discussion ............................................. 117
Experiment 10 ............................................................. 127
  Experiment 10 Method ............................................... 129
  Experiment 10 Results .............................................. 130
  Experiment 10 Discussion ......................................... 131
Experiment 11 ............................................................. 132
  Experiment 11 Method ............................................... 133
  Experiment 11 Results .............................................. 134
  Experiment 11 Discussion ......................................... 138
Experiment 12 ............................................................. 139
  Experiment 12 Method ............................................... 140
  Experiment 12 Results .............................................. 141
  Experiment 12 Discussion ......................................... 145
Experiment 13 ............................................................. 150
  Experiment 13 Method ............................................... 150
  Experiment 13 Results .............................................. 151
  Experiment 13 Discussion ......................................... 159
General Discussion ..................................................... 160
CHAPTER 4: CONCLUSIONS............................................... 166
The Role of LF Structure in Real-time Semantic Inferencing .......... 167
  LF Complexity Leads to Processing Difficulty ................. 167
  LF Structure is Crucial for Semantic Inferencing ............. 168
The Role of Semantics in Determining Well-formedness of LF Structure ............................................... 169
  Inverse-scope Interpretation is Available Following Ellipsis Sentence ............................................... 170
  Scope Economy Reflects a Processing Preference ............... 171
Future Research .......................................................... 172
Conclusions ............................................................... 173
References ................................................................. 175

Note: Appendix is available at http://aix1.uottawa.ca/~canderso/diss
List of Figures

Figure 1. Mean residual reading time for disambiguating sentence, Experiment 4A. ...... 50
Figure 2. Mean residual reading times for disambiguating sentence, Experiment 4B. ...... 53
Figure 3. Mean residual reading times for disambiguating sentence, Experiment 5A. ...... 61
Figure 4. Mean residual reading times for disambiguating sentence, Experiment 5B. ...... 65
Figure 5. Mean residual reading time for quantified sentence, contingent on comprehension question response, Experiment 6. ............................................ 71
Figure 6. Mean residual reading times for words in quantified sentence, contingent on comprehension question response, Experiment 6. ............................................ 72
Figure 7. Mean residual reading time for quantified sentence, Experiment 7. ............... 76
Figure 8. Mean residual reading times for ambiguous quantified sentence, contingent on comprehension question response, Experiment 7. ............................................ 77
Figure 9. Mean residual reading time for quantified sentence, Experiment 8. .............. 84
Figure 10. Mean residual reading time for disambiguating sentence, Experiment 8. ...... 85
Figure 11. Mean residual reading time for quantified sentence, Experiment 7. .......... 112
Figure 12. Mean residual reading time for ellipsis sentence, Experiment 7. .............. 114
Figure 13. Mean residual reading times for quantified sentence, Experiment 11........... 135
Figure 14. Mean residual reading time for ellipsis sentence, Experiment 11.............. 136
Figure 15. Mean residual reading time for quantified clause, Experiment 12............ 142
Figure 16. Mean residual reading time for ellipsis clause, Experiment 12................. 143
Figure 17. Mean residual reading time for entire quantified sentence, Experiment 13. 152
Figure 18. Mean residual reading time for entire ellipsis sentence, Experiment 13. 153

Figure 19. Mean residual reading time for each word in quantified sentence, Experiment 13. 156

Figure 20. Mean residual reading time for each word in ellipsis sentence, Experiment 13. 158
List of Tables

Table 1. Experiment 2 Responses to Comprehension Question................................. 40

Table 2. Experiment 4A Mean residual reading times for a...every quantified sentence (ms) ................................................................. 48

Table 3. Experiment 4A Proportion of responses to comprehension question following a...every sentence .......................................................................................... 49

Table 4. Experiment 4B Mean residual reading times for every...a quantified sentence (ms) .................................................................................................................................. 51

Table 5. Experiment 4B Proportion of responses to comprehension question following every...a sentence .............................................................................................. 54

Table 6. Experiment 5A: Proportions of responses to comprehension question following a...every sentence in context .................................................................................. 63

Table 7. Experiment 5B Proportions of responses to comprehension question following every...a sentence in context .................................................................................. 66

Table 8. Experiment 7 Proportion of responses to comprehension question ............... 79

Table 9. Experiment 9 responses to comprehension question ..................................... 105

Table 10. Predicted Results for Self-Paced Reading Experiments ................................. 110

Table 11. Experiment 7 responses to comprehension question ..................................... 116

Table 12. Relationship between residual reading times and plausibility ranking ........ 131

Table 13. Experiment 11 responses to comprehension question .................................. 138

Table 14. Experiment 12 responses to comprehension question .................................. 145

Table 15. Experiment 13 responses to comprehension question .................................. 155
CHAPTER 1:
UNDERSTANDING SCOPE AMBIGUITY
Introduction

The research presented in this dissertation consists of a set of psycholinguistic experiments designed to investigate the nature of the representation and computation of quantifier scope relations in natural language. The phenomenon of quantifier scope ambiguity offers a rich opportunity for linguistic investigation since it lies squarely at the interface between form and meaning. There is still considerable debate among syntacticians and semanticists as to whether a sentence’s surface structure maps directly and transparently onto a semantic representation, or whether some linguistic level of representation intervenes between the visible surface form and the interpretation of a sentence. A scope-ambiguous sentence, where multiple truth conditions can arise from a single surface form, can serve as a test case for theories of the syntax-semantics interface. Some current theories of quantifier interpretation hold that the locus of quantifier scope ambiguity is at logical form (LF), a level of representation that mediates between surface syntax and truth-conditional semantics, while in theories that posit no such level, the ambiguity arises in the derivation of the surface syntax, or in the mapping from the syntax to the semantic representation.

Quantifier scope ambiguity is also important for psycholinguistics since it provides an opportunity to study the extent to which linguistic structure and conceptual knowledge contribute to on-line interpretive decisions. Any psycholinguistic insights that we can gain into the real-time comprehension of scope-ambiguous sentences might help adjudicate between linguistic theories that posit a covert level of syntactic representation as an interface to the semantics and those that do not. Based on the results of a set of questionnaire and self-paced reading experiments, I will argue for the central role of linguistic structure in sentence comprehension, and for the importance of psycholinguistic principles in evaluating theoretical proposals.

The structure of the dissertation is the following. In Chapter 1 I describe three linguistic accounts of scope ambiguity. I also introduce the main theories of sentence processing, with a focus on the place of syntactic structure in each model of real-time comprehension. Chapter 2 presents a set of experiments from whose results I conclude that computing quantifier scope relations relies crucially on linguistic structure and not just general conceptual knowledge. I propose a principle of Processing Scope Economy that claims that the processor prefers to construct the simplest possible scope configuration, and that more complex scope configurations incur a processing cost. In Chapter 3 I examine Fox’s theory of Scope Economy (2000), a linguistic account that
proposes that syntactic quantifier scope relations are constrained by their truth conditions. The results of the experiments that I report in Chapter 3 show that my principle of Processing Scope Economy accounts better for the facts about the real-time comprehension of doubly quantified sentences than Fox’s grammatical principle does. I argue that the syntactic well-formedness of quantified sentences is not determined by their semantics, and that apparent grammatical Scope Economy effects are a side effect of how parallel syntactic structures are processed in real-time. In Chapter 4 I consider the consequences of my proposal for linguistic and psycholinguistic theories of quantifier scope.

The Facts about Scope Ambiguity

The study takes its starting point from the observation that a sentence such as (1) containing two quantified expressions is ambiguous.

(1) A climber scaled every cliff.

One interpretation of (1) is that there exists a single energetic climber who scaled all the cliffs in the relevant domain. This interpretation can be represented using logical notation as shown in (2).

(2) $\exists x \ [\text{climber}(x) \land \exists y \ [\text{cliff}(y) \implies \text{scaled}(x,y)] ]$

The interpretation represented in (2) is called the surface-scope interpretation because the quantifiers in the semantic representation appear in the same hierarchical configuration as in the surface syntax given in (1), such that the first quantifier (the existential) has scope over the second (the universal).

The second possible interpretation of (1), represented in (3), corresponds to a scenario in which each cliff was scaled by some possibly different climber.

(3) $\exists x \ [\text{cliff}(x) \implies \exists y \ [\text{climber}(y) \land \text{scaled}(y,x)] ]$

This interpretation is referred to as the inverse-scope interpretation because the order of the quantifiers in the semantic representation is inverted from the order of the quantified
DPs in the surface syntax, such that the universal, which appears second in the surface syntax, has scope over the first-position existential.

Many linguists have argued that the ambiguity of a doubly quantified sentence such as (1) arises at the syntactic level of Logical Form (LF). In other words, the single surface syntactic structure of (1) maps onto more than one possible LF structure, and these multiple LFs lead to multiple distinct interpretations of the sentence. On the other hand, others have proposed that scope ambiguity is actually a surface-structure ambiguity arising from the multiple possible derivations of the surface syntax, while still others have argued that a single, unambiguous syntactic structure can map onto multiple distinct truth conditions. I describe representative examples of each of these accounts of quantifier scope ambiguity below.

Technical and Notational Matters

Type Composition

Throughout the theoretical discussion I shall assume the correctness of Heim & Kratzer’s (1998) version of Montague’s type-driven compositional semantics (Montague, 1974), whereby the truth-conditions of a sentence are derived from a composition of the syntactic constituents of the sentence. The system contains two basic semantic types – individuals labeled e, and truth-values labeled t. The individual type corresponds to entities in the world, such as people and objects. Truth-values are either 0 or 1, that is, either false or true.

This semantic system also includes functions that map between types: if $\square$ and $\square$ are types, then $<\square, \square>$ is a type whose denotation is a function from expressions of type $\square$ to expressions of type $\square$. For example, a function $<e,t>$ maps from individuals to truth-values. To see how function application represents semantic composition in natural language, let us consider sentence (4).

(4) Vladimir snores.

Assume that the proper name Vladimir picks out exactly one person in our universe; that person is an individual of type e. The intransitive verb snores serves as a function from individuals to truth-values, that is, the function of type $<e,t>$ depicted in (5).
Snores thus takes Vladimir as its argument and maps to 1 if it is the case that Vladimir snores, and 0 in the case that Vladimir does not snore. This semantic composition can be illustrated using a tree diagram as shown in (6):

Transitive verbs, of course, require a more complex semantic type. Sentence (7) includes two individuals, Vladimir and Katya.

The individuals Vladimir and Katya have semantic type e, and the semantic type of the entire sentence is a truth-value t. Examining the tree diagram in (8),

we can see that Vladimir must combine with a constituent of type <e,t> to give a truth-value for the sentence. The semantic type of loves, then, must map from individuals to properties, that is, from type <e> to type <e,t>. We can conclude that transitive verbs have the semantic type <e,<e,t>>.
**Lambda Notation**

One further notational matter will allow us to conveniently describe the semantics of quantifiers. This presentation is again based on Heim & Kratzer (1998). A function can be represented using the notation shown in (9).

\[ [a : j . g] \]

In this notation, \(a\) stands for the argument of the function, \(j\) for the domain\(^1\), and \(g\) for the value. Let us suppose that our discourse universe \(D\) contains exactly two individuals, Vladimir and Katya, and that Vladimir snores but Katya does not. So the function in (10) maps to 1 if \(x = \text{Vladimir}\) and maps to 0 if \(x = \text{Katya}\).

\[ [\exists x : x \in D. x \text{ snores}] \]

A two-place predicate like *loves* can also be represented using \(\lambda\) notation, as shown in (11).

\[ [\exists x . \exists y . y \text{ loves } x] \]

The function in (11) maps to 1 when it applies to two individuals \(x\) and \(y\) and \(y\) loves \(x\), and otherwise maps to 0. I take advantage of \(\lambda\) notation below in characterizing the linguistic representations of scope ambiguous sentences.

**The Semantics of Quantifier Scope**

**Generalized Quantifiers**

While the DPs *Vladimir* and *Katya* denote individuals in the world and thus have the semantic type \(e\), not everything that is a syntactic DP denotes an individual. For

---

\(^1\) It is often possible and convenient to omit the specification of the domain, especially when it is as simple as “\(x\) is an individual in the domain of discourse”.
example, quantificational DPs such as *every salmon, most spiders, or some students* do not denote individuals; it is therefore inappropriate for them to have type \(<e>\). Mostowski (1957) and Lindström (1966) proposed that quantified DPs be treated instead as generalized quantifiers with the semantic type \(<<e,t>,t>,t>\), such that a quantified DP maps from predicates to truth-values. In this way, a DP serves as a function that takes a predicate (often a VP) as its argument. The tree diagram in (13) illustrates the Montagovian semantic composition for sentence (12).

(12) Every linguist snores.

\[
\begin{tikzpicture}
  \node (e) {\textit{every}};
  \node (l) {\textit{linguist}} at (2,0);
  \node (s) {\textit{snores}} at (4,0);
  \draw (e) -- (l) -- (s);
\end{tikzpicture}
\]

Just as above in example (6), the intransitive verb *snores* has the semantic type \(<e,t>,t>\), since it maps from an individual to a truth-value (1 if the individual snores, 0 if not). The noun *linguist* also maps from individuals to truth-values (1 if the individual is a linguist, 0 if not), so it also has the type \(<e,t>\). The subject DP *every linguist* is not an individual but a generalized quantifier, which means that it takes the intransitive verb *snores* as its argument, unlike in (6), where the subject *Vladimir* was an individual that served as the argument to the function *snores*. The quantified DP therefore has the semantic type \(<<e,t>,t>,t>\) and maps from one-place predicates to truth-values. We can conclude, then, that the determiner *every* has the type \(<<e,t>,<e,t>,t>,t>\) and the denotation shown in (14):

(14) \[ [\text{every}] = \\square f : f \in D_{<e,t>} \cdot [ \square g : g \in D_{<e,t>} \cdot \text{for all } x \in D_e \text{ such that } f(x)=1, g(x)=1 ] \]

In (12)–(13), *every* denotes the relation between linguists and things that snore, both of which have type \(<e,t>\).
Treating quantified DPs as generalized quantifiers of semantic type $<<e,t>,t>$ is adequate for quantified phrases in subject position, but is problematic for quantifiers in object position, such as in (15).

(15) Vivek fixed every computer.

We saw above in (8) that transitive verbs have the semantic type $<e,<e,t>>$, mapping from individuals to one-place predicates. Obviously this is not the right type to combine with a generalized quantifier, as (16) illustrates.

(16)

Clearly a quantifier of type $<<e,t>,t>$ cannot be interpreted by semantic composition if it is in object position. There are three possibilities for making the object DP interpretable within this semantic system. One is to allow the quantified phrase to move to a position where it may combine with a constituent of the appropriate type. This is the approach taken by May (1977; 1985). Another is to leave the quantified phrase in object position, assigning to it some semantic type of a higher order than $<<e,t>,t>$ and possibly also giving the verb a more complex semantic type, as proposed by Partee and by Hendriks (Hendriks, 1987; Partee, 1986; Partee & Rooth, 1983). The third choice is to allow more than one way to combine the syntactic constituents, which leads to more than one option for semantic composition, as proposed by Steedman (2000; 2003). All three of these proposals are described below.
Interpreting Quantifiers in Object Position

Quantifier Raising at LF

In the example derivations that follow, I use Heim & Kratzer’s modernized version of May’s theory, which takes advantage of the VP-internal subject hypothesis (Kitagawa, 1986; Koopman & Sportiche, 1991; Kuroda, 1988), whereby a verb’s subject argument originates inside the VP and moves to a position higher in the tree for case or agreement purposes.

May (1977) argued that the syntactic operation of quantifier raising (QR) takes a quantified DP from its uninterpretable object position in the surface structure and moves it covertly to an LF position where it can be interpreted. Heim and Kratzer (1998, p. 186) give a definition for a Predicate Abstraction Rule that produces the correct semantic types in a variable-binding configuration. The resulting structure, in which every denotes the relationship between two one-place predicates, namely, computers and things that Vivek fixed, is shown in (17).

(17)

Interpreting Quantifiers in Object Position

Quantifier Raising at LF

May (1977) argued that the syntactic operation of quantifier raising (QR) takes a quantified DP from its uninterpretable object position in the surface structure and moves it covertly to an LF position where it can be interpreted. Heim and Kratzer (1998, p. 186) give a definition for a Predicate Abstraction Rule that produces the correct semantic types in a variable-binding configuration. The resulting structure, in which every denotes the relationship between two one-place predicates, namely, computers and things that Vivek fixed, is shown in (17).

(17)
As (17) shows, the object DP every computer is raised to a position adjoining the phrase where it originated. It leaves a trace of type $<$e$>$ that is bound by a variable binder (indicated by “1” in the tree) adjoined to the phrase immediately below the raised DP. The constituent containing the trace is an expression of type $<$e,$t$$>$ with the meaning $[\lambda x . \text{Vivek fixed } x]$. This function serves as the argument to the QRed generalized quantifier which maps from functions of type $<$e,$t$$>$ to truth values.

The QR account provides a natural explanation for the ambiguity of doubly quantified sentences such as (18).

(18) A climber scaled every cliff.

The two interpretations of (18) result from the two possible LF landing sites of the object phrase that undergoes QR.

If the DP every cliff undergoes the shortest possible QR move it is still inside VP at LF, as shown in (19), while a climber must move out of VP to the appropriate subject position.

(19)
In this position, *every cliff* is interpreted inside the scope of the DP *a climber*, giving rise to the surface-scope interpretation where a single climber scaled all the cliffs.

On the other hand, if an additional QR move raises the object DP *every cliff* to a position higher than *a climber*, as shown in (20), *every cliff* is interpreted with wider scope than *a climber*, which leads to the inverse-scope interpretation where climbers vary with cliffs.

An alternate possibility exists for achieving inverse scope. It is possible that sentence (18) has the configuration shown in (19), where *a climber* occupies the subject position in the surface syntax but is reconstructed at LF, possibly via Quantifier Lowering (QL) (May, 1977, 1985), inside the scope of *every cliff*. The DP is interpretable in this
position because it has the semantic type of a generalized quantifier, \(<e,t>,t>\). On this account, the syntactic representation of the inverse-scope interpretation is no more complex than the surface-scope representation, but its derivation is more complex since one additional operation is needed to reconstruct the subject DP in its base position, as shown in (21):

(21)

To summarize, on the QR account the two interpretations of a doubly quantified sentence like (18) result from the two possible LF configurations of the quantifiers. If a climber is in a higher position than every cliff at LF, then it takes scope over every cliff, leading to the surface-scope interpretation. But if every cliff is in an LF position higher than a climber, either because it undergoes additional QR or because a climber undergoes QL, then it takes scope over a climber, leading to the inverse-scope interpretation. This syntactic account of quantifier scope ambiguity depends crucially on the existence of LF, a covert level of syntactic representation that serves as the interface to the semantic component of the grammar. The inverse-scope interpretation differs from the surface-scope interpretation in that its LF has a more complex syntactic representation or derivation, even though the surface syntax is the same for both interpretations.
Flexible Types

Flexible types approaches (Hendriks, 1987; Partee, 1986; Partee & Rooth, 1983) allow quantified DPs to be interpreted in object position by “allowing each category to correspond to a family of types rather than just a single type” (Partee & Rooth, 1983, p. 357). In other words, there is a systematic ambiguity of types at least for quantifiers and determiners, and possibly for all DPs.² Heim & Kratzer succinctly capture this ambiguity for determiners:

(22) For every lexical item \( [\cdot,\cdot,\cdot] \) with a meaning of type \( <<e,t>,<e,t>,t>> \), there is a (homophonous and syntactically identical) item \( [\cdot,\cdot,\cdot] \) of type \( <<e,t>,<e,e,t>,<e,t>> \) (Heim & Kratzer, 1998, p. 182)

With rule (22) as part of the semantics the lexicon needs to list only determiners of type \( <<e,t>,<e,t>,t>> \), each of which can be lifted to a more complex type as needed according to the rule.

Let us consider how type-raising solves the problem of interpreting a quantifier in object position by returning to sentence (23).

(23) Vivek fixed every computer.

We saw above that the quantified phrase every computer could not combine with the transitive verb fixed since fixed has type \( e,<e,t> \) while every computer has type \( <<e,t>,t> \), as illustrated in (24).

(24) \[
\begin{array}{c}
<<e,t>,t> \\
<<e,t>,<e,t>,t>> \\
\text{every} \\
<<e,t>,t>> \\
\text{computer}
\end{array}
\]

² Partee and Rooth (1990) also propose a type-shifting approach to conjunctions, which is not relevant for my present purposes.
When every is in a subject DP, the denotation given in (25) provides the right semantic types for function application.

\[(25)\]  
\[
\text{in a subject DP} \\
[[\text{every}]] = \\
\begin{array}{c}
\square f : f \sqsubseteq D_{e,t} . \\
\square g : g \sqsubseteq D_{e,t} . \\
\forall x \sqsubseteq D_e \text{ such that } \\
f(x)=1, g(x)=1
\end{array}
\]

But once we apply the type-shifting rule to the determiner every, then it has the denotation given in (26), which allows the quantified phrase every computer to take a transitive verb of type \(<e,<e,t>>\) as its argument and give a one-place predicate of type \(<e,t>\), as illustrated in (27).

\[(26)\]  
\[
\text{in a narrow-scope object DP} \\
[[\text{every}]] = \\
\begin{array}{c}
\square f : f \sqsubseteq D_{e,t} . \\
\square g : g \sqsubseteq D_{e,t>}, \\
\forall x \sqsubseteq D_e \text{ such that } \\
f(x)=1, g(y)(x)=1
\end{array}
\]

\[(27)\]  
\[
\text{Vivek} \\
\begin{array}{c}
\llbracket x . x \text{ fixed every computer} \rrbracket
\end{array}
\]

Under the flexible-types approach, the ambiguity of a doubly-quantified sentence is attributed to the type ambiguity of the quantifiers. Example (29) illustrates the
composition for the surface-scope interpretation of (28). Notice that, syntactically, it is isomorphic with the singly-quantified sentence (27), since the object DP every cliff takes the verb scaled as its argument to create an expression of type $\langle e, t \rangle$. This expression indicates the property of having scaled every cliff, and serves as the argument for a climber.

\begin{equation}
(28) \quad \text{A climber scaled every cliff.}
\end{equation}

In order to achieve the inverse-scope interpretation, where every cliff takes scope over a climber, the determiner every must undergo an additional instance of type-shifting. The resulting denotation is given in (30).

\begin{equation}
(30) \quad \text{in a wide-scope object DP}
\end{equation}

\[
[[\text{every}]] = \left[ f : f \bowtie D_{\langle e, t \rangle} \cdot [ g : g \bowtie D_{\langle e, e \rangle} \cdot [ Q : Q \bowtie D_{\langle e, e, t \rangle} \cdot \text{for all } D \text{ such that } f(x)=1, Q(g(x))=1 ] ] \right]
\]

The function composition shown in (31) illustrates that every cliff takes scaled as its argument and creates an expression that takes the quantifier a climber as its argument. This composition gives every cliff scope over a climber, leading to the inverse-scope interpretation. Contrast this with (29), where the expression $[[x . x scaled every cliff]]$ served as the argument to a climber.
The flexible types account of quantifier scope ambiguity preserves syntactic simplicity at the expense of introducing semantic complexity. The inverse-scope interpretation of a doubly quantified sentence has the same syntax as the surface-scope interpretation, but a more complex semantic representation, since additional type-shifting operations are needed for the object phrase to take scope over the subject. The sentence is ambiguous because quantifier-determiners theoretically have infinitely many semantic types.

**Combinatory Categorial Grammar**

Combinatory Categorial Grammar (CCG) (Steedman, 2000, 2003; Steedman & Baldridge, 2003) posits only one level of representation, such that the interpretation of a sentence is read directly from the derivation of the surface syntax. There are no covert movement operations in this framework.

CCG is crucially different from flexible-types approaches to quantifier interpretation because it requires that every syntactic category map only to a single semantic type, and vice versa. However, CCG includes a syntactic type-raising operation in the lexicon. This operation allows an item of a given syntactic category to be raised to a more complex syntactic category, with the result that the item thereby also receives a
new semantic type. Quantified phrases in object position thus pose no problem for the semantics, since the syntactic type of the object DP can be raised to make the phrase interpretable.

Within CCG, the options for the scope of a quantified expression fall out naturally from the options for syntactic composition. In (32), for example, if the object DP every cliff combines syntactically with the verb scaled and this constituent then combines with the DP three climbers, then the subject three climbers takes scope over every cliff.

(32)  | Three climbers | scaled | every cliff |
      | S/(S\NP)      | (S\NP)/NP | (S\S/NP) |
      | S\NP          | S         |          |

Example (33) illustrates a second option for syntactic composition. The subject DP three climbers can combine first with the verb scaled, after which the resulting constituent combines with the object DP every cliff. Under this derivation, every cliff takes scope over the subject three climbers, leading to the inverse-scope interpretation.

(33)  | Three climbers | scaled | every cliff |
      | S/(S\NP)      | (S\NP)/NP | (S\S/NP) |
      | S\NP          | S         |          |

The sentence’s semantics can be read directly from the syntactic derivation, and neither derivation is more complex than the other.

While the ambiguity of a genuine doubly quantified sentence in CCG is a direct result of the syntactic type-raising mechanism and the option for multiple grammatical syntactic derivations, Steedman argues that the ambiguity of a sentence like (34) does not result from type-raising, since only every is a quantificational determiner, while the indefinite a is instead referential.

(34)  | A climber scaled every cliff.
Several other scholars have also taken this approach to indefinites (Fodor, 1982; Fodor & Sag, 1982; Park, 1995; VanLehn, 1978). In CCG the indefinite is represented in the syntax as a generalized Skolem term. This Skolem term operates on a variable by associating properties with it (e.g., in sentence (34) the property of being a climber) but, until is it specified, does not refer to any particular entity. The interpretation of sentence (34) depends on when in the syntactic derivation the Skolem term is specified.

Specification is “an ‘anytime’ operation that can occur at literally any point during grammatical derivation” (Steedman, 2003, p. 7). If the Skolem term is specified early in the derivation, then a unique climber is identified and the indefinite functions as a constant. This derivation leads to the interpretation where a single climber scaled all the cliffs. If, however, specification of the Skolem term is delayed until after the sentence’s constituents have combined syntactically, then the indefinite is interpreted inside the scope of the universal every cliff as a variable bound by the universal, leading to the interpretation where climbers vary with cliffs.

Steedman’s model of the grammar incorporates a model of sentence comprehension in which all possible structures and interpretations of an ambiguous string are constructed as the sentence is perceived incrementally. The competing parses are evaluated for plausibility with respect to the preceding context and to real-world knowledge. There is thus no structural reason for one interpretation of a quantified sentence to be preferred over another. It is only contextual plausibility that determines which interpretation is ultimately assigned to an ambiguous sentence.

The CCG account of scope ambiguity is quite similar to the choice function portion of Reinhart’s (1997) account. Reinhart proposes that indefinites are ambiguous between generalized quantifiers and choice functions. As a generalized quantifier, an indefinite takes scope according to its LF position, as the QR account predicts. As a choice function, on the other hand, an indefinite can take unboundedly wide scope while remaining in its surface position, depending on the point in the derivation at which the function is existentially closed. It may be impossible to distinguish empirically between the choice function account and the CCG account.

Summary

I have presented summaries of three theories that account for the ambiguity of a doubly quantified sentence like (35).
A climber scaled every cliff.

The QR account relies on the existence of a covert level of representation, LF, where the quantified phrases can be arranged in two distinct syntactic configurations, each of which maps directly to a distinct set of truth conditions. The inverse-scope LF is more complex (either in its representation or its derivation) than the surface-scope LF. Within the flexible types account, on the other hand, there is only one possible syntactic representation, but the mapping from the syntax to the semantics is complicated by the option for type-raising in the lexicon. The inverse-scope representation is more complex than the surface-scope because the quantifiers have a more complex semantic type. Finally, the CCG system has a simple syntax that maps directly onto the semantics. According to CCG the interpretation of the ambiguous sentence (35) depends on the timing of steps in the syntactic derivation.

At the outset I shall adopt the QR account as my working theory, including the assumption that the inverse-scope representation includes an additional QR operation that raises the object DP above the subject at LF. Nevertheless, I shall return to these three accounts to evaluate them in light of the empirical evidence that emerges from the experiments.

Quantifier Scope and Sentence Processing

Scope ambiguity offers a particularly promising avenue for research in sentence processing since it potentially involves covert structure. Surface-identical scope-ambiguous sentences thus offer an ideal test case for theories of human sentence processing by making it possible to focus on the role of information that is not in the surface syntax in real-time sentence comprehension. Psycholinguistic evidence that perceivers compute covert structure in real time would lend strong support to models of sentence comprehension where abstract linguistic structure plays a central role, since covert structure cannot be recovered from information in the surface signal such as co-occurrence frequencies, templates, or prosody. Such evidence would also offer validation to theories of the syntax-semantics interface that postulate a level of representation that mediates between the syntax and the semantics.

Conversely, if we fail to find evidence that perceivers do not compute covert structure in real time, this would weaken the case for structure-driven models of
comprehension by suggesting that covert structure is not computed automatically the way that surface syntactic relations are. Such a result would also widen the gap between theoretical models of linguistic (syntactic and semantic) competence and models of the real-time comprehension of syntactic objects.

Serial vs. Parallel Processing

While the literature offers many models of human sentence processing, research tends to cluster around two poles of a spectrum. At one end are serial, structure-driven models such as the Garden Path model of Frazier and her colleagues (Ferreira & Clifton, 1986; Frazier, 1987, 1990; Frazier & Clifton, 1996; Frazier & Fodor, 1978). In this model, structural principles guide the processor in pursuing a single parse for a sentence. The preferred analysis always has the simplest possible structure. Other sources of information such as contextual plausibility are used to evaluate the parse and to reanalyze it if necessary. Although the model was originally intended to describe the parsing of surface syntactic structure, recent additions to the theory (Frazier, 1999; Tunstall, 1998) have proposed that LF parsing is guided by the same principles. Structure-driven models predict that processing difficulty will be associated with more complex syntactic structures and with reanalysis of an already-constructed parse.

At the other end of the spectrum are parallel, constraint-based models of sentence comprehension (MacDonald, 1994; Spivey & Tanenhaus, 1998; Tanenhaus, Spivey-Knowlton, & Hanna, 2000; Trueswell, 1996; Trueswell & Tanenhaus, 1994). In these models, multiple sources of information, such as syntactic category information, contextual plausibility, and lexical co-occurrence frequencies, among others, contribute to the activation of the many possible analyses of a sentence. In such a model, processing is predicted to be relatively fast and easy when several sources of information support one parse. The model predicts processing difficulty when two alternatives are approximately equally activated, such that they compete with each other.

One well-known parallel model of sentence processing emphasizes the role of referential context in the choice between alternative analyses. In the model proposed by Steedman and his colleagues (Altmann & Steedman, 1988; Crain & Steedman, 1985), competing analyses are evaluated in parallel for plausibility with respect to the discourse context. The preferred analysis is determined according to the “principle of parsimony” quoted in (36):
A reading which carries fewer unsupported presuppositions will be favoured over one that carries more. (Altmann & Steedman, 1988, p. 203)

The effect of the principle is to favour the analysis that requires the smallest number of changes to the existing discourse context. In the absence of context, the surface-scope interpretation of (37) is more parsimonious, since it causes one individual (a climber) to be added to the discourse model, while the inverse-scope interpretation requires the accommodation of several climbers into the model.

(37) A climber scaled every cliff.

However, if (37) appears in a context that has already introduced multiple climbers into the discourse model, then the inverse-scope interpretation is “referentially supported” (Altmann & Steedman, 1988, p. 204) and is predicted to be no more costly to the processor than the surface-scope interpretation. In Chapter 2 I present evidence that assigning inverse scope to a doubly quantified sentence incurs a processing cost even in a context that reliably supports the inverse-scope interpretation.

I shall assume at the outset that the processor operates serially, driven by structural principles, for parsing both surface syntactic structure and LF structure. I return to evaluate this assumption in the light of the results of the experiments presented in Chapter 2.

Linguistic Structure and Conceptual Structure

On the QR account of scope ambiguity, the ambiguity is primarily structural – the two different interpretations arise from two different syntactic configurations at LF. On the other hand, if the truth-conditions of a sentence containing an indefinite and a universal quantifier are not strictly determined by the syntactic configuration, as is the case in Steedman’s most recent account of scope ambiguity (Steedman, 2003), it is intuitively plausible that perceivers could parse the surface syntax and then compute scope relations through conceptual, non-linguistic inferencing such as is described by Fodor (1982). In other words, the “specification” of an indefinite might be a conceptual operation rather than a linguistic one. Indeed, some psychologists argue that syntactic structure plays only a minor role in sentence comprehension. For example, van Berkum,
Brown & Hagoort (1999) describe a theory of sentence comprehension in which interpretive decisions are conditioned not only by a continually updated discourse model, but by shared experiences between the speaker and listener and by cultural conventions. They argue that “there is no functionally distinct separate stage during which ... compositionally constrained word meaning is exclusively evaluated with respect to ‘local sentence meaning,’ independent of the context in which that sentence occurs” (van Berkum et al., 1999, p. 663).

Studying the comprehension of scope-ambiguous sentences thus offers an opportunity to investigate the nature of semantic inferencing altogether. If the on-line comprehension of scope-ambiguous sentences involves processing costs independent of those introduced by contextual constraints, empirical evidence for such costs would support a model of sentence comprehension in which abstract linguistic structure plays a crucial role. Such evidence would also support the QR account of quantifier scope ambiguity, in which the difference between the two interpretations is syntactic in nature.

The experiments that I present in Chapter 2 show that perceivers experienced processing difficulty when they assigned the inverse-scope interpretation to a doubly quantified sentence, even when there was no competing alternative interpretation and the discourse context supported the inverse scope interpretation. I argue that these results support a serial model of sentence processing that relies on structural principles to construct LF representations of quantifier scope relations.

Sentence Processing and Linguistic Theory

Quite apart from what we can learn about sentence processing by conducting experiments studying on-line comprehension, the experiments that I present here are valuable test cases for the linguistic theories summarized above. The data on which these theories depend are, for the most part, introspective grammaticality judgments arising from the intuitions of just a small number of native speakers, often the linguists themselves who developed the theories. Experimental data provide independent, objective evidence gathered from a large number of native speakers to substantiate these intuitive judgments. The self-paced reading data gathered from naïve participants provide measures that are unlikely to be subject to perceivers'conscious control. Such data are especially valuable when the crucial grammaticality judgments are contested, as in the case of doubly-quantified sentences followed by ellipsis sentences which are the basis for the theory of Scope Economy (Fox, 2000). In fact, the results of the experiments that I
present in Chapter 3 show that perceivers often find these constructions acceptable, a fact which calls into doubt the empirical foundation of the theory of Scope Economy.

A Proposal for Understanding Scope Ambiguity in Real Time

In Chapter 2 I present several on-line and off-line experiments that investigate perceivers’ comprehension of scope-ambiguous sentences in real time. The results of these experiments show that assigning an inverse-scope interpretation to a doubly quantified sentence incurs a processing cost, even when that interpretation is supported by the discourse context. Based on these results I propose a principle of Processing Scope Economy (PSE) whereby the processor’s preference is to construct a parse with the fewest syntactic operations. The PSE principle serves as an alternative to Fox’s grammatical principle of Scope Economy, which prevents a scope-shifting operation from occurring unless it has some semantic effect. In Chapter 3 I show that the PSE principle offers a better account of the psycholinguistic facts than the grammatical Scope Economy principle does, while avoiding the theoretical complications of Fox’s principle.
CHAPTER 2:
COMPUTING SCOPE RELATIONS
Introduction

While theoretical accounts of the linguistic representation of quantifier scope abound, the most prominent of which I described in Chapter 1, there exists little psycholinguistic work investigating how quantifier scope relations are computed in real time. Indeed, because quantifier scope ambiguity is not due to observable differences in surface syntax, there is some doubt as to whether computing quantifier scope relations involves processing syntactic structure at all. Various scholars have argued that the observed scope preferences arise from properties of the particular quantifiers, from the grammatical or thematic roles of the quantified expressions, from the relative complexity of the mental models associated with the competing interpretations, or simply from the relative plausibility of the interpretations with respect to context and to real-world knowledge.

While many of these factors no doubt contribute to the on-line comprehension of quantified sentences, I present in Chapter 2 data from several experiments that combine to show that linguistic structure plays a critical role in processing scope-ambiguous sentences. My experiments investigate the on-line comprehension of two types of doubly quantified sentences, examples of which are given in (38) and (39).

(38) A climber scaled every cliff.

(39) Every player rubbed a lucky charm.

In (38), *a climber* has scope over the universal *every cliff* in the surface syntax. If the LF scoping is the same, the resulting interpretation is the surface-scope one where a single climber scaled all the cliffs. If *every cliff* takes scope over *a climber* at LF, the result is the inverse-scope interpretation where each cliff was scaled by some possibly different climber. In example (39), on the other hand, the interpretation where the universal *every player* has wide scope is the surface-scope interpretation, while in the inverse-scope interpretation a *lucky charm* has wide scope, such that there is a single charm that all the players rubbed. I refer to these two kinds of doubly quantified sentences according to their surface ordering of quantifiers, as *a...every and every...a* sentences, respectively.

The results of the experiments presented here show that assigning an inverse-scope interpretation to a doubly quantified sentence consumes more processing resources than
assigning surface scope even when extra-linguistic factors conspire to make the inverse-
scope interpretation the preferred one. Assigning inverse scope incurs a processing cost
when the discourse context supports the inverse-scope interpretation, when the inverse-
scope interpretation is referentially simpler than the surface-scope interpretation, and even
when no competing interpretations are available. I attribute this processing cost to the
greater linguistic complexity of the inverse-scope representation and conclude that
computing quantifier scope relations in real time crucially involves linguistic processing,
not just general conceptual inferencing.

**Background: Theories of Quantifier Scope Comprehension**

Most of the existing studies of quantifier scope comprehension reveal a preference
for the surface-scope interpretation of a doubly quantified sentence. Several different
principles have been proposed in the literature to account for this preference. I present
here a summary of the highlights of these proposals.

**Hierarchical Principles Interact to Determine Scope Preferences**

Two early studies that gathered interpretive judgments about scope preferences
(Ioup, 1975; VanLehn, 1978) proposed hierarchies that accounted for perceivers’ scope
preferences in terms of the intrinsic properties of quantifier-determiners. In both studies,
each occupied the highest position in the hierarchy, indicating its strong preference to have
wide scope over other elements in a sentence. Every occupied the next highest position and
other quantifiers came lower in the hierarchy. Both studies also acknowledged the
influence of the role of the quantified expression in the sentence in influencing scope
preferences. Ioup proposed a Grammatical Hierarchy whereby topics have the greatest
preference to take wide scope, followed by subjects and then objects. In a similar vein,
VanLehn proposed a C-command hierarchy to capture his observation that a quantified
phrase prefers to take scope over the elements that it C-commands at surface structure.
Other researchers have noted that these two hierarchies often overlap with and make the
same predictions as an animacy hierarchy (Catlin & Micham, 1975; Micham, Catlin,
VanDerveer, & Loveland, 1980) and as a principle of simple linear order (Johnson-Laird,
Byrne, & Tabossi, 1989, 1992).

If the inherent properties of the quantifier and the role or position of the
quantifier in the sentence both play a role in determining a sentence’s preferred scoping,
then the predictions diverge for the two kinds of sentences that I consider here. The
Grammatical and C-command hierarchies both predict a preference for surface scope in both *a...every* and *every...a* sentences. However, the tendency of *every* to take wide scope within the Quantifier Hierarchy predicts a preference for surface scope in *every...a* sentences but for inverse scope in *a...every* sentences. If both these hierarchies are at work, we should observe a strong preference for the surface-scope interpretation in *every...a* sentences, but more mixed preferences in *a...every* sentences. The results of my experiments show, rather, that perceivers prefer surface-scope interpretations for both kinds of sentences and that this preference is at least as strong and possibly somewhat stronger for *a...every* sentences than for *every...a* sentences.

**Non-Linguistic Information Determines Scope Preferences**

Fodor (1982) argued that scope relations are represented in the mind not in linguistic form, but rather as models of the world that can be schematized with simple diagrams. A sentence like (40), then, has only one syntactic representation, that is, its surface structure, but has two potential conceptual representations corresponding to the two models shown in (41) and (42).

(40) A climber scaled every cliff.

(41) climber
    cliff
    cliff
    cliff
(42) climber
    cliff
    cliff
    cliff
    cliff
    cliff

Notice that the two distinct conceptual representations do not map directly to the truth-conditions of the two interpretations. The model shown in (42) necessarily maps to the
inverse-scope interpretation, but the truth conditions of the inverse-scope interpretation are such that there need not be multiple distinct climbers as long as each cliff is scaled by someone, possibly the same climber for all the cliffs. So model (41) could also serve as a conceptual representation for the inverse-scope interpretation of sentence (40). The surface-scope interpretation of (40), on the other hand, where a single climber scales all the cliffs, corresponds only to model (41). For Fodor (1982), the conceptual representation that a perceiver develops for a given sentence is of greater interest than the logical entailments of linguistic semantic representation. In other words, the fact that the surface-scope interpretation of (40) entails an interpretation where each cliff is scaled is less interesting than the fact that a distinct conceptual representation containing multiple climbers is possible.

Since interpretation occurs incrementally in Fodor’s system, the linear order of quantifiers in the surface syntax plays a role in determining scope preferences. It is therefore relatively easy to interpret every with wide scope when it is encountered early in the sentence, since everything that follows every can be interpreted in its scope. On the other hand, giving every wide scope in an a...every sentence like (40) should be relatively more difficult and thus dispreferred, since a climber is initially “assigned a singular representation, which must then be revised when the universal quantifier is encountered” (Fodor, 1982, p. 148). This non-linguistic system of representation thus predicts a preference for the surface-scope interpretation in both the kinds of sentences that I am concerned with, although the surface-scope preference for a...every sentences should be stronger. In other words, the inverse-scope interpretation should be more strongly disfavoured in a...every sentences than in every...a sentences. On the other hand, if the discourse context leads the perceiver to construct a mental model that includes multiple climbers, then presumably there is little reason to initially assign a singular representation for the existential quantifier. It should therefore be easier to assign the inverse-scope interpretation of (38) following a context that supports it, since the mental model of multiple climbers already exists.

Fodor’s account of the mental representation of quantifiers is similar in theme to the later models of sentence comprehension proposed by Steedman and his colleagues (Altmann & Steedman, 1988; Crain & Steedman, 1985). According to their principle of parsimony, the preferred interpretation for an ambiguous sentence (not necessarily a quantified sentence) is the interpretation that carries the fewest unsupported presuppositions, that is, the interpretation that requires the fewest changes to the discourse model. If a sentence like (40) occurs in isolation, the principle leads the processor to prefer the surface-scope interpretation, since it requires accommodating only a single climber into
the discourse model, as opposed to the multiple climbers introduced by the inverse-scope interpretation. But if the preceding context has already introduced multiple climbers, then the inverse-scope interpretation of (40) should be no less parsimonious than the surface-scope interpretation. The model thus predicts little difficulty in assigning inverse scope if this interpretation is supported by the discourse context. I will show in Experiments 5 and 7 that a favourable discourse context does not mitigate the cost of assigning inverse scope to an a...every sentence, contrary to the predictions of the principle of parsimony and of Fodor’s theory.

**Multiple Constraints Interact to Determine Scope Preferences**

The early psycholinguistic studies of interpretive preferences for scope-ambiguous sentences drew conclusions based on perceivers’ introspective grammaticality judgments. In Kurtzman and MacDonald’s (1993) study, participants were timed while they read pairs of sentences such as those in (43)–(44) and judged whether the second sentence was an acceptable continuation for the first.

(43) A kid climbed every tree.
The kid was full of energy.

(44) A kid climbed every tree.
The kids were full of energy.

The singular continuation in (43) corresponds to the surface-scope interpretation of the doubly quantified sentence and the plural continuation in (44) corresponds to the inverse-scope interpretation. The experiment included every...a sentences as well as a...every sentences. While this task still relies on conscious judgments of linguistic intuition, the time constraint of the task should allow for a more accurate measure of perceivers’ real-time comprehension of quantified sentences than introspective judgments do.

Kurtzman and MacDonald’s study tested several of the principles described above, such as the Grammatical Hierarchy and the C-command Hierarchy, as well as other related structural principles, in order to evaluate which principles the parser uses in making interpretive decisions about doubly quantified sentences. While some of their results showed the same preference for surface-scope interpretations observed in previous studies, the overall pattern of data led the authors to argue for an account in which multiple principles interact to influence scope preferences. The two primary principles that
Kurtzman and MacDonald identify are the Thematic Hierarchy Principle and the Single Reference Principle. According to the Thematic Hierarchy Principle, agentive phrases have the greatest preference to take wide scope, followed by experiencers and then patients, while the Single Reference Principle captures the intuition of Fodor (1982) described above that an existential \( a \) will be interpreted as having a single referent if it is encountered before every in a sentence.

In Kurtzman and MacDonald’s model, if the two (or more) principles support the same scope relation, then there will be a clear and strong preference for that interpretation. If the principles conflict, however, both possible interpretations are activated and there may be only a weak preference, if any, for one or the other interpretation. In this parallel-processing model, processing difficulty is predicted when the two interpretations are approximately equally activated such that they compete with each other. In Experiment 7, however, I show that the cost of assigning inverse scope is incurred even for an unambiguous inverse-scope sentence, that is, even when there is no alternative interpretation competing with the inverse-scope representation. The result suggests that competition between alternative interpretations is not the only factor in the comprehension of doubly quantified sentences.

**LF Structure Determines Scope Preferences**

The remaining accounts of the real-time comprehension of quantifier scope ambiguity are those that emphasize the role of linguistic structure. At least two theoretical proposals (Pritchett & Whitman, 1995; Reinhart, 1997) predict that surface-scope interpretations are preferred because they have a more economical LF structure, hypothesizing that sentences with more complex LFs consume more processing resources. Tunstall (1998) formalized this notion in processing terms with her Principle of Scope Interpretation (PSI), given below in (45).

(45) **Principle of Scope Interpretation** (Tunstall, 1998, p. 56)
The default relative scoping in a multiply quantified sentence is computed from the required LF-structure of that sentence, where the required LF is determined by required grammatical operations acting on the S-structure. The default scoping is the preferred scoping unless there is evidence to go beyond it.
The PSI, which follows naturally from general principles of processing economy, makes explicit the processor’s preference for surface-scope configurations in the absence of compelling evidence for inverse scope.

I propose a principle that is closely related to the PSI: the principle of Processing Scope Economy.

(46) Processing Scope Economy
The human sentence processing mechanism prefers to compute a scope configuration with the simplest syntactic representation (or derivation). Computing a more complex configuration is possible but incurs a processing cost.

The two principles agree that, all else being equal, the processor prefers surface-scope parses. The PSI, however, makes no predictions about the cost associated with computing a dispreferred scope configuration, and even hints that the surface-scope preference might be reversed by the right kind of evidence. The PSE principle, in contrast, is explicit in claiming that the surface-scope preference can be overridden, but that doing so incurs a processing cost even when the evidence of plausibility and context support the inverse-scope configuration.

Chapter Outline

In this chapter I present three questionnaire experiments and five self-paced reading experiments, the results of which show a cost for assigning inverse scope to a doubly quantified sentence not only when it appears in isolation, but also when the discourse context supports the inverse-scope interpretation, when the inverse-scope interpretation has the referentially simpler representation, and when there is no competing representation. I argue that these results support the principle of Processing Scope Economy, and that the cost of inverse scope is a structural cost associated with constructing a relatively complex LF. I conclude that interpreting doubly quantified sentences necessarily involves processing linguistic structure.
Experiments

Off-line Questionnaires

Three questionnaires measured perceivers’ off-line interpretive preferences for ambiguous doubly quantified sentences. Experiment 1 presented the a...every sentences in isolation, while Experiment 2 presented them following a paragraph that supported either the surface-scope or the inverse-scope interpretation. Experiment 3 investigated the every...a sentences embedded in contexts.

Experiment 1

Experiment 1 was a questionnaire designed to measure a baseline for perceivers’ preferred interpretations of ambiguous doubly quantified sentences in the absence of contextual information.

Experiment 1 Method

Stimuli

The stimuli consisted of 24 ambiguous doubly quantified sentences, each of which was followed by two possible paraphrases of the sentence. All of the sentences had the form illustrated in (47):

(47) A cashier greeted every customer.

a. One cashier greeted customers.

b. Several cashiers greeted customers.

For each sentence, the subject was a singular, indefinite DP (e.g., a cashier, a climber), the verb was an active, eventive verb in the simple past tense (e.g., greeted, climbed), and the object was a DP universally quantified with the determiner every (e.g., every customer, every cliff). A complete list of the stimuli is provided in the Appendix. The 24 stimulus items were included in a list with 96 filler items that had different kinds of ambiguities. Each
filler item consisted of an ambiguous sentence followed by two potential paraphrases. The filler items were ambiguous because of various ellipsis constructions, as illustrated in (48)–(51) or because of the interaction of numeral quantifiers with temporal adverbials (52).

(48) At the fundraiser, Valerie thanked the volunteers, and at the party, Lillian.

(49) Melissa assumed the secretary got a raise and Ernie did too.

(50) André surprised Jeffrey during the debate more frequently than Philippe.

(51) Solomon offended Rick on Tuesday, not Samuel.

(52) Four nurses waiting behind the desk were later drinking coffee.

The order of the items within the list was randomized. The order of presentation of the two possible responses was counterbalanced across two different versions of the list.

Procedure

Participants were instructed to circle the paraphrase that corresponded to their initial interpretation of the sentence. The participants were given as much time as needed to complete the questionnaire, but were asked to complete the questionnaire without deliberating very long over the answers. They typically took no more than twenty minutes to complete the questionnaire.

Participants

Thirty-eight Northwestern undergraduates participated in the experiment for course credit. All participants were native speakers of English.
Experiment 1 Results

Perceivers chose the paraphrase corresponding to the surface-scope interpretation in 81% of cases, a value that is significantly different from the chance value of 50% (z-test, p<0.001). Twelve participants chose the surface-scope paraphrase for all of the items, and one item (A student solved every puzzle.) earned a surface-scope response from all participants. No participant had a surface-scope response rate of less than 46%, and no item had a surface-scope response rate of less than 50%.

Experiment 1 Discussion

The results of Experiment 1 show that when a doubly quantified a...every sentence is ambiguous there is nonetheless a default preference for the surface-scope interpretation of that sentence. This result is consistent with those reported in the literature, as well as most of the proposed principles governing quantifier scope comprehension. For example, in the LF representation of the surface-scope interpretation, the phrase that takes wide scope (a cashier in sentence (47)) is the subject and the agent of the sentence, and also c-commands the other quantified phrase. The surface-scope interpretation also has a simpler LF than the inverse-scope interpretation.

While Experiment 1 shows a clear preference for surface scope when the doubly quantified sentence is presented in isolation, the results also make it clear, however, that the inverse-scope interpretation is not completely inaccessible to perceivers (at least in a forced-choice task), since participants chose this interpretation on 19% of trials. In short, the so-called ambiguous quantified sentences are indeed ambiguous, but the preferred interpretation is the surface-scope interpretation.

Experiment 2

In Experiment 2 the doubly quantified sentences from Experiment 1 were embedded in discourse contexts that supported either the surface-scope interpretation or the inverse-scope interpretation. This off-line questionnaire established a baseline for the effects of the contexts on perceivers’ interpretive preferences, which is valuable because these contexts are used later in on-line comprehension experiments.
Experiment 2 Method

Stimuli

The 24 ambiguous doubly quantified *a...every* sentences were embedded in contexts about six sentences long. The contexts were designed to support either the surface-scope or the inverse-scope interpretation of the quantified sentence. Each context described a scenario where a group of people took part in some activity. The inverse-scope context differed from the surface-scope context by describing the scenario in such a way that the activity was likely to be undertaken by multiple members of the group, whereas the surface-scope context made it more likely that a single member of the group would take part in the activity. For example, a sample surface-scope context is shown in (53):

(53) The members of the gourmet club decided to put out a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. The president of the club requested that a volunteer test the recipes to make sure that the instructions were correct. After a short discussion, a member of the club tested every recipe.

In example (53), the club’s president asks for a single volunteer to test recipes, making it likely that the doubly quantified sentence, *a member tested every recipe*, describes a scenario with only a single person doing the testing. The scenario is set up differently in the inverse-scope context shown in (54):

(54) The members of the gourmet club decided to put out a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. A member of the club tested every recipe.

In the inverse-scope scenario, all the people who nominate recipes are supposed to test them. The doubly quantified sentence, *a member tested every recipe*, therefore quite plausibly describes a scenario where multiple people test recipes.
Notice, however, that both contexts introduce multiple individuals (i.e., “the members of the gourmet club decided...”) before the ambiguous doubly quantified sentence. According to the principle of parsimony (Altmann & Steedman, 1988; Crain & Steedman, 1985), therefore, an interpretation with multiple recipe-testers should be no less parsimonious than the interpretation with a single tester, since the multiple participants have already been introduced in the context.

The ambiguous doubly quantified sentence appeared in two conditions: surface-scope biased and inverse-scope biased. There were also two baseline conditions, one where the quantified sentence was unambiguous surface scope in the surface-scope context, and one where the quantified sentence was unambiguous inverse scope in the inverse-scope context. The experiment thus had a 2x2 design, with two different contexts and two levels of ambiguity. The 24 items were arranged in four different lists in a Latin Square design, such that each version of the list contained six items in each of four conditions. Each subject saw only one version of the list, and therefore saw each stimulus item in only one of the four conditions.

An example of the entire paradigm for one stimulus item is given in (55)–(58). The full set of stimuli is provided in the Appendix.

(55) Surface scope, unambiguous
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination. One weekend, the climbing equipment shop sponsored a show to demonstrate the sport. While an announcer described the techniques, the climbing expert scaled every cliff.

(56) Surface scope biased, ambiguous
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination. One weekend, the climbing equipment shop sponsored a show to demonstrate the sport. While an announcer described the techniques, an experienced climber scaled every cliff.
Inverse scope, unambiguous
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination. One weekend, the climbing equipment shop sponsored a race between climbing enthusiasts. While an official timed the event, a different climber scaled every cliff.

Inverse scope biased, ambiguous
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination. One weekend, the climbing equipment shop sponsored a race between climbing enthusiasts. While an official timed the event, an experienced climber scaled every cliff.

Each paragraph was followed by a comprehension question and two possible answers, as shown in example (59).

Once a month, the city tested its emergency alert systems. The regular schedule for the air-raid sirens was to test them on the first of the month. At exactly ten in the morning on that day, an employee sounded every siren.

How many employees sounded sirens?

A. One.  
B. Several.

The order of presentation of the possible answers was counterbalanced across the items. I took the answer One to indicate that the perceiver had assigned a surface-scope interpretation to the quantified sentence, while the answer Several indicated an inverse-scope interpretation. This strategy provides a conservative measure of the true number of inverse-scope interpretations assigned, since the strict interpretation of the inverse-scope configuration (in (59), roughly, for every siren, some employee sounded it) is consistent with a scenario where every siren is sounded by the same employee or the scenario where several employees sound sirens. While it is possible, then, that some portion of the One responses represent cases where perceivers had assigned an inverse-scope interpretation, the Several responses necessarily represent inverse-scope interpretations. Since Experiment 1 showed that the surface-scope interpretation is the default interpretation, it is the number of items
assigned an inverse-scope interpretation that will be of most interest in the present set of experiments. Using this conservative measure averts the danger of overestimating the number of inverse-scope interpretations.

In the unambiguous surface-scope condition, the comprehension question was not a how many question, but rather a question asking for information about the story told in the paragraph, as shown in example (60). One of the two answer choices was correct and the other was incorrect.

(60) Once a month, the city tested its emergency alert systems. The regular schedule for the air-raid sirens was to test them on the first of the month. At exactly ten in the morning on that day, the safety officer sounded every siren.

Who sounded sirens?

A. The mayor. B. The safety officer.

In this condition, then, the question response did not distinguish between two possible interpretations, but simply indicated whether perceivers had understood the paragraph.

The items were arranged in pseudo-random order, such that no two items of the same condition appeared adjacent to each other. Included in the questionnaire were 24 filler items from a different experiment. The filler items consisted of short paragraphs that included a sentence with a comparative construction like the example given in (61).

(61) Pauline assisted Wanda more efficiently than Lisa this past week on the special issue of the paper.

The comparative sentence was ambiguous between the subject and object interpretation of the unelided remnant (Lisa in (61)).

Procedure

The stimuli were printed in paper questionnaires several pages long. Participants were instructed to circle the answer that matched their first interpretation of the sentence. Participants had as much time as they needed to complete the questionnaire, but were
asked to complete it without deliberating very long over the answers. They typically took no more than twenty minutes to complete the questionnaire.

Participants

Twenty-seven participants completed the questionnaire, 17 women and ten men. Eleven participants were Northwestern University undergraduates and 16 were adult volunteers. The participants ranged in age from 19 to 70 years. The undergraduates received course credit for their participation. Data from three subjects were excluded from the analysis: one who was not a native speaker of English, one who had already participated in a related experiment, and one who did not complete the entire questionnaire.

Experiment 2 Results

As shown below (Table 1), the participants nearly always chose the expected response in the two unambiguous conditions. In the ambiguous conditions, subjects chose the surface-scope interpretation for 81% of items in the surface-scope context, but chose the inverse-scope interpretation for only 53% of items in the inverse-scope context.
Table 1.

Experiment 2 Responses to Comprehension Question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response</th>
<th>Inverse-scope response</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.81</td>
<td>0.19</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.04</td>
<td>0.96</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.47</td>
<td>0.53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correct response</th>
<th>Incorrect response</th>
</tr>
</thead>
<tbody>
<tr>
<td>unambiguous surface scope</td>
<td>0.99 0.01</td>
</tr>
</tbody>
</table>

Comparing the two ambiguous conditions, there was a significant effect of context on the response chosen ($t_{1}(46)=5.8, p<0.0001; t_{2}(46)=4.1, p=0.0002$). Perceivers chose the surface-scope response *One* more frequently following the surface-scope context than following the inverse-scope context. Comparing the two conditions with the inverse-scope context, there was a significant effect of the ambiguity of the quantified sentence ($t_{1}(46)=9.4, p<0.0001; t_{2}(46)=6.4, p<0.0001$). Perceivers chose the inverse-scope response *Several* more frequently when the quantified sentence was unambiguous inverse scope than when it was ambiguous.
Experiment 2 Discussion

It is clear from these results that the unambiguous conditions are genuinely unambiguous in the expected way, that is, that perceivers chose the correct interpretation in unambiguous surface-scope conditions and the inverse-scope interpretation in unambiguous inverse-scope conditions. This fact confirms that the subjects responded to the questionnaire as expected and not randomly.

The more interesting result is the response to the ambiguous conditions. In the surface-scope supporting condition, 81% of items were assigned a surface-scope interpretation. Interestingly, this is the same proportion of surface-scope interpretations as in Experiment 1, where the quantified sentences appeared in isolation. Apparently the surface-scope context did not make a surface-scope interpretation any more likely than the default preference. Even in this context, the quantified sentence was ambiguous, since 19% of the surface-scope biased items were assigned an inverse-scope interpretation. In other words, the inverse-scope interpretation was disfavoured but not impossible in the surface-scope biasing conditions.

In the inverse-scope supporting context, only 53% of items were assigned an inverse-scope interpretation. It appears that, while the context played a role in the interpretive decision (since the proportion of responses is significantly different from in the surface-scope context), there was still a strong default preference for the surface-scope interpretation, since it was chosen in 47% of trials with an inverse-scope context. Notwithstanding this default interpretive preference, comparing these results with the results of Experiment 1 makes it clear that the inverse-scope biasing context did affect the interpretation of the ambiguous quantified sentence, since in the absence of context in Experiment 1 the inverse-scope interpretation was assigned only 19% of the time. In short, the inverse-scope context boosted the rate of selecting the inverse-scope interpretation from 19% to 53%.

Overall, the results of Experiment 2 show that the unambiguous sentences were indeed unambiguous, that the contexts influenced the interpretation of the ambiguous sentences, but also that the ambiguous sentences retained some degree of ambiguity in these different contexts.
Experiment 3

In Experiment 3 an off-line questionnaire investigated perceivers’ interpretive preferences for the every...a sentences embedded in contexts. This questionnaire gathered ratings using a five-point scale rather than a forced-choice task, in an attempt to reduce any bias potentially introduced by the forced-choice task. For example, perceivers might settle on an inverse-scope interpretation very rarely in a natural setting, yet be reluctant to choose the surface-scope interpretation repeatedly in an experimental setting. If this were the case, the number of inverse-scope responses might be artificially inflated by a forced-choice task. The gradient scale task was intended to reduce this risk.

Experiment 3 Method

Stimuli

The experiment stimuli consisted of ten every...a sentences, each of which was embedded in a paragraph designed to support either the surface-scope or the inverse-scope interpretation. Examples are shown in (62) and (63).

(62) surface-scope context:
Ball players can be quite superstitious, particularly during a winning streak. Some players believe that their habits are more effective than their teammates' quirks. Before the Marauders' final game, every player rubbed a lucky charm.

(63) inverse-scope context:
Ball players can be quite superstitious, particularly during a winning streak. If a player doesn't participate in the team ritual, his teammates get very angry. Before the Marauders' final game, every player rubbed a lucky charm.

Following each paragraph was a five-point scale with two unambiguous paraphrases of the doubly quantified sentence, one at either end, as shown in (64).
(64)  All the players rubbed the same lucky charm.  1 2 3 4 5
Each player rubbed a different lucky charm.

The paraphrase at the low end of the scale corresponded to the inverse-scope interpretation of the every...a sentence ([ ] > [ ]), while the paraphrase at the high end corresponded to the surface-scope interpretation ([ ] > [ ]).

The items were arranged in a list along with the 24 a...every items that were used in Experiment 2. (The data from the a...every items are not reported here.) The items were arranged in random order in the list. Two versions of the list were prepared in a split-half design. The order of presentation of the items in one list was reversed from the order in the other.

Procedure

The stimuli were printed in paper questionnaires several pages long. Participants were instructed to rate which of the two interpretations of the quantified sentence was more likely, and circle the point on the scale that corresponded to their rating. Participants had as much time as they needed to complete the questionnaire, but were asked to complete it without deliberating very long over the answers. They typically took no more than twenty minutes to complete the questionnaire.

Participants

Twenty-three Northwestern University undergraduates participated in Experiment 3 for course credit, six men and 17 women. Data from five participants were excluded from the analysis: three who had already participated in a related experiment and two who were not native speakers of English.

Experiment 3 Results

The mean rating was calculated for each item, where a high rating indicates that the surface-scope interpretation was more likely and a low rating that the inverse-scope interpretation was more likely. For items in the surface-scope context, the mean rating was 4.3. For items in the inverse-scope context, the mean rating was 3.3. The difference in
mean rating between the two contexts was significant ($t_{1(17)}=4.1$, $p<0.001$; $t_{2(9)}=2.4$, $p=0.04$).

**Experiment 3 Discussion**

It is clear that the two different contexts influenced perceivers’ interpretations of the ambiguous *every...a* sentences, since the mean ratings were significantly different between the two contexts. However, in the inverse-scope context the mean rating of 3.3 was above the half-way point on the five-point scale. Although the inverse-scope context led perceivers to rate the inverse-scope paraphrase as more likely than in the surface-scope context, nevertheless the mean rating was closer to the surface-scope interpretation than to the inverse-scope interpretation. In other words, perceivers found the surface-scope paraphrase more likely than the inverse-scope paraphrase even in the inverse-scope context. Clearly the inverse-scope interpretation was quite strongly dispreferred for the *every...a* sentences, although the different method used in this experiment makes it impossible to compare these results directly with those of Experiments 1 and 2.

**Self-Paced Reading Experiments**

Five self-paced reading experiments investigated the real-time comprehension of doubly quantified *a...every* and *every...a* sentences in isolation and in context. Experiments 4 and 6 presented the quantified sentences in isolation and Experiments 5, 7 and 8 presented them in biasing contexts.

**Experiment 4**

Experiment 4 presented *a...every* sentences and *every...a* sentences in a word-by-word self-paced reading task. Each sentence was followed by a singular or plural sentence that disambiguated to either the surface-scope or inverse-scope interpretation, as in Kurtzman and MacDonald (1993).
Experiment 4 Method

Stimuli

Experiment 4 had two subparts. In part A, the stimuli consisted of the 24 ambiguous \textit{a...every} sentences. The stimuli in part B were the 10 \textit{every...a} sentences. In both parts, each sentence was followed by a sentence that disambiguated to either surface scope or inverse scope. The stimuli were divided into presentation regions of a single word or phrase. In examples (65)–(68), the ‘|’ symbol demarcates the presentation regions.

\begin{itemize}
  \item [(65)] An | experienced | climber | scaled | every | cliff.
  \hspace{1cm} The climber | was | very skilled.
  \item [(66)] An | experienced | climber | scaled | every | cliff.
  \hspace{1cm} The climbers | were | very skilled.
  \item [(67)] Every | historian | examined | a | document.
  \hspace{1cm} The document | was | a stirring speech.
  \item [(68)] Every | historian | examined | a | document.
  \hspace{1cm} The documents | were | in good condition.
\end{itemize}

Every item was followed by a \textit{how-many} question with two possible answers that corresponded to the surface-scope and inverse-scope interpretations of the quantified sentence. For example, items (65)–(66) were followed by question (69), and items (67)–(68) by question (70).

\begin{itemize}
  \item [(69)] How many climbers scaled cliffs?
  \hspace{1cm} One. \hspace{2cm} Several.
  \item [(70)] How many documents were examined?
  \hspace{1cm} One. \hspace{2cm} Several.
\end{itemize}
For an *a...every* item like (65)–(66), the answer *One* to question (69) corresponds to the surface-scope interpretation, where a single climber scales all the cliffs, while the answer *Several* corresponds to the inverse-scope interpretation where each cliff is scaled by a different climber. In contrast, for the *every...a* items like (67)–(68), the answer *One* corresponds to the inverse-scope interpretation, where a single document is examined, and the answer *Several* to the surface-scope interpretation where each historian examines a different document.

The stimuli were arranged in a list that included 66 ambiguous items of similar structure that were used in other experiments. Sentence (71) is an example of one of the filler items.

(71) Michael misled the banker about the remaining funds and Arnold about the incoming checks.

Sentence (71) is ambiguous as to the status of *Arnold*, which could be the subject of an elided element “Arnold misled the banker about the incoming checks” or the object of an elided element “Michael misled *Arnold* about the incoming checks”. Half of the filler items were followed by comprehension questions. The items in the list were arranged in pseudo-random order. Four versions of the list were created the items in a different order in each version.

*Procedure*

Participants read the stimulus items on a computer screen in a self-paced reading task. Before the actual experiment began, participants read ten practice items. Before each item appeared on the computer screen, a set of dashes showed the position of the item on the screen. Participants pressed the space bar on the computer’s keyboard to make each region of text appear. As each region of text was displayed, the previous region reverted to dashes. Participants were instructed to press the space bar at a rate that allowed them to read at their normal speed. The computer recorded the length of time that each region of text was displayed on the screen.

After each item, a comprehension question was presented along with two possible answers, as illustrated above in (69) and (70). The order of presentation of the two possible answers was counterbalanced across the items and conditions. The answer keys ‘D’ and ‘K’ on the computer’s keyboard were marked with coloured tape. Participants used
the ‘D’ key to select the answer displayed on the left side of the screen and the ‘K’ key to select the right-hand answer. The computer recorded the answer selected and the response latency. Participants were not given feedback about their responses.

The computer prompted the participants to take a break three times during the experiment, at the one-quarter, one-half and three-quarter points. At these breaks they could rest as long as they chose, after which they started the experiment again by pressing a key. Participants were also allowed to rest between trials if they chose. The entire experiment, including instructions, consent, the practice session, and breaks, took about 35 minutes to complete.

Participants

Forty-three Northwestern University undergraduates participated in the experiment, 18 men and 25 women. The data from 14 participants were excluded from the analysis: five who were not native speakers of English, eight who had already participated in a related experiment, and one who talked to the experimenter while participating.

Experiment 4 Results

The results reported here are residual reading times, calculated using a deviation from regression algorithm (Ferreira & Clifton, 1986). Reading times for each participant were regressed on the length in characters of each segment that the participant read. The reading time residuals provide individually corrected estimates of participants’ expected reading times for the regions of differing lengths. Thus, for a given region, zero represents the expected reading time for a segment of that length in characters, averaged across participants. Negative numbers represent faster reading times than expected and positive numbers represent slower reading times than expected for a region of that length.
Sub-Experiment 4A

I report first the results for the *a...every* items. The table below (Table 2) gives the mean residual reading times for each word in the quantified sentence and Figure 1\(^3\) shows the reading times for the disambiguating sentence, averaged over subjects and items.

**Table 2.**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>adjective</th>
<th>subject noun</th>
<th>verb</th>
<th>every</th>
<th>object noun</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-6</td>
<td>-128</td>
<td>-68</td>
<td>-54</td>
<td>-44</td>
<td>156</td>
</tr>
</tbody>
</table>

Although Figure 1 shows that the plural disambiguating sentence was read more slowly than the singular, there is no evidence that this difference is significant in the first two regions (all $F$’s$<1$). At the final region, the difference between singular and plural approached significance only in the items analysis ($t_{1(58)}=1.4$, $p=0.17$, $t_{2(44)}=1.9$, $p=0.06$). However, when the residuals were calculated over the entire sentence instead of for each region, the difference in reading time between the singular and plural disambiguating sentence was significant ($t_{1(58)}=2.9$, $p=0.005$; $t_{2(44)}=2.5$, $p=0.02$).

The following table (Table 3) gives the proportions of responses to the comprehension question following the *a...every* sentences.

---

3 In this and in every figure, the error bars indicate the standard error of the mean.
Table 3.

Experiment 4A Proportion of responses to comprehension question following a...every sentence.

<table>
<thead>
<tr>
<th></th>
<th>Surface scope response One</th>
<th>Inverse scope response Several</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular disambiguating sentence</td>
<td>0.87</td>
<td>0.13</td>
</tr>
<tr>
<td>plural disambiguating sentence</td>
<td>0.59</td>
<td>0.41</td>
</tr>
</tbody>
</table>
Figure 1. Mean residual reading time for disambiguating sentence, Experiment 4A.
The disambiguating sentence had a significant effect on the responses to the comprehension question: participants chose the surface-scope response *One* after 87% of trials with a singular disambiguating sentence, but on only 59% of trials with a plural disambiguating sentence. The difference between the two conditions is statistically significant ($t(1)(58)=4.3$, $p<0.001$; $t(2)(44)=8$, $p<0.001$). There is limited evidence that the proportion of responses following the plural disambiguating sentence is different from a chance rate of 50% ($t(2)(9)=1.6$, *one-tailed* $p=0.06$).

While the plural sentence, which was expected to disambiguate to the inverse-scope interpretation, did increase the proportion of inverse-scope responses from 13% to 41%, perceivers still chose the inverse-scope response on fewer than half of the trials. Clearly the plural sentence did not disambiguate the quantified sentence entirely.

**Sub-Experiment 4B**

The following table (Table 4) provides the residual reading times for the quantified sentences in the *every...a* items. The reading times for the disambiguating sentence are depicted in Figure 2.

| Table 4. |
|-----------------|-----------------|-----------------|-----------------|
| *Every*         | *subject*       | *verb*          | *a*             | *object*        |
| -101            | -99             | -22             | 33              | 75              |

Although Figure 2 shows that the singular sentence that disambiguated to inverse scope was read more slowly than the plural, surface-scope disambiguating sentence by the end of the sentence, this difference was not significant at any region. At the first region, the reading times were virtually identical for the singular and plural subject NPs ($F's<1$, n.s.). At the second region, the plural verb was read more slowly than the singular...
(F1(1,58)=2.4, p=0.13; F2(1,18)=1.4, p=0.25). At the final region, the singular sentence was read more slowly than the plural (F1(1,58)=1.5, p=0.22; F2(1,18)=2.4, p=0.14). When the residuals were calculated over the entire sentence instead of for each region, there was no evidence for a difference in reading time between the singular and plural disambiguating sentences (t1(58)=0.79, p=0.43; t2(18)=0.86, p=0.86).
Figure 2. Mean residual reading times for disambiguating sentence, Experiment 4B.
The following table (Table 5) gives the proportion of responses to the comprehension question in part 4B of the experiment.

Table 5.

Experiment 4B Proportion of responses to comprehension question following every...a sentence.

<table>
<thead>
<tr>
<th></th>
<th>Surface scope response Several</th>
<th>Inverse scope response One</th>
</tr>
</thead>
<tbody>
<tr>
<td>plural disambiguating sentence</td>
<td>0.91</td>
<td>0.09</td>
</tr>
<tr>
<td>singular disambiguating sentence</td>
<td>0.18</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Notice that, for the every...a sentences, the difference in the proportions of responses between the singular and plural disambiguating sentences was more pronounced than for the a...every sentences. Perceivers chose the surface-scope response Several following the plural disambiguating sentence on 91% of trials, significantly more frequently than following the singular disambiguating sentence, where the proportion of surface-scope responses was only 18% (t1(58)=15, p<0.0001; t2(18)=16, p<0.0001). Apparently the singular sentence was more effective at disambiguating an every...a sentence to inverse scope than the plural sentence was at disambiguating an a...every sentence to inverse scope.

**Experiment 4 Discussion**

Sub-Experiment 4A showed that a plural sentence that disambiguated to inverse scope was read more slowly than a singular, surface-scope disambiguating sentence following an ambiguous a...every sentence. This fact provides some tentative support for
the notion that assigning inverse scope incurs a processing cost, which is unsurprising since
the inverse-scope interpretation is dispreferred by nearly all of the principles described
above (e.g., the Grammatical, C-command, and Thematic Hierarchies), has a conceptual
representation with greater referential complexity, and an LF representation with greater
structural complexity. It is not consistent, however, with the Quantifier Hierarchy which
predicts that every has a strong tendency to take scope over other quantifiers. On the other
hand, the slower reading times at the plural sentence could simply indicate perceivers’
surprise at encountering an unexpected plural if they had committed to the surface-scope
interpretation of the quantified sentence, which is entirely likely in light of the results of
the questionnaire in Experiment 1. In other words, the cost associated with the plural
sentence does not necessarily indicate that the inverse-scope interpretation itself is costly,
but simply that unexpected information incurs a cost. The every…a items in Sub-
Experiment 4B also showed slower reading times for the inverse-scope disambiguating
sentence, in this case the singular sentence, but the difference was not significant.

The more interesting result is the pattern of question responses. For the a…every
sentences, perceivers chose the surface-scope response following 87% of surface-scope
disambiguating sentences and also following 59% of inverse-scope disambiguating
sentences. In other words, the plural disambiguating sentence did not lead them to select
the inverse-scope interpretation very often. In contrast, for the every…a sentences, the
comprehension question responses were more consistent with the disambiguating
sentences. Perceivers chose the surface-scope response following 91% of plural, surface-
scope disambiguating sentences and chose the inverse-scope response following 81% of
singular, inverse-scope disambiguating sentences.

The results suggest that the preference for surface scope is stronger in a…every
sentences than in every…a sentences, since the reading times for the inverse-scope
disambiguating sentence are significantly slower only for the a…every items, and since the
proportion of inverse-scope responses is low for the a…every items. This pattern of results
might be an artifact of the truth conditions of the two kinds of sentences. In an a…every
sentence like (72), the inverse-scope configuration where every cliff takes scope over a climber
does not entail the existence of multiple climbers.

(72) A climber scaled every cliff.

As long as every cliff was scaled by some climber, possibly the same climber for all cliffs,
then the truth conditions of the inverse-scope interpretation are satisfied. It is possible
that a perceiver assigns the inverse-scope interpretation to the quantified sentence but
leaves the number of climbers in the discourse model unspecified (or at the simplest option, namely, one climber). The plural disambiguating sentence then causes processing difficulty not because it compels the parser to adopt the inverse-scope interpretation, but because it presupposes the existence of multiple climbers, which must then be accommodated in the discourse model.

For an every...a sentence like (73), the situation is the reverse.

(73) Every player rubbed a lucky charm.

Both the singular and the plural continuations are consistent with the surface-scope configuration where every player takes scope over a lucky charm. As long as every player rubs some lucky charm, the sentence is true whether they each rub a different one or they all rub the same one. It is possible, then, that the singular sentence did not actually disambiguate to an inverse-scope representation, but simply to the single-charm scenario consistent with the surface-scope representation. If this is the case, then the lack of a difference in reading times between the singular and plural “disambiguating” sentences is not very surprising.

What is most surprising in these results is the high proportion of singular responses to the comprehension question following the plural continuation to the a...every sentences. While the singular continuation can be consistent with either a surface-scope or an inverse-scope representation, the plural continuation is compatible only with an inverse-scope interpretation that includes multiple climbers. Yet apparently this interpretation is so strongly dispreferred that even a plural disambiguating sentence does not often lead perceivers to construct a representation with multiple climbers. I return to this puzzle in the discussion of the results of Experiment 5.

Experiment 5

Experiment 5 presented two sets of sentences in contexts that supported either the surface-scope interpretation or the inverse-scope interpretation. The experiment included sentences with both configurations of quantifiers: a...every and every...a. Each doubly quantified sentence was embedded in a context and followed by a singular or plural disambiguating sentence. Reading times were measured for the entire quantified sentence and the entire disambiguating sentence.
Experiment 5 Method

Stimuli

Experiment 5 took the pairs of quantified and disambiguating sentences used in Experiment 4 and embedded them in contexts that supported either the surface-scope interpretation or the inverse-scope interpretation. These contexts were the same ones that were used in the norming questionnaires in Experiments 2 and 3.

There were two sets of stimulus items: 24 with a...every sentences and 10 with every...a sentences. Each item appeared in one of two contexts, and the disambiguating sentence for each item was either singular or plural. The experiment thus had a 2x2 design, with two possible contexts and two possible disambiguating sentences for each item. The four conditions of a single stimulus item are illustrated for an a...every item in (74)-(77); the full set of stimuli is provided in the Appendix.

(74) **Surface-scope context, singular disambiguating sentence**
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination.
One weekend, the climbing equipment shop sponsored a show to demonstrate the sport.
While an announcer described the techniques, an experienced climber scaled every cliff.
The climber was very skilled.
The shop's sales increased substantially the next weekend.

(75) **Surface-scope context, plural disambiguating sentence**
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination.
One weekend, the climbing equipment shop sponsored a show to demonstrate the sport.
While an announcer described the techniques, an experienced climber scaled every cliff.
The climbers were very skilled.
The shop's sales increased substantially the next weekend.
Inverse-scope context, singular disambiguating sentence
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination. One weekend, the climbing equipment shop sponsored a race between climbing enthusiasts. While an official timed the event, an experienced climber scaled every cliff. The climber was very skilled. The shop's sales increased substantially the next weekend.

Inverse-scope context, plural disambiguating sentence
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination. One weekend, the climbing equipment shop sponsored a race between climbing enthusiasts. While an official timed the event, an experienced climber scaled every cliff. The climbers were very skilled. The shop's sales increased substantially the next weekend.

Example (74) illustrates the way that the stimuli in Experiment 5 were presented: each line in the example corresponds to one presentation region in the self-paced reading experiment. Notice that the whole-sentence presentation is different from the word-by-word presentation used in Experiment 4.

Just as in Experiment 4, every item was followed by a how many question with two possible answers that corresponded to the surface-scope and inverse-scope interpretations of the quantified sentence. For example, the a...every items like (74)–(77) were followed by question (78), where the response One corresponds to the surface-scope interpretation and the response Several to the inverse-scope interpretation.

(78) How many climbers scaled cliffs?

One. Several.
For the every...a item depicted in (79), the comprehension question is given in (80). Noted that for these items, the question response One corresponds to the inverse-scope interpretation and the response Several to the surface-scope interpretation.

(79) The granddaughter of a famous politician recently donated some old papers to the university archives. Much of the collection was junk, but there were some items of interest. Every historian examined a document. The document was a stirring speech. The university was lucky to have received the donation.

(80) How many documents were examined?

One. Several.

The stimuli were arranged in a list that included 24 items of similar structure that were used in other experiments. The filler items were the same as the fillers used in Experiment 2: each filler item consisted of a sentence that was ambiguous because of an elided constituent, embedded in a context paragraph and followed by a comprehension questions. Four different versions of the list were created using a Latin Square design, such that each list contained eight or nine items in each of the four conditions. The items were arranged in a different, pseudo-random order in each version of the list.

Procedure

The procedure was identical to that used in Experiment 4.

Participants

Forty-one Northwestern University undergraduates participated in the experiment for course credit, 11 men and 30 women. The data from 11 participants were excluded from the analysis: four who were not native speakers of English, four who had already participated in a related experiment, and three for whom the experiment was unexpectedly interrupted.
Experiment 5 Results

There were three measures for each stimulus item: the reading time for the ambiguous quantified sentence, the reading time for the disambiguating sentence, and the proportion of responses to the comprehension question.

Sub-Experiment 5A

The mean residual reading time for the a...every quantified sentence, averaged across subjects and items, was 298 milliseconds in the surface-scope context and 334 milliseconds in the inverse-scope context. There was no evidence of an effect of context on the reading times ($F^2\less 1$).

The graph given below (Figure 3) shows the residual reading times for the disambiguating sentence in the four conditions.
Figure 3. Mean residual reading times for disambiguating sentence, Experiment 5A.
The plural disambiguating sentence, which disambiguated to the inverse-scope interpretation, was read significantly more slowly than the singular, surface-scope disambiguating sentence ($F_{1}(1,31)=9.5, p<0.01; F_{2}(1,23)=4, p=0.06$). While the reading times for the disambiguating sentence were slower following the inverse-scope context than following the surface-scope context, this difference approached significance only in the analysis by subjects ($F_{1}(1,31)=3.2, p=0.08; F_{2}(1,23)=2.1, p=0.16$). There was no evidence of an interaction between the context and the number of the disambiguating sentence ($F_{1}, F_{2} <1$).

The following table (Table 6) provides the proportions of responses to the comprehension question. There was a significant main effect of the number of the disambiguating sentence on the proportion of responses chosen ($F_{1}(1,31)=31, p<0.001; F_{2}(1,23)=83, p<0.001$). Perceivers chose the surface-scope response One more often when the disambiguating sentence was singular and the inverse-scope response Several more often when the disambiguating sentence was plural. Nevertheless, the singular sentence appears to have been a more effective disambiguator than the plural, since it led to 70-75% surface-scope response, while the plural sentence led to only 55% inverse-scope responses. While neither the singular nor the plural sentence disambiguated the quantified sentence entirely, the singular sentence did so more consistently. Just as in Experiment 4, perceivers may have occasionally selected the singular response when they had computed an inverse-scope representation, since the singular response is consistent with the truth conditions of the inverse-scope interpretation. There was no evidence for an effect of the context on the proportion of responses ($F_{1}, F_{2} <1$) nor of an interaction between context and number ($F_{1}, F_{2}, <1$).
Table 6.

Experiment 5A: Proportions of responses to comprehension question following a...every sentence in context.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response One.</th>
<th>Inverse-scope response Several</th>
</tr>
</thead>
<tbody>
<tr>
<td>surface-scope context, singular disambiguating sentence</td>
<td>0.75</td>
<td>0.25</td>
</tr>
<tr>
<td>surface-scope context, plural disambiguating sentence</td>
<td>0.45</td>
<td>0.55</td>
</tr>
<tr>
<td>inverse-scope context, singular disambiguating sentence</td>
<td>0.70</td>
<td>0.30</td>
</tr>
<tr>
<td>inverse-scope context, plural disambiguating sentence</td>
<td>0.44</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Sub-Experiment 5B

The mean residual reading time for the every...a quantified sentence, averaged across subjects and items, was 136 milliseconds in the surface-scope context and 137 milliseconds in the inverse-scope context. There was no evidence for an effect of context on the reading times ($F's<1$).

The residual reading times for the disambiguating sentence in the four conditions are depicted below (Figure 4).

The singular disambiguating sentence, which disambiguated to the inverse-scope interpretation, was read more slowly than the plural, surface-scope disambiguating sentence
(F1(1,63)=12, p<0.001; F2(1,9)=2.7, p=0.14). This difference is significant only in the analysis by subjects, probably because of the small number of items. There was no evidence of an effect of the context or of an interaction between the context and the number of the disambiguating sentence (all F's <1).
Figure 4. Mean residual reading times for disambiguating sentence, Experiment 5B.
The following table (Table 7) provides the proportions of responses to the comprehension question.

Table 7.

Experiment 5B Proportions of responses to comprehension question following every...a sentence in context.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response Several</th>
<th>Inverse-scope response One</th>
</tr>
</thead>
<tbody>
<tr>
<td>surface-scope context, plural disambiguating sentence</td>
<td>0.87</td>
<td>0.13</td>
</tr>
<tr>
<td>surface-scope context, singular disambiguating sentence</td>
<td>0.29</td>
<td>0.71</td>
</tr>
<tr>
<td>inverse-scope context, plural disambiguating sentence</td>
<td>0.77</td>
<td>0.23</td>
</tr>
<tr>
<td>inverse-scope context, singular disambiguating sentence</td>
<td>0.31</td>
<td>0.69</td>
</tr>
</tbody>
</table>

There was a significant main effect of the number of the disambiguating sentence on the proportion of responses chosen ($F_{1}(1,63)=240$, $p<0.0001$; $F_{2}(1,9)=62$, $p<0.0001$). Perceivers chose the surface-scope response *Several* more often when the disambiguating sentence was plural and the inverse-scope response *One* more often when the disambiguating sentence was singular. There was no evidence for an effect of the context on the proportion of responses ($F_{1}$, $F_{2}<1$). There was, however, near-significant evidence of an interaction between context and number ($F_{1}(1,63)=3.7$, $p=0.06$; $F_{2}(1,9)=4.4$, $p=0.07$). Notice that both the plural and singular sentences were about equally effective at disambiguating the quantified sentence, since perceivers chose the response that corresponded to the disambiguating sentence about 70% of the time regardless of the
direction of the disambiguation. This result is slightly different from that observed for the
*a...every* sentences, where a singular sentence disambiguated to the surface-scope
interpretation more effectively than a plural sentence did to the inverse-scope
interpretation.

**Experiment 5 Discussion**

The results of Experiment 5 are remarkably similar to those of Experiment 4: the
sentence that disambiguated to inverse scope was read more slowly for both *a...every* and
*every...a* sentences even when the quantified sentence was embedded in a context that
supported the inverse-scope interpretation. Apparently the processing cost associated with
assigning the inverse-scope interpretation was not mitigated by the helpful discourse
context. Recall that the principle of parsimony (Altmann & Steedman, 1988; Crain &
Steedman, 1985) argued that the inverse-scope interpretation is difficult because it requires
accommodation of more discourse entities (e.g., multiple climbers in (81)) into the
discourse representation than the surface-scope interpretation does.

\[(81) \quad \text{A climber scaled every cliff.}\]

The model predicts that this difficulty would be alleviated if the discourse representation
already included the multiple entities, since the quantified sentence would then not
introduce any new presuppositions. The present data show, however, that assigning the
inverse-scope interpretation incurs a processing cost even when the discourse context has
already introduced the multiple entities. Clearly the relative complexity of constructing the
mental representation is not the only source of the cost of assigning inverse scope.

Indeed, if the referential complexity of the discourse representation associated with
a doubly quantified sentence were the chief factor, we would predict no difficulty at all in
assigning the inverse-scope interpretation to an *every...a* sentence. In a sentence like (82),

\[(82) \quad \text{Every player rubbed a lucky charm.}\]

the inverse-scope interpretation involves fewer entities, since there is only a single charm
that all the players rub. The principle of parsimony cannot account for the slower reading
times at the singular disambiguating sentence following a *every...a* sentence in an inverse-
scope context.
On the other hand, there was no difference in reading times for the quantified sentence itself in the two contexts, for either the *a...every* or the *every...a* items. It is possible that perceivers simply committed to the surface-scope interpretation for the most items, regardless of context, which led to similar reading times for both contexts. Perceivers then experienced processing difficulty upon encountering the inverse-scope disambiguating sentence that required them to reactivate an already discarded representation. However, comparing the comprehension question response data for the *a...every* sentences to those from Experiment 4 weakens this account. In Experiment 4, where the quantified sentence appeared without any context, the plural continuation sentence led perceivers to select the plural inverse-scope response on only 41% of trials. In Experiment 5, where the quantified sentence followed an inverse-scope supporting context, perceivers selected the plural response following 56% of plural continuations and even following 30% of singular continuations. Both experiments showed evidence of processing difficulty at the plural continuation sentence. If this difficulty were due simply to the cost of activating the discarded inverse-scope representation, we should expect the cost to be lower in Experiment 5, where the context activated the inverse-scope representation enough to allow perceivers to choose the plural response more frequently than in Experiment 4. But the degree of processing difficulty appears to be the same for the plural continuations to *a...every* sentences in both experiments, and indeed was somewhat greater following the inverse-scope context than the surface-scope context. It seems unlikely, then, that the inverse-scope supporting context that led to a greater proportion of plural responses in Experiment 5 than in Experiment 4 could have failed to activate the inverse-scope representation for parsing the ambiguous quantified sentence. Rather, perceivers may have simply delayed in resolving the scope ambiguity until they encountered the continuation sentence, at which point the cost of assigning the inverse-scope interpretation was incurred.

Nevertheless, the proportions of responses to the comprehension questions for both *a...every* and *every...a* items, where one quarter to one half of responses were inconsistent with the context, even following a matching continuation sentence, suggest that the contexts themselves may not have been very strongly biasing. Experiments 6 and 7, which make it possible to compare reading times according to what interpretation perceivers assigned, shed some light on this issue.

**Experiment 6**

Experiments 6 and 7 investigated the on-line comprehension of doubly quantified *a...every* sentences that were not followed by a singular or plural continuation sentence.
Because the quantified sentences were not disambiguated by the following sentence, the responses to the comprehension questions are a more trustworthy indicator of the interpretation that perceivers assigned to the quantified sentence than in Experiments 4 and 5 where the quantified sentence was followed by a singular or plural disambiguating sentence. It is possible in Experiments 6 and 7, therefore, to compare reading times for inverse-scope interpreted quantified sentences against those for surface-scope interpreted sentences. These comparisons show that assigning inverse scope to a doubly quantified sentence incurred a processing cost, and that this cost was not mitigated by a favourable discourse context.

Experiment 6 Method

Stimuli

The 24 stimulus items in Experiment 6 consisted of the same doubly quantified *a...every* sentences used in Experiment 4, followed by a sentence that did not disambiguate the quantified sentence. The stimuli were presented one word at a time. Each pair of sentences was followed by a comprehension question with two possible answers, as shown in (83).

(83) How many paratroopers jumped from planes?

One.  Several.

The stimuli were arranged in pseudo-random order in a list that included 68 items used in other experiments. Forty-eight of these items were ambiguous because they contained elided constituents, like those included in Experiments 2 and 4. Twenty filler items contained local ambiguities like the one in (84).

(84) Friends in Washington heard the news about Bill Clinton and other friends heard the rumor about Hillary was not fully accurate.

In sentence (84), the noun phrase following the verb *heard* could be the object of the verb or the subject of a sentential complement clause. The temporary ambiguity is resolved by
the end of the sentence. Forty-eight of the filler items were followed by a comprehension question.

*Procedure*

The procedure was identical to that used in Experiment 4.

*Participants*

Thirty-six Northwestern University undergraduates participated in the experiment for course credit.

*Experiment 6 Results*

Residual reading times were calculated for each word and over the entire quantified sentence and the entire ellipsis sentence. The reading times plotted in the two figures below (Figure 5 and Figure 6) compare trials on which perceivers gave the surface-scope response *One* to the comprehension question versus trials on which they gave the inverse-scope answer *Several*. Figure 5 shows the reading times calculated over the entire quantified sentence, while Figure 6 shows them for each word.
Figure 5. Mean residual reading time for quantified sentence, contingent on comprehension question response, Experiment 6.
Figure 6. Mean residual reading times for words in quantified sentence, contingent on comprehension question response, Experiment 6.
Both figures show clearly that the quantified sentence was read more slowly when perceivers assigned it an inverse-scope interpretation than when they assigned a surface-scope interpretation. The effect of interpretation was statistically significant ($F_1(1,63)=6.4$, $p=0.01$; $F_2(1,43)=4.3$, $p=0.05$) for the residuals calculated over the entire sentence. Figure 6 shows that this pattern holds for every word in the sentence, although the differences were not significant at any word in the sentence.

**Experiment 6 Discussion**

The analysis of the reading times according to perceivers’ answers to the comprehension question shows that perceivers read the quantified sentence more slowly when they assigned it the inverse-scope interpretation than when they assigned it the surface-scope interpretation. Unlike in Experiment 4, the cost of the inverse-scope interpretation shows up during the quantified sentence, and not just at the following disambiguating sentence. We can conclude that the cost is not simply that of reactivating the inverse-scope representation upon reading the plural sentence once the processor has committed to the surface-scope interpretation of the quantified sentence. Rather, it appears that assigning an inverse-scope interpretation to a quantified sentence can incur a cost at the quantified sentence itself.

The question then arises of why this cost was observed in Experiment 6 but not in Experiments 4 and 5. It is possible that the difference in results is a methodological artifact. In addition to the ambiguous doubly quantified sentences, Experiment 6 included quantified sentences that were unambiguous, while in Experiments 4 and 5 all of the quantified sentences were ambiguous. Perhaps the surrounding unambiguous items in Experiment 6 made perceivers more likely to commit early to a single interpretation of the ambiguous doubly quantified sentence, incurring a cost for assigning inverse scope, while perceivers in Experiments 4 and 5 were content to leave the ambiguity unresolved until encountering the singular or plural continuation sentence.

**Experiment 7**

Experiment 6 presented the 24 doubly quantified a...every sentences embedded in the biasing contexts from Experiment 5A. Because the experiment was designed to investigate the Scope Economy proposal dealt with in Chapter 3, the quantified sentences
were followed by a VP-ellipsis sentence. The ellipsis sentence is not relevant for present purposes, but I return to consider that portion of the Experiment 7 data in Chapter 3, focusing here on the quantified sentences in context. Experiment 7 also includes two unambiguous baseline conditions.

Experiment 7 Method

Stimuli

The 24 stimulus items used in Experiment 7 were a modification of the *a...every* items in contexts from Experiment 5A. The items included the same 24 quantified sentences, appearing in one of two contexts, which supported either the surface-scope interpretation or the inverse-scope interpretation. The sentence following the quantified sentence was not a singular or plural disambiguating sentence, but rather an ellipsis sentence.

Experiment 7 also included two baseline conditions, shown below in (85) and (86):

(85) The climbing expert scaled every cliff.

(86) A different climber scaled every cliff.

Sentence (85) is unambiguous surface scope and was embedded in the surface-scope context, while sentence (86) is unambiguous inverse scope and was embedded in the inverse-scope context. The experiment thus had a 2x2 design with two types of context and two levels of ambiguity.

Like Experiment 5, the stimuli in Experiment 7 were presented in regions one clause or sentence in length. The quantified sentence occupied a single presentation region. All of the experimental items were followed by a how-many comprehension question like the ones shown above in (78).

The 24 stimuli were arranged in a list along with 24 filler items of similar structure that were used in a different experiment. The filler items did not include a comprehension question. The items were arranged in pseudo-random order such that no two items in the same condition were adjacent to each other. Four different versions of the list of stimulus items were created, with the items appearing in a different order in each version of the list.
Procedure

The procedure was the same as for Experiment 5.

Participants

Twenty-four Northwestern University students, four men and 20 women with an age range of 17 to 46 years, participated in the experiment and received course credit. All participants were native speakers of English.

Experiment 7 Results

Figure 7 gives the residual reading times for the quantified sentence in Experiment 7 across the four conditions. All reading times are reported in milliseconds and are averaged over subjects and items.

The reading times for the quantified sentence showed a significant main effect of context ($F_1(1,23)=18, p<0.001$; $F_2(1,23)=17, p<0.001$). Sentences in the inverse-scope conditions (both ambiguous and unambiguous) were read more slowly than those in the surface-scope conditions. There was no evidence for an effect of ambiguity ($F_1(1,23)=1.4, n.s.; F_2(1,23)=1.6, n.s.$) or for an interaction between the two factors ($F_1, F_2<1$).

As in Experiment 6, the reading times for the ambiguous doubly quantified sentences were also analyzed with respect to the response to the comprehension question. Those results are depicted below (Figure 8).
Figure 7. Mean residual reading time for quantified sentence, Experiment 7.
Figure 8. Mean residual reading times for ambiguous quantified sentence, contingent on comprehension question response, Experiment 7.
The quantified sentence was read significantly more slowly on trials in which perceivers gave the inverse-scope response than on trials that earned a surface-scope response ($F_{1}(1,46)=7.2$, $p=0.01$; $F_{2}(1,46)=6.4$, $p=0.02$). This pattern is very similar to that observed above in Experiment 6, where no context paragraph preceded the quantified sentence. The proportions of responses to the comprehension questions are given in Table 8.

Comparing the two ambiguous conditions, there was a significant effect of biasing context on the proportion of surface-scope responses chosen ($t_{1}(46)=3.5$, $p=0.001$; $t_{2}(46)=3.3$, $p=0.002$). Participants chose the response One more frequently following the surface-scope context than following the inverse-scope context. Comparing the two conditions with the inverse-scope supporting context, there was a significant of ambiguity ($t_{1}(46)=6.1$, $p<0.0001$; $t_{2}(46)=6.2$, $p<0.0001$). Participants chose the surface-scope response One more often when the quantified sentence was ambiguous than when it was unambiguous inverse-scope.
Table 8.

Experiment 7 Proportion of responses to comprehension question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response</th>
<th>Inverse-scope response</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.68</td>
<td>0.32</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.08</td>
<td>0.92</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.44</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Correct response | Incorrect response

| unambiguous surface scope | 0.95 | 0.05 |

Experiment 7 Discussion

Experiment 6 clearly showed that assigning an inverse-scope interpretation to an ambiguous doubly quantified sentence incurred a processing cost. The results of Experiment 7 show that this fact holds true even when the quantified sentence follows a context that supports the inverse-scope interpretation. Because the inverse-scope supporting context does not mitigate the cost of assigning the inverse-scope interpretation, I conclude that this cost is not simply a result of the referential complexity of the inverse-scope representation, since the multiple entities necessary for the inverse-scope interpretation had already been introduced by the context when the processor encountered the quantified sentence. These data are therefore contrary to the predictions of the principle of parsimony (Altmann & Steedman, 1988; Crain & Steedman, 1985).
The unambiguous baseline conditions in Experiment 7 provide another valuable comparison, since they show that the inverse-scope interpretation incurs a cost even when it is the only possible interpretation. As Figure 7 shows, the quantified sentence was read just as slowly when it was unambiguous inverse-scope as when it was ambiguous in an inverse-scope context. This fact is compelling evidence against Kurtzman and MacDonald’s (1993) parallel processing model. In their model, processing difficulty arises when the two representations are equally weighted thanks to the input of the various competing constraints that govern interpretive processes. Because both representations are equally activated, the processor has difficulty committing to one of them. In the unambiguous inverse-scope condition, no other interpretation was possible for the quantified sentence. There should be no competition between representations, and there is no reason to predict processing difficulty. But in fact perceivers did experience significant processing difficulty at the unambiguous inverse-scope quantified sentence, difficulty which could not have arisen from competition with an alternative representation.

The one difference between the unambiguous inverse-scope quantified sentence and the ambiguous quantified sentences is that the unambiguous sentence contains the adjective *different*. It is possible that this adjective has properties that contribute to the processing load. Experiment 8 tested this possibility.

**Experiment 8**

Experiment 8 was designed to test the possibility that the processing cost observed at the unambiguous inverse-scope quantified sentence in Experiment 7 was attributable to some difficulty associated with processing the adjective *different*. Using eight quantified sentences embedded in contexts and followed by disambiguating sentences, I found that quantified sentences containing *different* were processed no more slowly than sentences without *different*.

**Experiment 8 Method**

**Stimuli**

The experiment used eight sets of stimulus items, each of which consisted of a doubly quantified sentence embedded in a context and followed by a disambiguating sentence. In a 2x2 design I manipulated the order of the quantifiers in the sentence and
the presence of different. There were thus four conditions for each item, as illustrated below in (87)–(90).

(87) The members of the gourmet club decided to publish a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. Every member tested a recipe. The recipes were very tasty. When the cookbook was finally published, it was a big success.

(88) The members of the gourmet club decided to publish a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. Every member tested a different recipe. The recipes were very tasty. When the cookbook was finally published, it was a big success.

(89) The members of the gourmet club decided to publish a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. A member tested every recipe. The members made valuable corrections. When the cookbook was finally published, it was a big success.

(90) The members of the gourmet club decided to publish a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. A different member tested every recipe. The members made valuable corrections. When the cookbook was finally published, it was a big success.
In conditions 1 and 2, (87) and (88), the quantifiers were in the every...a order, while in conditions 3 and 4, (89) and (90), the quantifiers were in the a...every order. Conditions 2 and 4 included different but conditions 1 and 3 did not. In every condition, the context supported the interpretation of the doubly quantified sentence that was consistent with different. In other words, the context, which was the same across all four conditions, supported the surface-scope interpretation of the every...a sentences and the inverse-scope interpretation of the a...every sentences. The following disambiguating sentence was always a plural sentence, again consistent with the different interpretation. In short, all circumstances favoured the interpretation of the quantified sentence that was consistent with different, so if perceivers experienced processing difficulty at the quantified sentence in conditions 2 and 4 as compared with conditions 1 and 3, this difficulty could be attributed to the presence of different and not to the context or to the direction of the disambiguation.

The stimuli were presented in regions one clause or sentence long. The quantified sentence and disambiguating sentence were each presented in a single region.

The eight stimulus items were arranged in pseudo-random order in a list that included 45 other items that were used in other experiments. Four versions of the list were prepared, with the items in a different order in each list.

Procedure

The procedure was the same as that used in the previous self-paced reading experiments.

Participants

Thirty-two Northwestern University undergraduates participated in Experiment 8, 11 males and 21 females. The data were analyzed from 24 of the participants. Of the eight participants whose data were not analyzed, four were not native speakers of English, three had already participated in a related experiment and had been debriefed, and one self-identified as having dyslexia.
Experiment 8 Results

Mean residual reading times were calculated across subjects and items for the quantified sentence and the following disambiguating sentence. The reading times for the quantified sentence are shown in Figure 9.

As Figure 9 shows, the sentences containing different were read no more slowly than those without different, for either configuration of quantifiers. There was no evidence for an effect of the presence of different or of the order of quantifiers on the reading times (all F's <1). Clearly, different does not add processing difficulty to a sentence.

Figure 10 shows the residual reading times for the disambiguating sentence. The figure shows clearly that the plural sentence was read more slowly following the a...every sentences, where it disambiguated to inverse scope, that in the every...a sentences where it disambiguated to surface scope. Furthermore, the disambiguating sentence was read more slowly following an ambiguous quantified sentence than following an unambiguous sentence that contained different.
Figure 9. Mean residual reading time for quantified sentence, Experiment 8.
Figure 10. Mean residual reading time for disambiguating sentence, Experiment 8.
There was a significant main effect of the order of the quantifiers in the quantified sentence ($F(1,23)=7.6, p=0.01; F(1,7)=5, p=0.06$) on the reading times for the disambiguating sentence. The effect of the presence of different in the quantified sentence was significant only in the analysis by items ($F(1,23)=2.2, p=0.15; F(1,7)=8.1, p=0.03$). There was no evidence of an interaction between the two factors.

Experiment 8 Discussion

There was no difference in the reading times for the quantified sentence in any of the four conditions. In other words, the sentences that contained different were not read significantly more slowly than those that did not. The results of Experiment 8 make it clear that different does not introduce a processing cost, even to an inverse-scope quantified sentence. We can conclude, then, that the slow reading times for the unambiguous inverse-scope quantified sentence were due to the difficulty of assigning the inverse-scope interpretation and not to some property of the adjective different.

General Discussion

The results of all of the self-paced reading experiments showed that perceivers experienced processing difficulty when they assigned an inverse-scope interpretation to a doubly quantified sentence. Experiments 5 and 7 presented the quantified sentences in contexts that supported either the surface-scope or the inverse-scope interpretation. The cost associated with assigning inverse scope was observed even for sentences following an inverse-scope supporting context that introduced multiple entities into the discourse context. The cost also showed up in Experiment 5 for every...a sentences, whose inverse-scope representation is referentially simpler than the surface-scope interpretation because it requires instantiating fewer entities in the discourse model. The predictions of the principle of parsimony (Altmann & Steedman, 1988; Crain & Steedman, 1985) are thus not borne out, and I conclude that the cost of assigning inverse scope is not simply the cost associated with constructing a discourse model that includes multiple entities.

Experiments 6 and 7 showed that the cost of inverse scope is incurred at the quantified sentence and not just at the disambiguating sentence, as it was in Experiments 4 and 5. This result indicates that the processing difficulty is not simply that of reactivating an interpretation that had been discarded at the close of the previous sentence (Bock, Loebell, & Morey, 1992; Jarvella, 1971; Just & Carpenter, 1980; Lombardi & Potter,
1992; Potter & Lombardi, 1998; Sachs, 1967), but arises during the processing of the sentence that receives the inverse-scope interpretation.

Experiment 7 showed, furthermore, that even an unambiguous inverse-scope quantified sentence incurs a processing cost, contrary to the predictions of a parallel-processing model (Kurtzman & MacDonald, 1993) where processing difficulty indicates competition between representations. For a sentence like (91) embedded in a supportive context,

(91) A different paratrooper jumped from every plane.

the Quantifier Hierarchies of Ioup and VanLehn predict a preference for the inverse-scope interpretation since every tends to take wide scope, the principle of parsimony predicts relatively easy processing since the discourse representation already includes multiple paratroopers, and the parallel-processing model predicts easy processing since the sentence is unambiguous. Nevertheless, perceivers read sentence (91) significantly more slowly than its surface-scope counterparts.

Taken together, all of these facts point to an account of quantifier scope comprehension that predicts a cost for interpreting quantifiers with a scope configuration other than their configuration in the surface syntax, such as the principle of Processing Scope Economy. According to the PSE principle, an inverse-scope configuration incurs a processing cost because its abstract linguistic structure is more complex than a surface-scope configuration. The principle is thus consistent with the Quantifier Raising (May, 1977, 1985) and Flexible Types (Partee, 1986; Partee & Rooth, 1983) accounts of quantifier scope ambiguity, since inverse-scope representations are more complex than surface-scope ones in both of these theories. In contrast, in the CCG theory (Steedman, 2000, 2003), where the interpretation of a scope-ambiguous sentence depends on the timing of steps in the derivation, neither representation is more complex than the other. The results of Experiments 4–8 demonstrate that the on-line comprehension of an inverse-scope quantified sentence gives rise to difficulty beyond what can be explained by extra-grammatical factors. I conclude, therefore, that processing quantifier scope is a crucially grammatical operation.
CHAPTER 3:
SCOPE ECONOMY IN REAL TIME
Several scholars have observed that a scope-ambiguous sentence is sometimes restricted to the surface-scope interpretation when a VP-ellipsis sentence with a singular definite subject follows it (Hirschbühler, 1982; Sag, 1976; Williams, 1977). For example, sentence (92) in isolation can mean either that there is a single student who loves all the films (the surface-scope interpretation) or that there are several students who each love a different film (the inverse-scope interpretation). But when sentence (93) follows, only the surface-scope interpretation of (92) is possible.

(92) Some student loves every Hitchcock film.

(93) Marta does, too.

Although it has been widely assumed in the literature that the ellipsis sentence (93) is grammatical only on the surface-scope reading of (92), this intuition is not universally accepted. For example, Johnson & Lappin (1997) provide an example (94) that appears not to share the disambiguating effect found in (92)–(93).

(94) At least two cabinet ministers bear responsibility for each government department, and the Prime Minister does too.

(Johnson & Lappin, 1997, p. 311)

Sentence (94) seems perfectly acceptable on the interpretation where each department is the responsibility of a different pair of cabinet ministers but all are ultimately the responsibility of the Prime Minister. Steedman (p.c., 26 Oct 2003) suggests that the disambiguating effect of (93) is “illusory”, arising as a result of real-world knowledge and contextual plausibility. One advantage to using psycholinguistic experimentation to investigate the theoretical proposals about quantifier scope ambiguity is that the data obtained are relatively objective in comparison with intuitive grammaticality judgments. Such objectivity is particularly valuable in a case such as this where the grammaticality judgments are contested.

Fox (1995; 2000) has proposed an account, building on May’s (1977; 1985) Quantifier-Raising (QR) theory of quantifier interpretation\(^4\), that locates the disambiguating effect of ellipsis in the syntax of these constructions. This chapter describes

\(^4\) A detailed description of QR is provided in Chapter 1.
a set of experiments that investigate Fox's account by studying the real-time comprehension of constructions like (92)–(93). Based on the results of the experiments, which indicate that ellipsis sentences like (93) do not consistently disambiguate doubly quantified sentences, I argue for an alternative account wherein the effect arises because of the properties of the human sentence processing mechanism and more general cognitive principles.

Theoretical Background

Fox's (1995; 2000) account of the interaction of VP-ellipsis with quantifier scope ambiguity begins with the following reasoning: Because a doubly quantified sentence like (92) has two different interpretations, we can infer that there exist two distinct LF configurations for a sentence with that surface structure. In (95), which represents schematically\(^5\) one possible LF configuration for sentence (92), the quantified DP some student occupies a position higher in the tree than the DP every film. The phrase some student thus has scope over every film, giving rise to the surface-scope interpretation where a single student loves all the films.

\[(95)\]

\[
\text{some student} \quad \text{loves} \quad \text{every film}
\]

\(^5\)In Chapter 3 I use simplified tree diagrams like (95) to represent the configurations of the quantifiers at LF. The discussion of quantifier scope relations in Chapter 1 provides fully detailed depictions of the LF representations.
On the other hand, if every film is QRed to an LF position higher than some student, as illustrated in (96), then every film takes scope over some student, leading to the inverse-scope interpretation where each film is loved by some possibly different student.

(96)

As I illustrated in Chapters 1 and 2, on the QR account of quantifier scope ambiguity the two interpretations of a doubly quantified sentence result from two distinct LF configurations. Sentence (92) is scopally informative since its interpretation necessarily reveals what the configuration of quantified phrases is at LF.

In principle, even a sentence like (97), which contains only one quantifier phrase, should have the option for two distinct LFs, since QR allows every computer either to be inside the scope of Vivek or to take scope over Vivek, as shown in (98)–(99).

(97) Vivek fixed every computer.

However, sentence (97) does not allows us to infer what position the DP every computer occupies at LF, since the truth conditions are the same for both LF configurations, as the paraphrases in (98)–(99) illustrate.
Since the interpretation of (97) could have resulted from either configuration of the subject and object DPs, the sentence is *scopally uninformative*. This kind of sentence is also referred to as *scopally commutative*. Fox (2000) proposed that only the simpler of the two configurations, i.e., (98), is licensed in a scopally uninformative sentence. He proposed a principle of grammatical Scope Economy, quoted in (100).

(100) **Scope Economy** (Fox, 2000, p. 23)

SSOs [scope-shifting operations] that are not forced for type considerations must have a semantic effect.

In other words, QR is permitted only if the truth conditions of the resulting LF would be different from the truth conditions of the un-QRed configuration. For sentence (97), then, the inverse-scope LF configuration (99) is ungrammatical because the QR operation that
raises every computer over Vivek is semantically vacuous: the truth conditions of configuration (99) are no different from those of (98). I shall refer to Fox’s principle by the name of Grammatical Scope Economy (GSE).

The empirical evidence for GSE is quite subtle. The very sentences that the principle constrains, that is, singly quantified sentences like (97), do not allow us to observed the effects of GSE precisely because they are scopally uninformative. If it is ungrammatical for every computer to take scope over Vivek, semantic intuition does not reveal that fact, since we cannot observe the LF of the sentence. We can observe GSE at work only indirectly, for example in a VP-ellipsis context like the one repeated in (101)–(102).

(101) Some student loves every Hitchcock film.

(102) Marta does [love every Hitchcock film], too.

Let us assume that the elided (i.e., unpronounced) constituent loves every Hitchcock film is present in the LF representation of sentence (102), (Hankamer & Sag, 1976; Kennedy & Merchant, 2000; Sag, 1976). Scope Economy prevents every film from taking scope over Marta, since the QR operation that raises every film to the higher position would have no semantic effect. The LF depicted in (103)b is ungrammatical, and (103)a is the only possible LF for the sentence.

(103) a. 

```
   Marta
     /     
   loves  every film
```

b. 

```
   every film
     /     
   Marta
     /     
   loves  every film
```

* [GSE]

Fox’s analysis relies on the claim that the relationship between the ellipsis sentence and its antecedent sentence is constrained by Parallelism, such that “the scope-bearing elements in the antecedent sentence must receive scope parallel to that of the corresponding elements in the ellipsis ... sentence” (2000, p. 31). In other words, if the
ellipsis sentence has surface scope, then the antecedent sentence cannot have inverse scope, and vice versa. Scholars have argued for the existence of structural parallelism between an ellipsis sentence and its antecedent elsewhere in the literature (Fiengo & May, 1994; Williams, 1977); indeed, the requirement is independently motivated in sentences that do not contain quantification, such as (104) from Fox (1995, p. 306).

(104) John likes flying planes, and Bill does too.

The constituent likes flying planes is ambiguous; its two interpretations are, roughly, likes to fly planes and likes planes that fly. But whichever interpretation the phrase is assigned in the first conjunct, the second, elided conjunct must also have that interpretation. This fact is presumably due to the parallelism requirement on VP-ellipsis constructions. Related formulations of the parallelism principle are proposed by Tancredi (1992) and Rooth (1992) to account for deaccenting phenomena. Merchant (2001) argues that the licensing condition on VP-ellipsis is not structural isomorphism between the elided and antecedent VPs, but rather a semantic condition, such that the two VPs must entail each other. Assuming that quantifier scope ambiguity arises from the possibility for more than one structural configuration of the quantifiers at LF, then Merchant’s semantic condition has the effect of requiring that the quantified phrases in the doubly quantified and ellipsis sentences be in parallel syntactic configurations.

Whatever the precise formulation of the Parallelism principle may be, some degree of parallelism seems to be required between an ellipsis sentence and its antecedent sentence.

Returning to the example sentences (101)–(102), GSE does not apply to sentence (101), since the two different scope configurations for the sentence are truth-conditionally distinct. Rather, the configuration of (101) is constrained by Parallelism, which requires the LFs of the two sentences to be isomorphic with each other in order for the VP-ellipsis construction to be licensed. The result is that the doubly quantified sentence is restricted to the LF shown in (105)a.
The LF of the ellipsis sentence is constrained by Scope Economy, with the result that the LF of the quantified sentence is constrained indirectly via Parallelism with the ellipsis sentence. It follows, then, that (101) has only the surface-scope interpretation where a single student loves all the films.

A major theoretical consequence of introducing Scope Economy as a grammatical principle that conditions operations like QR is that the syntactic well-formedness of an LF configuration is determined by aspects of the sentence’s meaning. Ever since Chomsky’s Syntactic Structures, the prevailing view has been that “grammar [syntax] is autonomous and independent of meaning” (1957, p. 17). Under this assumption of modularity, the only interaction of the two components occurs at the interface where the semantics interprets the structure that the syntax has generated. Fox agreed that allowing syntactic operations to be subject to “every property that relates to the way a sentence can contribute to thought or communication” (2000, p. 68) was theoretically undesirable, and proposed instead that the language faculty includes a “deductive system” that can prove whether two potential (syntactic) representations are logically equivalent. Thus, syntactic well-formedness is not sensitive to all considerations of sentence meaning, but only to those properties that can be used to prove logical equivalence, such as quantification, negation, and conjunction, among others. Nevertheless, the long-cherished idea of the autonomy of the syntax is seriously undermined by the introduction of Scope Economy into the grammar.

Because the combination of GSE and Parallelism make the interpretation of a doubly quantified sentence in this environment dependent on the structure and interpretation of the sentence that follows it, Fox’s proposal naturally raises a set of processing questions about the time-course of interpretive decisions. In real time, these sentences are perceived by a reader or listener one after the other. If an ellipsis sentence is
to constrain the interpretation of the preceding quantified sentence, then the mental representation of the quantified sentence must be open to modification well after the point where the sentence has been perceived and parsed. For this to be possible, either the processor must delay in assigning an interpretation to the scope-ambiguous sentence, leaving it open to the influence of the ellipsis sentence, or else the processor must be able to revise its earlier decision at a later point once the ellipsis sentence is parsed. For a processor that obeys grammatical principles, as I assume, GSE will have the consequence of prompting reanalysis when the processor encounters a GSE violation. In other words, if a principle such as Parallelism prompts the processor to build an ungrammatical scope configuration, the processor will attempt to reanalyze the preceding material. The goal of the reanalysis is to construct a configuration that obeys GSE. In the case of a doubly quantified sentence followed by an ellipsis sentence, the only configuration that obeys both GSE and Parallelism is one where both sentences have surface scope. If the quantified sentence has inverse scope then the processor must reanalyze it with surface scope.

An alternative hypothesis is that the principle of economy in quantifier scope relations is located not in the grammar but in the processor, as a parsing preference for simpler scope configurations. The idea that the processor prefers to build the simplest possible structure is very common in models of the human sentence processing mechanism (Frazier, 1987, 1990, 1999; Frazier & Fodor, 1978; Gibson, 1998; Gorrell, 1995; Marcus, 1980). This alternative is embodied in the principle of Processing Scope Economy that I proposed in Chapter 1, repeated below in (106).

(106) Processing Scope Economy
The human sentence processing mechanism prefers to compute a scope configuration with the simplest syntactic representation (or derivation). Computing a more complex configuration is possible but incurs a processing cost.

6 Most studies of reanalysis during on-line comprehension examine the processor’s attempt to repair or rebuild the structure of the sentence that it is currently parsing when it encounters a local ungrammaticality. In such cases, the “preceding material” to be reanalyzed consists of some subset of whatever has already been parsed in the present sentence. The particular constructions I consider here are unusual in that, if the quantified sentence has been assigned inverse scope, arriving at a grammatical parse for the ellipsis sentence requires the processor to reanalyze the preceding quantified sentence, whose parse has already been completed. There may be some qualitative differences between this kind of reanalysis and the better-studied, within-sentence reanalysis; I address these potential differences later in this chapter.
The PSE principle proposes that the processor prefers to construct surface-scope configurations, but can construct an inverse-scope configuration if this is required, either to arrive at the desired interpretation of a sentence, or to satisfy some requirement of the grammar such as Parallelism. There is therefore no prediction that a doubly quantified sentences will always have surface scope, but only that the inverse-scope computation consumes more processing resources than the surface-scope because its syntactic representation is more complex.

While I am concerned at present with the particular characterization of the processor’s structural preferences proposed in (106), it may well be the case that PSE is simply one instantiation of a more global processing preference for simpler structures. GSE, on the other hand, has less generality, since it is formulated to deal with the special case of semantically vacuous QR.

The questionnaires and self-paced reading experiments described in this chapter test the predictions of the two principles. The results of the experiments lead me to conclude that Scope Economy is not a component of the grammar of quantification, but that its effects originate in the architecture of the sentence processing mechanism.

Experiments

Three off-line questionnaires and four self-paced reading experiments investigated the comprehension and interpretation of doubly quantified sentences followed by ellipsis sentences. In all of the experiments, the target sentences were embedded in contexts that supported either the surface-scope or the inverse-scope interpretation.

Stimuli

In most of the experiments described here, the core stimuli consisted of 24 pairs of sentences like (107)–(108): a quantified sentence followed by a VP-ellipsis sentence.

(107) An experienced climber scaled every cliff.

(108) The instructor did, too.
The pairs of sentences were embedded in contexts about six sentences long. The contexts were designed to support either the surface-scope or the inverse-scope interpretation of the quantified sentence. The ambiguous sentence (107) thus appeared in two conditions: surface-scope biased and inverse-scope biased. There were also two baseline conditions, one where the quantified sentence was unambiguous surface scope (109) in the surface-scope context, and one where the quantified sentence was unambiguous inverse scope (110) in the inverse-scope context.

(109) The climbing expert scaled every cliff.

(110) A different climber scaled every cliff.

The subject of the ellipsis sentence (108) was always a definite NP, such that it would be expected to disambiguate the quantified sentence to the surface-scope interpretation according to Scope Economy.

Each item thus appeared in one of four conditions: surface-scope unambiguous, surface-scope ambiguous, inverse-scope unambiguous, or inverse-scope ambiguous. The full paradigm of one stimulus item is shown in (111)–(114). The complete set of stimuli is provided in the Appendix.

(111) **Surface scope, unambiguous**

With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination.

One weekend, the climbing equipment shop sponsored a show to demonstrate the sport.

While an announcer described the techniques, the **climbing expert scaled every cliff.**

The shop owner did, too.

The shop’s sales increased substantially the next weekend.
(112) **Surface scope biased, ambiguous**
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination.
One weekend, the climbing equipment shop sponsored a show to demonstrate the sport.
While an announcer described the techniques, **an experienced climber scaled every cliff.**
The shop owner did, too.
The shop's sales increased substantially the next weekend.

(113) **Inverse scope, unambiguous**
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination.
One weekend, the climbing equipment shop sponsored a race between climbing enthusiasts.
While an official timed the event, **a different climber scaled every cliff.**
The shop owner did, too.
The shop's sales increased substantially the next weekend.

(114) **Inverse scope biased, ambiguous**
With the increased popularity of adventure sports, the cliffs outside Campbellton were becoming a popular destination.
One weekend, the climbing equipment shop sponsored a race between climbing enthusiasts.
While an official timed the event, **an experienced climber scaled every cliff.**
The shop owner did, too.
The shop's sales increased substantially the next weekend.

In the questionnaires and the self-paced reading experiments, a comprehension question like (115) followed each stimulus item.
(115) How many climbers scaled cliffs?

One. Several.

The response One corresponded to the surface-scope interpretation of the quantified sentence, and the response Several to the inverse-scope interpretation.

Predictions

PSE and GSE make different predictions about the three sets of results of the self-paced reading experiments: (a) reading times for the quantified sentence, (b) reading times for the ellipsis sentence, and (c) the proportions of responses to the comprehension questions. The two principles also make predictions about the proportions of responses to the comprehension questions in the questionnaire experiments. I shall consider first the predictions about the reading times.

PSE predicts that assigning an inverse-scope interpretation to (107) would be more difficult than assigning surface scope since inverse scope is the more complex, dispreferred configuration. In the self-paced reading experiments described in this chapter, processing difficulty should show up as relatively slow reading times. GSE does not make any specific prediction about the comprehension of the quantified sentence (107), but is not inconsistent with the prediction made by PSE.

Both PSE and GSE make the same prediction about the comprehension of the ellipsis sentence (108), but for different reasons. If a perceiver has assigned inverse scope to the quantified sentence, then she should experience processing difficulty at the ellipsis sentence. According to GSE, this processing difficulty would be due to ungrammaticality since assigning inverse scope to the ellipsis sentence, as Parallelism requires, would violate Grammatical Scope Economy. This ungrammaticality prompts the processor to attempt reanalysis of the quantified sentence, consuming processing resources (Ferreira & Henderson, 1991, 1993; Sturt, Pickering, & Crocker, 1999, 2000). According to PSE, the ellipsis sentence should cause processing difficulty on the inverse-scope interpretation of the quantified sentence simply because the processor obeys Parallelism in assigning inverse scope to the ellipsis sentence, which incurs a processing cost because of the complexity of the representation. The processing cost of assigning inverse scope to the ellipsis sentence
should be comparable to the cost of assigning inverse scope the quantified sentence according to PSE.

The two principles make diverging predictions about the proportions of responses to comprehension questions in the questionnaires and self-paced reading experiments. GSE predicts that the processor will reanalyze the quantified sentence to assign it surface scope upon encountering the ellipsis sentence, in order to satisfy Scope Economy and Parallelism. Perceivers should therefore select the surface-scope response one to the comprehension question if they have read and parsed the ellipsis sentence. PSE, in contrast, predicts that the processor will assign surface scope wherever possible, since this is the preferred configuration, but will assign inverse scope if it is necessary. In cases where the quantified sentence has an inverse-scope interpretation, the ellipsis sentence must also have inverse scope to satisfy Parallelism. There is no reason for the processor to abandon an inverse-scope interpretation of the quantified sentence under PSE, and we would therefore expect to see a similar proportion of inverse-scope responses following the ellipsis sentence as in contexts without the ellipsis sentence.

In summary, the two principles make the following empirical predictions about the three experimental measures:

- PSE predicts slow reading times for the quantified sentence in inverse-scope contexts compared to surface-scope contexts. GSE makes no specific prediction about reading times for the quantified sentence.

- PSE predicts slow reading times for the ellipsis sentence in inverse-scope contexts, since Parallelism with the quantified sentence will require a costly inverse-scope configuration. GSE also predicts slow reading times at the ellipsis sentence in inverse-scope contexts because the ellipsis sentence triggers reanalysis of the quantified sentence to surface scope where GSE forbids an inverse-scope configuration.

- PSE predicts an effect of context on the responses to the comprehension question in both conditions with an ambiguous quantified sentence. The proportions of responses should be similar to the proportions for items with no following ellipsis sentence. GSE predicts a high proportion of surface-scope responses to the comprehension question in both conditions with an ambiguous quantified sentence since the inverse-scope interpretation is ungrammatical.
Experiment 9

Experiment 9 was a paper questionnaire that served as a direct follow-up to Experiment 2, presented in Chapter 2. In Experiment 2, the quantified sentence was presented following a context paragraph, but without any subsequent sentence. In the present experiment, the quantified sentence was presented embedded in a context paragraph and was followed by an ellipsis sentence and one additional sentence.

Experiment 9 Method

Stimuli

In Experiment 9, the quantified sentence was presented in its biasing context and was followed by a VP-ellipsis sentence with a definite NP subject, as illustrated above in (107)–(108). An additional sentence followed the ellipsis sentence. The only difference between the stimuli and those in Experiment 2 was the inclusion of the ellipsis sentence and the final sentence. GSE predicts more surface-scope responses than in Experiment 2 in the ambiguous conditions since the ellipsis sentence should disambiguate the quantified sentence to the surface-scope interpretation. According to PSE, the results of the two questionnaires should be similar, since there is no impetus to revise an inverse-scope interpretive decision.

The items were arranged in pseudo-random order, such that no two items of the same condition appeared adjacent to each other. Included in the questionnaire were 24 filler items from different experiments that included sentences with a different kind of ambiguity. Each paragraph was followed by a comprehension question and two possible answers, as shown in example (116).

(116) Once a month, the city tested its emergency alert systems. The regular schedule for the air-raid sirens was to test them on the first of the month. At exactly ten in the morning on that day, an employee sounded every siren.

How many employees sounded sirens?

A. One. B. Several.
The order of presentation of the possible answers was counterbalanced across the items. I took the answer *One* to indicate that the perceiver had assigned a surface-scope interpretation to the quantified sentence, while the answer *Several* indicated an inverse-scope interpretation. This strategy provides a conservative measure of the true number of inverse-scope LFs constructed, since the strict interpretation of the inverse-scope configuration (in (116), roughly, *for every siren, some employee sounded it*) is consistent with a scenario where every siren is sounded by the same employee or the scenario where several employees sound sirens. While it is possible, then, that some portion of the *One* responses represents cases where perceivers had constructed an inverse-scope LF, the *Several* responses necessarily represent inverse-scope LFs. Since the predictions of the two principles diverge precisely on this issue of the number of inverse-scope responses, it seems reasonable to use this conservative measure which averts the danger of overestimating the number of inverse-scope interpretations.

In the unambiguous surface-scope condition, the comprehension question was not a *how many* question, but rather a question asking for information about the story told in the paragraph, as shown in example (117). One of the two answer choices was correct and the other was incorrect.

(117) Once a month, the city tested its emergency alert systems. The regular schedule for the air-raid sirens was to test them on the first of the month. At exactly ten in the morning on that day, the safety officer sounded every siren.

*Who sounded sirens?*

A. The mayor. B. The safety officer.

In this condition, then, the question response did not distinguish between two possible interpretations, but simply indicated whether perceivers had understood the paragraph.

*Procedure*

The stimuli were printed in paper questionnaires several pages long. Participants were instructed to circle the answer that matched their first interpretation of the sentence. Participants had as much time as they needed to complete the questionnaire, but were
asked to complete it without deliberating very long over the answers. They typically took no more than twenty minutes to complete the questionnaire.

Participants

The participants were 38 Northwestern University undergraduates with an age range of 18 to 36 years who received course credit for their participation. Nine participants were males and 29 were females. The data from two participants were excluded from the analysis: one who did not complete the entire questionnaire, and one who was not a native speaker of English.

Experiment 9 Results

As shown below (Table 9), participants nearly always chose the expected response in the unambiguous conditions. In the ambiguous conditions, perceivers chose the response supported by the context about two-thirds of the time: 66% surface-scope responses in the surface-scope context, and 67% inverse-scope response in the inverse-scope context, in spite of the presence of the ellipsis sentence.
Table 9.

Experiment 9 responses to comprehension question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response One</th>
<th>Inverse-scope response Several</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.66</td>
<td>0.34</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.07</td>
<td>0.93</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.33</td>
<td>0.67</td>
</tr>
<tr>
<td>Correct response</td>
<td>0.95</td>
<td>0.05</td>
</tr>
<tr>
<td>Incorrect response</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparing the two ambiguous conditions, the effect of the context was significant ($t_{1}(70)=6.3$, $p<0.0001$; $t_{2}(46)=4.2$, $p=0.0001$). Perceivers chose the response *One* more often in the surface-scope context and the response *Several* more often in the inverse-scope context. Comparing the two conditions with the inverse-scope context, the ambiguity of the quantified sentence had a significant effect ($t_{1}(70)=6.3$, $p<0.0001$; $t_{2}(46)=5.2$, $p<0.0001$). Perceivers chose the response *Several* more often when the quantified sentence was unambiguous inverse scope than when it was ambiguous.

**Experiment 9 Discussion**

The results of Experiment 9 are very similar to those of Experiment 2, which is presented in Chapter 2. It is noteworthy that the presence of the ellipsis sentence did not
have a strong disambiguating effect. GSE predicted that the ellipsis sentence, which has only the surface-scope interpretation, would eliminate the inverse-scope interpretation in the ambiguous conditions. However, after reading the ellipsis sentence perceivers gave the inverse-scope response to 67% of ambiguous, inverse-biased items, even more frequently than in Experiment 2 where the ellipsis sentence was not present.

In the unambiguous inverse-scope condition, it is possible that subjects simply ignored the ungrammatical ellipsis sentence in answering the comprehension question, since there was no alternative interpretation of the quantified sentence. However, in the ambiguous inverse-biased condition, it is surprising that subjects chose the surface-scope interpretation – the only possible interpretation according to GSE – only 33% of the time, less frequently than in this condition in Experiment 2, where no ellipsis sentence was present to disambiguate.

One possible explanation for this result could be that participants simply skipped over the ellipsis sentence and the final sentence and answered the comprehension question immediately after reading the quantified sentence, since the question asked only about the interpretation of the quantified sentence. Some support for this explanation can be found in the surface-scope biased ambiguous condition, where perceivers chose the surface-scope interpretation only 66% of the time. In this condition, the context supported surface scope and the ellipsis sentence was expected to require surface scope, yet in one-third of cases the inverse-scope interpretation was chosen. This fact is surprising under GSE if perceivers were reading and understanding the ellipsis sentence, but is less surprising if perceivers simply stopped reading after the quantified sentence, leaving the ambiguity unresolved, and answered the comprehension question at that point. If participants were indeed ignoring the ellipsis sentence, then the results of Experiment 9 tell us no more than Experiment 2 did. But if we assume that participants were cooperative and completed the task in the intended way, then Experiment 9 suggests that the VP-ellipsis sentence is not as strongly disambiguating as GSE predicts, and the results thus represent some tentative support for PSE.

Self-paced Reading Experiment Design

Experiments 10–13 were self-paced reading studies in which participants read short paragraphs containing quantified sentences, like the examples shown in (111)–(114) above. The self-paced reading task is useful because it provides a fairly direct means of observing relative processing difficulty. If a perceiver’s reading time is slower than average for a
particular region, this indicates that the perceiver is having difficulty in parsing or comprehending that segment of the sentence or discourse. A further methodological advantage of reading time measures is that they are less subject to a perceiver’s conscious control than introspective grammaticality judgments, and might therefore provide a more realistic indicator of perceivers’ implicit linguistic knowledge than conscious judgments.

The stimulus items in Experiments 10–13 consisted of 24 paragraphs containing quantified and ellipsis sentences. Each item appeared in one of four conditions in a 2\(\times\)2 design with two types of supporting context and two levels of ambiguity. The complete paradigm of one example stimulus item is shown above in (111)–(114). The entire set of stimuli is provided in the Appendix.

Participants were assigned to counterbalancing conditions in a Latin Square design, such that each participant saw six items in each of the four conditions. No participant saw any item in more than one condition.

The experimental task involved a moving-window presentation of the paragraphs, with a window size several words long, making up a single phrase or clause. Participants pressed a key to reveal each region of text and conceal the previous region. The quantified sentence and the ellipsis sentence each appeared in a single presentation region. Following each self-paced reading item, participants answered a comprehension question like the one shown in (118).

(118) How many climbers scaled cliffs?

One. Several.

The answer One corresponded to the surface-scope interpretation to the quantified sentence and the answer Several to the inverse-scope interpretation.

In the unambiguous surface-scope condition, the comprehension question was not a how many question, but rather a question asking for information about the story told in the paragraph, as shown in (119). One of the two answer choices was correct and the other was incorrect.
(119) Who scaled the cliffs?

The climbing expert. The referee.

In this condition, then, the question response did not distinguish between two possible interpretations, but simply indicated whether perceivers had understood the paragraph.

The design of the experiments provided three measures of perceivers’ performance: reading times for the quantified and ellipsis sentences, and responses to the comprehension questions. The GSE and PSE principles make different predictions about the three measures.

The unambiguous surface-scope condition provided the baseline measures. In this condition, the context supported the surface-scope interpretation, the quantified sentence was unambiguously surface-scope, and the ellipsis sentence was compatible with surface scope. There is no reason to predict processing difficulty, so the reading times for the two critical regions should fast compared to the other conditions. If participants are attending to the stimuli, then the answers to the comprehension question should be correct in most instances.

In the ambiguous, surface-scope biased condition, the two principles predict that performance will be the same as the baseline for the ellipsis sentence and the question response. Because the ellipsis sentence is compatible with the surface-scope interpretation, no processing difficulty is predicted. And since the surface-scope interpretation is supported by the context and either preferred by the processor (under PSE) or required by the presence of the ellipsis sentence (under GSE), perceivers should choose the surface-scope response One to the comprehension question. Since the quantified sentence is ambiguous, it is possible that its reading time might be slower than the baseline if the processor delays in assigning an interpretation, or if the two scope representations compete with each other in parallel in the comprehension process. Such a delay could give rise to processing difficulty and therefore to slower reading times for the quantified sentence as a result of the burden of carrying unstructured material (Frazier & Rayner, 1987) or as a result of competition between the two possible representations (Gibson, 1991; Gorrell, 1987; Kurtzman, 1985; Spivey & Tanenhaus, 1998).

In the unambiguous inverse-scope condition, GSE makes no specific prediction about the quantified sentence, but PSE predicts a cost at the quantified sentence even
though it is unambiguous, since the inverse-scope configuration is dispreferred by the processor. Both principles predict a processing cost at the ellipsis sentence in this condition, but for different reasons. Under GSE the ellipsis sentence is ungrammatical with an inverse-scope quantified sentence and therefore triggers attempted reanalysis of the quantified sentence to a surface-scope configuration, consuming processing resources. Under PSE, Parallelism requires the processor to construct an inverse-scope representation for the ellipsis sentence following an inverse-scope quantified sentence, and computing this complex configuration consumes processing resources. As for the comprehension question, since the quantified sentence is unambiguously inverse scope, then perceivers will probably give the inverse-scope response Several. GSE might, however, predict that perceivers would occasionally choose the response One because of the ungrammaticality of the inverse-scope configuration.

Like the ambiguous surface-scope biased condition, the ambiguous inverse-scope biased condition might show evidence of processing difficulty at the quantified sentence if the processor delays in assigning an interpretation to the ambiguous sentence, as I suggested above. PSE predicts slow reading times for the quantified sentence since there is an inherent cost to assigning inverse scope (although we might expect the inverse-scope supporting context to mitigate this cost (Altmann & Steedman, 1988; Crain & Steedman, 1985)). In any case, if perceivers do indeed commit to the inverse-scope interpretation of the quantified sentence, both hypotheses predict processing difficulty at the following ellipsis sentence, either because it is ungrammatical and triggers costly revision (SGH) or because giving it an inverse-scope configuration is dispreferred (SPH).

The predictions diverge for the responses to the comprehension questions in this condition. GSE predicts that the ellipsis sentence will trigger reanalysis of the quantified sentence, with the result that it has the surface-scope configuration and interpretation. Perceivers should therefore choose the surface-scope response One on most trials in this condition. PSE, on the other hand, predicts no such reanalysis. Once the processor has assigned inverse scope to the quantified sentence, that interpretation is maintained, and perceivers should choose the inverse-scope response Several on as many trials as in the baseline Experiment 2.

The following table (Table 10) summarizes the reading time and question response predictions for the four conditions. Notice that the predictions of the two hypotheses diverge for the responses to the comprehension question in the ambiguous inverse-scope condition. GSE predicts no difference between the two ambiguous conditions in the comprehension question responses, regardless of the supporting context, since the ellipsis
sentence should trigger reanalysis to surface scope in every case. In contrast, PSE makes the positive prediction of an effect of context on the comprehension question responses for ambiguous items. Because both interpretations are allowed under PSE, the comprehension question responses should reflect the interpretation that perceivers assigned based on the supporting context.

Table 10.

Predicted Results for Self-Paced Reading Experiments

<table>
<thead>
<tr>
<th>Condition</th>
<th>Quantified sentence prediction</th>
<th>Ellipsis sentence prediction</th>
<th>Question response prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GSE</td>
<td>PSE</td>
<td>GSE</td>
</tr>
<tr>
<td>Unambiguous surface scope</td>
<td></td>
<td>fast</td>
<td>fast</td>
</tr>
<tr>
<td>Ambiguous surface scope supported</td>
<td></td>
<td>fast</td>
<td>fast</td>
</tr>
<tr>
<td>Unambiguous inverse scope</td>
<td></td>
<td>slow</td>
<td>slow</td>
</tr>
<tr>
<td>Ambiguous inverse scope supported</td>
<td></td>
<td>slow</td>
<td>slow</td>
</tr>
</tbody>
</table>

Experiment 7

In Chapter 2 I presented a subset of the results of Experiment 7, namely the reading times for the quantified sentence and the comprehension question data. I present those results again here for convenience, and also include the reading time data for the ellipsis sentence. The Procedure for Experiment 7 is described in Chapter 2.
Experiment 7 Results

The results reported here are residual reading times, calculated using a deviation from regression algorithm (Ferreira & Clifton, 1986). Reading times for each participant were regressed on the length in characters of each segment that the participant read. The reading time residuals provide individually corrected estimates of participants’ expected reading times for the regions of differing lengths. Thus, for a given region, zero represents the expected reading time for a segment of that length in characters, averaged across participants. Negative numbers represent faster reading times than expected and positive numbers represent slower reading times than expected for a region of that length.

The residual reading times for the critical regions across the four conditions are depicted in Figure 11. All reading times are reported in milliseconds and are averaged over subjects and items.
Figure 11. Mean residual reading time for quantified sentence, Experiment 7.
The reading times for the quantified sentence showed a significant main effect of context \((F_1(1,23)=18, p<0.001; F_2(1,23)=17, p<0.001)\). Sentences in the inverse-scope conditions (both ambiguous and unambiguous) were read more slowly than those in the surface-scope conditions. There was no evidence for an effect of ambiguity \((F_1(1,23)=1.4, \text{n.s.}; F_2(1,23)=1.6, \text{n.s.})\) or for an interaction between the two factors \((F_1, F_2<1)\).

This pattern of reading times coincides with the predictions of PSE described above (and is not inconsistent with GSE). The quantified sentence was read quickly in the surface-scope conditions and slowly in the inverse-scope conditions, even when it was unambiguous inverse scope. This result is consistent with the results of Experiments 2 and 4 reported in Chapter 2.

Figure 12 depicts the reading times for the ellipsis sentence.
Figure 12. Mean residual reading time for ellipsis sentence, Experiment 7.
The reading times for the ellipsis sentence also showed a significant main effect of context: the ellipsis sentence was read more slowly in the two inverse-scope conditions than in the surface-scope conditions ($F_{1}(1,23)=20, p<0.001$; $F_{2}(1,23)=12, p<0.01$). In spite of the large numerical difference between the unambiguous and ambiguous conditions in the inverse-scope context, there was little statistical evidence for an effect of ambiguity ($F_{1}(1,23)=3.2, p=0.09$; $F_{2}(1,23)=2.3, p=0.15$) and no evidence of an interaction between context and ambiguity ($F_{1}(1,23)=1.2, n.s.; F_{2}(1,23)=2.2, n.s.$). Comparing the two inverse-scope conditions, there was no evidence for an effect of ambiguity ($t_{1}(46)=1.5, p=0.14$; $t_{2}(46)=1.5, p=0.13$). On the other hand, the evidence for the effect of the supporting context on the reading times is clear. The predictions of both hypotheses for the ellipsis sentence are thus confirmed by these results.

The proportions of responses to the comprehension questions are given below (Table 11). Comparing the two ambiguous conditions, there was a significant effect of biasing context on the proportion of surface-scope responses chosen ($t_{1}(46)=3.5, p=0.001$; $t_{2}(46)=3.3, p=0.002$). Participants chose the response *One* more frequently following the surface-scope context than following the inverse-scope context. Comparing the two conditions with the inverse-scope supporting context, there was a significant effect of ambiguity ($t_{1}(46)=6.1, p<0.0001$; $t_{2}(46)=6.2, p<0.0001$). Participants chose the surface-scope response *One* more often when the quantified sentence was ambiguous than when it was unambiguous inverse-scope.
Table 11. Experiment 7 responses to comprehension question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response</th>
<th>Inverse-scope response</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.68</td>
<td>0.32</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.08</td>
<td>0.92</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.44</td>
<td>0.56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correct response</th>
<th>Incorrect response</th>
</tr>
</thead>
<tbody>
<tr>
<td>unambiguous surface scope</td>
<td>0.95</td>
</tr>
</tbody>
</table>

As predicted, in the unambiguous conditions participants nearly always selected the expected answer: 95% of unambiguous surface-scope items received the correct response, and 92% of unambiguous inverse-scope items received an inverse-scope response. In contrast, both surface- and inverse-scope responses were selected quite frequently in the two ambiguous conditions. Although GSE predicted that the inverse-scope interpretation of the quantified sentence would no longer be available after perceivers had read the ellipsis sentence, perceivers chose the inverse-scope response one-third of the time in the surface-scope supporting condition and more than half the time in the inverse-scope supporting condition - as often as they did in Experiment 2 where the stimuli did not include the ellipsis sentences. The proportions of responses to the comprehension questions thus support the PSE principle rather than GSE.
The results of Experiment 7 exhibit three important findings. First, as I argued in Chapter 2, the reading times for the quantified sentence revealed that assigning the inverse-scope interpretation incurs a processing cost even for an unambiguous sentence. I proposed that the linguistic representation corresponding to the inverse-scope interpretation is more complex than the representation that gives rise to the surface-scope interpretation. Second, the ellipsis sentence reading time data confirmed the predictions of both GSE and PSE by revealing processing difficulty when the quantified sentence was in an inverse scope context. However, the third finding, that perceivers frequently selected the inverse-scope interpretation in both conditions where the quantified sentence was ambiguous, supports the predictions of Processing Scope Economy and not Grammatical Scope Economy.

I shall consider first the reading time results for the ellipsis sentence, which are consistent with the predictions of both principles. According to PSE, the surface-scope configuration is preferred for the ellipsis sentence, since this is the simplest configuration necessary to produce a correct interpretation. The grammar requires that the LF of the ellipsis sentence be parallel with that of its antecedent, the quantified sentence. If the processor has already assigned inverse scope to the quantified sentence, then it must also assign inverse scope to the ellipsis sentence to satisfy the grammatical requirement for Parallelism. Since this configuration is more complex than the surface-scope configuration, it is dispreferred and requires more processing resources to construct. The processing cost shows up as slower reading times for the ellipsis sentence in the inverse-scope supporting context than in the surface-scope context.

Of course, the reading time results for the ellipsis sentence are also consistent with GSE, where the restriction on additional QR operations is localized in the grammar rather than the processor. The ellipsis sentence is simply ungrammatical if the quantified sentence has inverse scope, since there is no LF for the ellipsis sentence that satisfies both Parallelism and Grammatical Scope Economy. The processor must repair the analysis of the quantified sentence in order to construct a grammatical parse for the ellipsis sentence. We know from the literature (Ferreira & Henderson, 1991; Frazier, 1987; Frazier & Rayner, 1982; Meng & Bader, 2000) that the processor has difficulty when it encounters a local ungrammaticality that prompts it to revise its analysis of the current sentence. We also know that the processor responds to a global ungrammaticality in the same way as to a local ungrammaticality (Hopf, Bader, Meng, & Bayer, 2003; Kaan & Swaab, 2003). It would make sense, therefore, for the processor to have comparable difficulty upon
encountering an ungrammaticality in a sentence that is dependent on the previous one. The slow reading times observed for the ellipsis sentence in inverse scope contexts are thus expected if Scope Economy is a principle of the grammar.

There is nevertheless an alternative to both hypotheses that could account equally well for the reading time results. It is possible that the slower reading times in inverse-scope conditions resulted not from the cost of constructing dispreferred structure, nor from the reanalysis of the quantified sentence, but simply from the implausibility of the ellipsis sentence with respect to the preceding context. For example, the quantified sentence (a) in paragraph (120) is entirely plausible under the inverse-scope interpretation where there are multiple paratroopers each jumping from a different plane.

(120) One squadron of the Airborne Regiment was taking part in a training exercise. Their assignment was to secure a large area so that supplies could be delivered there. As the airplanes approached the target area, the soldiers prepared for their jump. Upon the signal,
   a. a paratrooper jumped from every plane.
   b. The squadron sergeant did, too.

It is decidedly less believable that the squadron sergeant jumped from multiple planes at once (as in sentence (b)). If the perceiver had assigned the inverse-scope interpretation to the quantified sentence in this context, the implausibility of the ellipsis sentence would be likely to cause processing difficulty (Pickering & Traxler, 1998; Speer & Clifton, 1998; Traxler & Pickering, 1996; Traxler, Pickering, & Clifton, 1998). If implausibility is the source of the slow reading times observed at the ellipsis sentence, then a relatively plausible ellipsis sentence should not cause the same degree of processing difficulty, even if it is ungrammatical (SGH) or dispreferred (SPH). The data from Experiment 7 do not make it possible to distinguish between the plausibility account and either the PSE or GSE account, but I report below the results of a questionnaire (Experiment 10) that collects plausibility ratings for the ellipsis sentence of each stimulus item in the two different contexts. The results of Experiment 10 suggest that the reading times were not affected by the relative plausibility of the ellipsis sentences.

The responses to the comprehension questions help to clarify some of the issues arising from the ellipsis sentence results. Recall that participants were presented with a question like (121) and two possible answers.
How many climbers scaled cliffs?

One. Several.

The answer One corresponds to the surface-scope interpretation and Several to the inverse-scope interpretation. (In the unambiguous surface-scope condition, the two possible answers were a correct and an incorrect one.) As I mentioned above, the number of Several responses represents a conservative measure of the number of inverse-scope interpretations assigned for two reasons:

First, the surface-scope interpretation $\square > \square$ (There exists a climber who scaled all the cliffs.) entails the truth of the inverse-scope interpretation $\square > \square$ (For every cliff, some climber scaled it.). In other words, the inverse-scope configuration has a more vague interpretation: it is true either if there were several climbers or if there was only one climber, as long as all the cliffs were scaled. Thus, the response One could indicate that the processor has constructed either the surface-scope configuration or an inverse-scope configuration with the interpretation that every cliff was scaled by the same climber. In contrast, the response Several necessarily indicates that the processor has constructed an inverse-scope configuration.

Second, Danny Fox (p.c., 3 January 2004) suggests that the answer Several is an infelicitous response to a how many question, which expects a definite answer. For this reason, perceivers might have had a bias toward choosing the more felicitous, definite response One, regardless of what interpretation they actually assigned to the quantified sentence. On the other hand, it is possible that the nature of the forced-choice task would have alleviated the tendency to disfavour the Several response, since participants might have been reluctant to choose the response One on trial after trial. In any case, it is possible that the comprehension question data underestimate the true number of cases where perceivers had constructed an inverse-scope configuration, but unlikely that they overestimate it. Since the most important finding from the responses to the comprehension questions was an unexpectedly high number of inverse-scope responses, a potential underestimate is not a serious concern.

In examining the question response data, we can observe first that perceivers selected the correct response on 95% of trials in the unambiguous surface-scope condition. This result confirms that participants were indeed performing the task in the expected way by answering the question correctly with respect to the context provided in the paragraph.
In the unambiguous inverse-scope condition, perceivers chose the inverse-scope response on 92% of trials, even though the combination of the ellipsis sentence with the inverse-scope quantified sentence was predicted to be ungrammatical by GSE. This result is not entirely surprising, since there is no alternative interpretation for the unambiguous quantified sentence.

The question response data in the ambiguous conditions are more telling. Although GSE predicted that the ellipsis sentence would eliminate the inverse-scope interpretation by triggering reanalysis of the quantified sentence to surface scope, perceivers still chose the inverse-scope response for nearly one third (32%) of surface-scope trials and more than half (56%) of inverse-scope trials. The proportions of surface- and inverse-scope responses are remarkably similar to those observed in Experiment 2, where perceivers chose the inverse-scope response for 19% of surface-scope trials and 53% of inverse-scope trials. This fact is all the more remarkable considering that Experiment 2 presented the ambiguous quantified sentences in context but without the following ellipsis sentence. GSE predicted that perceivers in the self-paced reading task, having read the disambiguating ellipsis sentence, would choose the inverse-scope response less frequently than in the questionnaire, but in Experiment 7 they actually chose it slightly more frequently. Although the ellipsis sentence caused processing difficulty for perceivers, it did not have the effect predicted by GSE of eliminating the inverse-scope interpretation of the quantified sentence.

According to PSE, however, the result is not surprising. Although inverse scope is the dispreferred configuration, it is possible for the processor to construct it where it is necessary for achieving the desired interpretation. Once the processor has constructed an inverse-scope configuration for the quantified sentence, then it must go on to assign inverse scope to the ellipsis sentence in order to satisfy Parallelism. The inverse-scope interpretation is thus still available when perceivers come to the end of the paragraph and answer the comprehension question. Under PSE, the similar proportion of responses in the ambiguous inverse-scope condition between Experiment 7 and Experiment 2 is not surprising but expected.

Nevertheless, there are other possible explanations for the high proportion of inverse-scope responses in the ambiguous conditions. The result could be attributed either to some element of the interpretive process, or to some aspect of the grammar of quantification and ellipsis. From a real-time point of view, the processor might simply have failed for some reason to reanalyze its original decision about the structure and interpretation of the quantified sentence, with the result that the inverse-scope
interpretation was retained even if it was no longer grammatical. On the other hand, the combination of an inverse-scope quantified sentence and an ellipsis sentence with a definite NP subject in these items might actually be grammatical, either because the ellipsis sentence was not actually scopally commutative, or because Grammatical Scope Economy does not constrain LF configurations in the way that Fox (2000) proposes.

It is possible that the 56% proportion of inverse-scope results simply indicates chance performance by the participants. Given the somewhat weak effect of the inverse-scope supporting context observed in Experiment 2, where perceivers chose the inverse-scope response on only 53% of trials, it is possible that the context is simply not very strongly biasing. Perhaps perceivers were confused by the combination of the doubly quantified sentence and ellipsis sentence and simply guessed at the comprehension question, leading to a chance rate of responses. If this is the case, then the comprehension question results do not support either hypothesis. But it is exactly the comparison with the results of Experiment 2 that makes the present results so compelling. GSE predicts that the ellipsis sentence should eliminate the inverse-scope interpretation altogether, leading to 0% Several responses. Instead, perceivers chose the inverse-scope response just as often as they did in Experiment 2. The inverse-scope context was strong enough to push perceivers to assign inverse scope more often than they did in Experiment 1, where there was no context (19% Several responses) and more often than the surface-scope context did (32% Several responses). The presence of the ellipsis sentence was not sufficient to reduce the proportions of inverse-scope responses as compared with Experiment 2, contrary to the predictions of GSE.

On the other hand, it is possible that the ellipsis sentence did indeed render the inverse-scope interpretation ungrammatical, but that the processor retained this interpretation regardless. The sentence processing literature includes reports of perceivers confidently retaining their initial interpretations of a temporarily ambiguous string in spite of its subsequent ungrammaticality, such as the cases described by Ferreira and her colleagues (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, Christianson, & Hollingworth, 2001). In those experiments, perceivers initially misanalyzed intransitive and reflexive verbs as transitive. In sentence (122), for example, the baby was initially parsed as the object of dressed rather than the subject of spit up.

(122) While Anna dressed the baby spit up on the bed.

Although there is no grammatical analysis of this sentence that includes the baby as the object of dressed, when participants were questioned after reading the entire sentence, they
frequently answered “yes” to the question “Did Anna dress the baby?” and gave high ratings for their confidence in the correctness of their answers. These two studies showed that it is possible for the processor to make and commit to interpretive decisions based on incompletely or incorrectly analyzed linguistic material.

The present situation might well be analogous to that reported by Christianson et al. and Ferreira et al. Perhaps the processor recognizes the ungrammaticality of its inverse-scope commitment when it encounters the ellipsis sentence, giving rise to the slower reading times for this region, yet still maintain the interpretive commitment that it has made for the quantified sentence. In this scenario perceivers choose the inverse-scope response to the comprehension question in spite of its ungrammaticality.

If the initial commitment to the inverse-scope interpretation of the quantified sentence is so strong that subsequent evidence of ungrammaticality fails to overturn it, we might expect that other kinds of subsequent conflicting evidence would be equally unsuccessful at triggering reanalysis. However, Experiment 5A (described in Chapter 2) provided evidence against this prediction. In that experiment, ambiguous quantified sentences were presented following either the surface-scope or inverse-scope context. The subsequent sentence did not contain ellipsis, but had either a singular or a plural definite NP subject, as shown in (124).

(123)  A climber scaled every cliff.

(124)  a. The climber was very skilled.
       b. The climbers were very skilled.

In the inverse-scope context, perceivers chose the surface-scope response 70% of the time after reading the singular sentence (124)a, compared with 44% of the time after reading the ellipsis sentence in Experiment 7. Clearly, the non-elided sentences with singular subjects in Experiment 5A were more effective at disambiguating to surface scope than the VP-ellipsis sentences were in Experiment 7.

We might therefore be doubtful that the present high proportion of inverse-scope responses are a result of the processor’s tenacity in preserving an initial incorrect analysis. Let us assume that the inverse-scope context led perceivers to the same proportion of initial inverse-scope commitments for the interpretation of the quantified sentence in both experiments. In Experiment 5A, the disambiguating sentence contained a singular definite NP coreferential with the subject of the quantified sentence. If a perceiver had committed
to the inverse-scope interpretation of (123) which includes multiple climbers, then (124)a would violate the uniqueness presupposition associated with the singular definite (Russell, 1905; Strawson, 1950). This presupposition failure apparently signals the processor to overturn its inverse-scope interpretation, resulting in a 70% rate of surface-scope responses. The ellipsis sentence, on the other hand, resulted in only 44% surface-scope responses. If the pragmatic violation of a singular definite serves as a sufficiently strong cue for the processor to overturn an inverse-scope commitment, why would the grammaticality violation of an ellipsis sentence not be strong enough to trigger reanalysis of the quantified sentence? The difference in the proportion of responses between the two experiments suggests that the processor’s tenacity is not solely responsible for its failure to reanalyze to surface scope in Experiment 7. Rather there is no evidence of reanalysis because there is no ungrammaticality to trigger reanalysis. As PSE predicts, an inverse-scope ellipsis sentence is dispreferred but grammatical. The ellipsis sentence is not a disambiguator, unlike the singular and plural definite sentences in (124).

There is another possible explanation for the persistence of the inverse-scope interpretation in spite of its predicted ungrammaticality under GSE. Perhaps the quantified sentence was simply inaccessible to revision by the time the ellipsis sentence was parsed because of the sentence boundary between the sentences. It is well established that the processor carries out many of the tasks of comprehending a sentence once it gets to the sentence’s end. According to Just and Carpenter, “[t]he processes that occur during sentence wrap-up involve a search for referents that have not been assigned, the construction of interclause relations … and an attempt to handle any inconsistencies” (1980, p. 345). Once the wrap-up is complete, the syntactic representation of a sentence decays (Jarvella, 1971; Lombardi & Potter, 1992; Potter & Lombardi, 1998; Sachs, 1967). In the present situation, then, it is possible that the processor recognizes the ungrammaticality of the inverse-scope configuration upon parsing the ellipsis sentence, as predicted by GSE, is prompted to revise the quantified sentence, but is unable to reanalyze the quantified sentence since its representation has already decayed from memory. If this is the case, then we should expect that the quantified sentence would be more available for revision if the disambiguating information were encountered within the same sentence. Experiment 12 tests this possibility by conjoining the quantified sentence and ellipsis sentence within a single sentence, as shown in as shown in (125).

(125)  An experienced climber scaled every cliff, and the instructor did, too.
If the ungrammaticality of the elided constituent prompts reanalysis, as GSE predicts, then the quantified clause should still be accessible in (125) since the disambiguating information is received while the sentence is still being parsed. If the sentence boundary between the quantified and ellipsis sentences is responsible for the high proportion of inverse-scope responses in Experiment 7, then we should observe fewer inverse-scope responses in Experiment 12.

The high number of inverse-scope responses to the comprehension question in Experiment 7 might also be explained by appealing to the grammar rather than to the workings of the processor. It is possible that the combination of an inverse-scope quantified sentence with an ellipsis sentence is actually grammatical in these items, perhaps because of some overlooked property of the sentences that allows them to escape the restrictions imposed by GSE. Recall that Grammatical Scope Economy disallows an inverse-scope configuration of quantifiers if the truth conditions are the same as in the surface-scope configuration. Fox (1995) provided some examples of cases where the inverse-scope interpretation appears to be possible even though GSE would be expected to constrain the second conjunct, citing the example from Lappin (1993) shown in (126).

\[(126) \quad \text{At least one Labour MP attended every committee meeting, and Bill did too.}\]

Intuition suggests that sentence (126) is quite acceptable on the interpretation where several MPs each attended a different meeting and Bill attended all the meetings. This intuition is surprising on Fox's account since GSE should prohibit inverse scope in the second conjunct, with the result that Parallelism prohibits inverse scope in the first conjunct.

Fox proposed that GSE does not operate on sentence (126) because the sentence contains a stage-level predicate rather than an individual-level predicate (Kratzer, 1995). A stage-level predicate introduces an event argument that is implicitly quantified over (Parsons, 1990). Thus, the ellipsis clause in (126) is actually doubly-quantified, and GSE does not prevent the QR operation that moves every meeting above Bill.

Since all the stimuli in Experiment 7 used eventive verbs, each one contained an implicit existential quantifier over events (EQE). It is therefore possible that Grammatical Scope Economy did not apply to these items, with the result that the inverse-scope responses are perfectly grammatical. If the high proportion of inverse-scope responses in Experiment 7 is a consequence of the eventive predicates in the stimulus items, then we
should observe the predicted disambiguating effect of VP-ellipsis in sentences with individual-level predicates such as (127), which do not contain event quantifiers.

(127)  A student knows every state capital.
       The teacher does too.

The pilot Experiment 11 described below used four stimulus items containing individual-level predicates like (127). However, the results of the pilot suggest that the ellipsis sentence did not eliminate the inverse-scope interpretation even in items with individual-level predicates.

I have suggested above several possibilities that could account for the results of Experiment 7 while still retaining GSE as a principle of the grammar. Nevertheless, the results quite straightforwardly support the predictions of the PSE principle. In an inverse-scope context, the processor must construct an inverse-scope configuration for the quantified sentence to provide the correct interpretation. According to PSE, this parse is relatively costly because the inverse-scope configuration is more complex and dispreferred. Next, in order to correctly parse the ellipsis sentence, the processor invokes Parallelism, constructing a second inverse-scope configuration. This parse is required by the grammar of VP-ellipsis, and is not ungrammatical since Scope Economy is not a principle of the grammar. The parse is, however, costly because the necessary inverse-scope structure is relatively complex and dispreferred by PSE. On this account, both the quantified sentence and the ellipsis sentence will exhibit slow reading times in inverse-scope contexts due to the complexity of the inverse-scope configurations. But because the inverse-scope configuration is permitted, the inverse-scope interpretation is available for the perceiver to choose in answering the comprehension question.

PSE also accounts for the troublesome counterexamples identified above like (128):

(128)  One cabinet minister bears responsibility for each government department, and the Prime Minister does too.

Under PSE, both conjuncts of the sentence can have inverse-scope configurations. In the first clause, the determiner each favours an inverse-scope representation (loup, 1975; Tunstall, 1998; VanLehn, 1978). This is dispreferred by PSE but allowed, introducing a processing cost. The second clause can have an inverse-scope configuration in order to satisfy Parallelism, and its truth conditions remain the same. Thus, each department is
under the guidance of a different minister, while still allowing the PM to have responsibility for all the departments.

One final potential account of the data remains. It is possible that Scope Economy is indeed a principle of the grammar, but that Parallelism is not a grammatical principle but a processing principle. If this were the case, then it would be possible, though dispreferred by the processor, for a quantified sentence to have an inverse-scope configuration while the following ellipsis sentence had a surface-scope configuration (or vice versa). In this way, the singly quantified ellipsis sentence would satisfy Grammatical Scope Economy, and the mismatching pair of sentences would cause processing difficulty because of their non-parallel configurations, which are dispreferred by Processing Parallelism. The inverse-scope response would be available because the quantified sentence retained its inverse-scope representation. An experiment to test this possibility would have to include pairs of sentences like (129)–(130):

(129)  A novice climber scaled every cliff.

(130)  An expert did, too.

For such sentences, Processing Scope Economy plus Grammatical Parallelism predict that both sentences must have the same configuration, whether surface scope or inverse scope. Assigning the inverse-scope interpretation to one should thus lead to a processing cost for both. In contrast, Grammatical Scope Economy plus Processing Parallelism predict that mismatched interpretations should be possible, though dispreferred. In this case, we should observe no processing difficulty at the indefinite ellipsis sentence following an inverse-scope quantified sentence because the ellipsis sentence does not violate Scope Economy. This experiment should be carried out in the future to test this possibility.

In summary, Experiment 7 revealed three important results:

- a doubly quantified sentence was read more slowly in an inverse-scope supporting context than in a surface-scope supporting context;
- a VP-ellipsis sentence was read more slowly following a quantified sentence in an inverse-scope supporting context than one in a surface-scope supporting context; and
• when the doubly quantified sentence was ambiguous, perceivers chose both
the inverse-scope and the surface-scope responses to a comprehension
question quite frequently, even after reading the ellipsis sentence.

I have considered here five potential explanations for these results:

• perceivers were responding at chance in the ambiguous inverse-scope
condition, a possibility that the comparison with the results of Experiments
1 and 2 makes unlikely;

• the ellipsis sentence is grammatical but implausible in an inverse-scope
context, a possibility examined below with Experiment 10;

• the ellipsis sentence is ungrammatical, as Fox's proposal claims, but the
ellipsis sentence fails to activate complete reanalysis of the quantified
sentence, a possibility that is tested by Experiments 4 and 5A;

• the ellipsis sentence is grammatical because it contains an implicit event
quantifier in addition to the universal quantifier, with the result that
Grammatical Scope Economy does not constrain its LF, a possibility
explored tentatively with Pilot Experiment 11; and

• the ellipsis sentence is grammatical because Scope Economy is not a
grammatical prohibition on QR but rather a general structural preference,
as predicted by PSE.

Each of these accounts makes concrete predictions that are tested further in the
experiments described below. The overall pattern of results presented thus far is most
consistent with the last of these accounts and the Processing Scope Economy principle,
suggesting that Scope Economy effects have their origin in the operation of the sentence
processing mechanism rather than in the principles of the grammar.

**Experiment 10**

I proposed above an alternative explanation for the significantly slower reading
times for the ellipsis sentence in the inverse-scope conditions in Experiment 7. I suggested
that the reading times might have been slower not because of the ungrammaticality of the
ellipsis sentence with respect to the inverse-scope quantified sentence, but because the
ellipsis sentence was implausible with respect to the context. For example, in both (131) and (132) below the ellipsis sentences are ungrammatical on the inverse-scope readings of their respective quantified sentences according to Grammatical Scope Economy, but the ellipsis sentence in (131) is considerably less plausible in its context than the one in (132) irrespective of grammaticality.

(131) One squadron of the Airborne Regiment was taking part in a training exercise. Their assignment was to secure a large area so that supplies could be delivered there. As the airplanes approached the target area, the soldiers prepared for their jump. Upon the signal,
   a. a paratrooper jumped from every plane.
   b. The squadron sergeant did, too.

(132) Mrs. Dunn’s third-grade class had to stay indoors at recess because of the bad weather. Mrs. Dunn put some puzzles on the blackboard for the children to solve. Because the puzzles were quite difficult, she assigned teams of students to the different puzzles. By the end of recess,
   a. a team had solved every puzzle.
   b. the teacher’s aide had, too.

In context (131), it is very unlikely that the sergeant jumped from all the planes “upon the signal”, but in context (132), it is entirely possible that the teacher’s aide solved all the puzzles during recess. The present concern, then, is that the differences in the reading times for the ellipsis sentence in Experiment 7 might have arisen not from the conflict between an inverse-scope quantified sentence and an ellipsis sentence that was grammatically required to have surface-scope, but because the events described in the inverse-scope quantified sentence with multiple actors were unlikely to have occurred with the single actor described in the ellipsis sentence.

In fact, the same concern holds for the reading times for the quantified sentence. If perceivers never entertained the inverse-scope interpretation, but instead attempted to assign the surface-scope interpretation in a context where a single actor was an unlikely subject, the implausibility of the surface-scope interpretation might have caused slow reading times for the quantified sentence. It is therefore crucial to obtain a measure of the plausibility of the sentence that describes the event in question as having only a single
actor, that is, a measure of the plausibility of a single actor performing, for example, multiple plane-jumping events or multiple puzzle-solving events. This section describes a questionnaire that was used to obtain this plausibility measure for the 24 stimulus items used in Experiment 7. If the difference in reading times observed in Experiment 7 is attributable to the relative plausibility of single-actor events in the two different contexts, then we should expect to find a correlation between the reading times and the plausibility ratings. In other words, greater implausibility should be correlated with slower reading times. In fact, no such correlation was found.

Experiment 10 Method

Stimuli

The 24 stimulus items were based on those used in self-paced reading Experiment 7. Each item contained the context sentences and a non-elided version of what was the ellipsis sentence in the self-paced reading stimulus. For example, for the stimulus item from Experiment 7 shown in (133) contained the quantified and ellipsis sentences following the context, the corresponding item in Experiment 10 is shown in (134).

(133) At the new hockey arena, the power kept failing in the locker rooms. The players persuaded the building manager to get someone to check the wiring for that section of the arena. Before the next hockey game, a qualified electrician tested every circuit. The building custodian did, too.

(134) At the new hockey arena, the power kept failing in the locker rooms. The players persuaded the building manager to get someone to check the wiring for that section of the arena. Before the next hockey game, the building custodian tested every circuit.

Because the quantified sentence was not included, there were only two versions of each item: one with the surface-scope context and one with the inverse-scope context. (Recall that in the self-paced reading experiment, the two additional conditions resulted from a manipulation of the quantified sentence, either ambiguous or unambiguous, following the same two contexts.)
Each item was followed by a dotted line marked with digits from one to five. The left end of the line was labelled *Implausible* and the right end *Plausible*. An example is shown in (135).

(135)  *Implausible*  1. . . . . 2. . . . . 3. . . . . 4. . . . . 5  *Plausible*

The 24 items in the two conditions were arranged in two lists, each of which contained half of the items from each condition. The items were arranged in pseudo-random order. There were no filler items in the list.

**Procedure**

Participants read an instruction sheet that included three example items with a similar format to the experimental items. They were instructed to rate how plausible the last sentence of the item was with respect to the rest of the paragraph. Participants could take as much time as they wanted to complete the questionnaire, but usually took no longer than 20 minutes.

**Participants**

Twenty-three Northwestern University students, six males and 17 females with an age range of 18 to 22 years, participated in the experiment for course credit. The data from five of the participants were not included in the analysis, since three had already participated in a related experiment and two were not native speakers of English.

**Experiment 10 Results**

The ratings for each item were recorded and the mean plausibility rating for each stimulus item calculated. The mean score for the surface-scope condition was 2.72 and for the inverse-scope condition was 2.50, a difference that is nearly significant ($t(23)=1.6$, $p=0.06$). It is not surprising that the plausibility scores were slightly lower for the inverse-scope condition, since the inverse-scope contexts supported a reading where multiple actors carry out the action, while the sentence that was rated had only a single actor as the subject.
The items were then arranged in rank-order according to their mean plausibility rating. Thus two ranked lists were created, one for surface-scope items and one for inverse-scope items. The item ranked 1 had been rated the most implausible and the item ranked 24 the most plausible. Least-squares regressions were performed of the residual reading times for the quantified and ellipsis sentences on the item ranks. Recall that the plausibility hypothesis predicted a negative correlation between reading times (where high numbers represent slow reading times) and plausibility scores (where a high rank represents high plausibility).

There was no evidence of a relationship between reading times and plausibility scores, as illustrated below (Table 12).

Table 12.

<table>
<thead>
<tr>
<th></th>
<th>surface scope</th>
<th>inverse scope</th>
<th>r²</th>
<th>F(1,22)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantified sentence</td>
<td></td>
<td></td>
<td>0.001</td>
<td>0.03</td>
<td>0.86</td>
</tr>
<tr>
<td>Ellipsis sentence</td>
<td></td>
<td></td>
<td>0.06</td>
<td>1.3</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
<td>3.3</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.03</td>
<td>0.65</td>
<td>0.43</td>
</tr>
</tbody>
</table>

**Experiment 10 Discussion**

If the slow reading times for the ellipsis sentence in the inverse-scope conditions observed in Experiment 7 had been due to the relative implausibility of the ellipsis sentence in that context, then we should have observed a correlation between the sentence’s plausibility and its reading time. The lack of such a correlation suggests that implausibility was not responsible for the slow reading times.

In the absence of a plausibility effect, it seems fairly safe to conclude that the reading times results of Experiment 7, which were consistent with the predictions of both
GSE and PSE, can be attributed to one of the two principles. According to GSE, the ellipsis sentence is ungrammatical with an inverse-scope quantified sentence, and the ungrammaticality prompts reanalysis of the quantified sentence. On this account the observed slow reading times are a symptom of the processing difficulty arising from the attempt to reanalyze the foregoing material. According to PSE, the ellipsis sentence can be assigned an inverse-scope interpretation, but to do so is dispreferred because of the relative complexity of the inverse-scope configuration. On this account the observed slow reading times represent the cost of constructing a dispreferred parse for the sentence.

Experiment 11

The pilot Experiment 11 was designed to test the possibility that the unexpected high number of inverse-scope responses in Experiment 7 was due to an implicit event quantifier that licensed QR of the direct object, thus escaping the restriction imposed by Grammatical Scope Economy. All the quantified and ellipsis sentences in Experiment 7 used eventive, stage-level predicates (Kratzer, 1995). According to Parsons (1990), an eventive predicate always contains an implicit existential quantifier over events (EQE). The result is that an apparently singly quantified sentence like (136):

\[(136)\text{ The shop owner scaled every cliff.}\]

actually contains two quantifiers, as shown in (137).

\[(137) \exists e \forall x. [\text{cliff}(x) \& \text{the shop owner scaled } x \text{ at } e]\]

Because (137) is doubly quantified, a different scope configuration results in different truth conditions\(^7\), as in (138):

\[(138) \forall x \exists e. [\text{cliff}(x) \& \text{the shop owner scaled } x \text{ at } e]\]

\(^7\) Actually, the option for more than one set of truth conditions depends on where the EQE originates. If it is generated below the subject, then the object quantified phrase will be allowed by GSE to move above the EQE, but not above the subject DP. Because it is not entirely clear where the EQE is generated, Experiment 11 tests individual-level predicates that do not contain event quantifiers.
Example (137) means, approximately, “There was some event of the shop owner scaling all the cliffs,” while (138) corresponds to, “For every cliff, there was some event of the shop owner scaling it.”

Since the two scope configurations have two different sets of truth conditions, Grammatical Scope Economy does not constrain the LF of a sentence that contains an EQE and another quantifier. The ellipsis sentence in an Experiment 7 item is thus permitted to have an inverse-scope configuration, with the result that the quantified sentence can have the inverse-scope interpretation. On this account, the inverse-scope interpretation is completely grammatical and the high proportion of inverse-scope responses in Experiment 7 is unsurprising.

In contrast to stage-level predicates, individual-level predicates do not contain event quantifiers. We should therefore expect Scope Economy to constrain singly quantified sentences with non-eventive predicates. Pilot Experiment 11 used four items analogous to the 24 items used in Experiment 7, but containing non-eventive individual-level predicates. If Scope Economy constrains the LFs of these sentences, then GSE predicts that the ungrammatical ellipsis sentence will trigger reanalysis to a grammatical surface-scope configuration, with the result that perceivers select a high proportion of surface-scope responses.

Experiment 11 Method

Stimuli

There were four stimulus items in Experiment 11. The target sentences (i.e. the quantified sentence and ellipsis sentence) in the stimuli contained individual-level predicates (e.g., loves, knows, speaks) rather than stage-level predicates (Kratzer, 1995). The four items were included in an experiment with 48 filler items that were used in a different experiment. The items were presented in the same four conditions as those in Experiment

---

8 Parsons suggested that stative predicates contain implicit quantifiers over states. If this were the case, then virtually every sentence with a single overt quantifier would actually be doubly quantified, since it would also contain an implicit existential quantifier over events or states. Since such a state of affairs would have the effect of making Scope Economy irrelevant to overtly quantified sentences, I shall not consider it any further here.
7: unambiguous surface scope, ambiguous surface scope, unambiguous inverse scope, and ambiguous inverse scope. The entire list of stimuli in the four conditions is provided in the Appendix.

Procedure

The procedure was identical to that used in Experiment 7.

Participants

Thirty-two Northwestern University undergraduates participated in Experiment 11, 11 males and 21 females. The data were analyzed from 24 of the participants. Of the eight participants whose data were not analyzed, four were not native speakers of English, three had already participated in a related experiment and had been debriefed, and one self-identified as having dyslexia.

Experiment 11 Results

The reading time results for Experiment 11 are depicted in Figure 13 and Figure 14.
Figure 13. Mean residual reading times for quantified sentence, Experiment 11.
Figure 14. Mean residual reading time for ellipsis sentence, Experiment 11.
The following table (Table 13) summarizes perceivers’ responses to the comprehension questions. Because there were only four items in this experiment, it was not possible to perform statistical analysis. Even a subject-by-subject inspection of the data does not allow for meaningful comparisons across the conditions, since each participant saw only one item in each condition. Visual inspection reveals that the pattern of results is similar to that observed in Experiment 7: participants read the ellipsis sentence more slowly in the inverse-scope contexts than in the surface-scope contexts, and selected a high proportion of inverse-scope responses to the comprehension question in both ambiguous conditions, even though GSE predicted that the inverse-scope interpretation would be eliminated by the presence of the ellipsis sentence.
Table 13.

Experiment 11 responses to comprehension question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response One</th>
<th>Inverse-scope response Several</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.13</td>
<td>0.87</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.54</td>
<td>0.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correct response</th>
<th>Incorrect response</th>
</tr>
</thead>
<tbody>
<tr>
<td>unambiguous surface scope</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Experiment 11 Discussion

While the results of the pilot Experiment 11 cannot be conclusive with such a small number of stimulus items, they are still suggestive. These stimuli all contained individual-level predicates, which are not eventive in nature. The ellipsis sentence in each item should therefore be genuinely singly quantified, and the well-formedness of the sentence should be conditioned by Grammatical Scope Economy. GSE predicts that the ellipsis sentence should trigger the processor to revise an inverse-scope quantified sentence and give it the surface-scope interpretation. We should therefore find a high proportion of surface-scope responses to the comprehension questions in these items.
However, even in these non-eventive items, participants chose the inverse-scope response about half the time in both ambiguous conditions. Clearly the inverse-scope interpretation is still available to perceivers, contrary to the prediction of GSE. The results of this pilot experiment suggest that we cannot dismiss the high proportion of inverse-scope responses observed in Experiment 7 by appealing to implicit event quantifiers contained in those items.

**Experiment 12**

I discussed above the possibility that the high proportion of inverse-scope responses to the comprehension question in the ambiguous conditions in Experiment 7 could be explained if the processor was unable to reanalyze the quantified sentence even though the ellipsis sentence prompted reanalysis. It is well known (Jarvella, 1971; Lombardi & Potter, 1992; Potter & Lombardi, 1998; Sachs, 1967) that the syntactic representation of a sentence decays from memory once the sentence’s end has been reached, since the processor engages in wrap-up inferencing at the end of the sentence (Just & Carpenter, 1980; Millis & Just, 1994). Based on the results of Experiment 5A reported in Chapter 2, it seems unlikely that the sentence boundary between the quantified sentence and the next sentence makes the quantified sentence completely inaccessible to reanalysis. In that experiment perceivers chose a surface-scope response 70% of the time when a non-elided sentence with a singular subject followed an ambiguous quantified sentence in the inverse-scope supporting context, as in (139), in contrast to only 44% of the time following the ellipsis sentence in the same condition in Experiment 7.

(139) A climber scaled every cliff.
The climber was very skilled.

Nevertheless, it is possible that reanalysis of the quantified sentence might be more easily accomplished by the processor if no sentence boundary intervened between it and the ellipsis sentence.

Experiment 12 tests this possibility by conjoining the doubly quantified sentences of Experiment 7 with VP-ellipsis constructions within the same sentence, as shown in as shown in (140).

(140) An experienced climber scaled every cliff, and
the instructor did, too.
Since the quantified sentence has not ended when the processor encounters the ellipsis sentence, the wrap-up inferencing presumably has not yet occurred. If, as GSE predicts, the ellipsis sentence prompts the processor to reanalyze an inverse-scope quantified sentence, this reanalysis might be possible as part of the inferencing that takes place during sentence wrap-up. If the sentence boundary between the quantified sentence and the ellipsis sentence in Experiment 7 was responsible for the high proportion of inverse-scope responses, then we should observe fewer inverse-scope responses in Experiment 12, since the processor will be able to access the representation of the quantified sentence in order to assign it surface scope.

Experiment 12 Method

Stimuli

The stimuli were the same as those used in Experiment 7 with the exception that the quantified sentence and ellipsis sentence were linked by the conjunction and, as illustrated above in example (140). The quantified clause and the conjunction and occupied one presentation region, while the ellipsis clause occupied the next region. Each item was followed by a comprehension question identical to that used in Experiment 7. The stimuli were arranged in a list including the same 24 filler items used in Experiment 7. The items were arranged in pseudo-random order, with no two items of the same condition adjacent to each other. Four versions of the list items were prepared as described above.

Procedure

The procedure was the same as that used for Experiment 7.

Participants

Thirty-five Northwestern University students, nine men and 26 women with an age range of 17 to 25 years, participated in the experiment for course credit. Data from seven participants was excluded from the analysis, since two had participated in a related experiment and five were not native speakers of English.
Experiment 12 Results

The residual reading times for the quantified and ellipsis clauses are depicted in Figure 15 and Figure 16.
Figure 15. Mean residual reading time for quantified clause, Experiment 12.
Figure 16. Mean residual reading time for ellipsis clause, Experiment 12.
The unambiguous inverse-scope quantified clause was read more slowly than in any other condition. There was a significant main effect of context ($F_{1}(1,27)=4.7$, $p=0.04$; $F_{2}(1,23)=8.5$, $p=0.008$) and a significant interaction between context and the level of ambiguity of the quantified clause ($F_{1}(1,27)=9.7$, $p=0.004$; $F_{2}(1,23)=8.4$, $p=0.008$). There was no evidence for a main effect of ambiguity on the reading times ($F_{1}(1,27)=2$, $p=0.17$; $F_{2}(1,23)=2.5$, $p=0.12$). This pattern of results is somewhat different from that observed in Experiment 7, where the quantified sentence was read more slowly in both inverse scope conditions, unambiguous and ambiguous, than in the two surface-scope conditions.

For the ellipsis sentence reading times, there were significant main effects of context ($F_{1}(1,27)=8.9$, $p=0.006$; $F_{2}(1,23)=6.5$, $p=0.02$) and of ambiguity ($F_{1}(1,27)=4.7$, $p=0.04$; $F_{2}(1,23)=3.8$, $p=0.06$). There was only slight evidence for an interaction between context and ambiguity ($F_{1}(1,27)=2.3$, $p=0.14$; $F_{2}(1,23)=3.5$, $p=0.07$). Comparing the two inverse-scope conditions, the effect of ambiguity was significant ($t_{1}(54)=1.9$, $p=0.06$; $t_{2}(46)=2.2$, $p=0.03$). To summarize, the ellipsis clause was read most slowly when the quantified clause was unambiguous inverse scope. Again, this pattern of results is slightly different from Experiment 7, where the ellipsis clause was read more slowly in both inverse-scope conditions than in the surface-scope conditions.

The following table (Table 14) shows the proportion of responses to the comprehension questions. Comparing the two ambiguous conditions, the effect of context was significant ($F_{1}(1,27)=7.3$, $p=0.01$; $F_{2}(1,23)=4.4$, $p=0.05$). Comparing the two inverse-scope conditions, the effect of ambiguity was significant ($F_{1}(1,27)=58$, $p<0.0001$; $F_{2}(1,27)=53$, $p<0.0001$). The proportions of responses to the comprehension questions are very similar to those in Experiment 7. Perceivers nearly always chose the expected response in the unambiguous conditions. And like in Experiment 7, perceivers chose the inverse-scope response half the time in the ambiguous inverse-scope condition and nearly one-third of the time in the ambiguous surface-scope condition, in spite of the presence of the ellipsis clause within the same sentence.
Table 14.

Experiment 12 responses to comprehension question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response</th>
<th>Inverse-scope response</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.69</td>
<td>0.31</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.12</td>
<td>0.88</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.51</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Correct response  Incorrect response
unambiguous surface scope  0.96  0.04

Experiment 12 Discussion

The comprehension question results in Experiment 12 are similar to those observed in Experiment 7. In the two unambiguous conditions participants nearly always chose the expected response, showing that they were performing the experimental task in the intended way. In the ambiguous conditions perceivers chose the surface-scope response only about two-thirds of the time following the surface-scope context, and about half the time following the inverse-scope context, again exactly as often as they did in Experiment 2 where no ellipsis sentence was present. It is clear that even when the ellipsis clause was within the same sentence as the ambiguous quantified clause, the inverse-scope interpretation was still available to perceivers on a substantial number of trials.
It is unlikely, then, that in the inverse-scope condition in Experiment 7 the processor attempted reanalysis upon encountering the ellipsis sentence but was thwarted by the sentence boundary after the quantified sentence. As we saw in Experiment 5A in Chapter 2, perceivers successfully reanalyzed an inverse-scope quantified sentence upon encountering disambiguating information in the form of a presupposition violation triggered by a singular definite subject DP in the next sentence, but in Experiment 12 there was no such reanalysis when the ellipsis sentence was encountered within the same sentence. Comparing the comprehension question responses from the two experiments, it appears that the ellipsis sentence is not very effective at triggering reanalysis, contrary to the predictions of GSE.

The reading time results for Experiment 12, on the other hand, are somewhat different from those observed in Experiment 7. At the quantified clause, we observe the slow reading time for the unambiguous inverse-scope condition that reflects the cost of assigning inverse scope that I argued for in Chapter 2. However, unlike in Experiment 7, the reading time for the quantified clause in the ambiguous inverse-scope condition is no slower than in the ambiguous surface-scope condition.

This difference is perhaps not surprising if the processor commits to the inverse-scope configuration (and thus incurs the associated cost) only when it is necessary. In the unambiguous condition, if the processing begins upon encountering different, the second word of the sentence, the cost of computing the inverse-scope configuration could be incurred early. In the ambiguous condition, on the other hand, both interpretations are possible. In this case, perhaps the cost of building the inverse-scope configuration is not incurred until the end of the sentence when the wrap-up inferencing takes place. Because the processor knows that the sentence has not yet ended at the end of the quantified clause (since the presentation region ends with the conjunction and), no wrap-up is required, so the processor does not commit to the costly inverse-scope interpretation and postpones the processing cost.

If the processor delays in committing to inverse scope until the end of the sentence, then we should observe slow reading times in the ambiguous inverse-scope condition at the ellipsis clause. But the ellipsis clause in the ambiguous inverse-scope condition was read no more slowly than in the ambiguous surface-scope condition. Only the unambiguous inverse-scope condition shows a slower reading time for the ellipsis clause. The result for the unambiguous condition is, of course, predicted by both GSE and PSE, just as in Experiment 7 above. What is more puzzling is the lack of a slow-down in the ambiguous inverse-scope condition. Considering only the reading time data, we might conclude that
perceivers had not assigned the inverse-scope interpretation at all, since both clauses were
read relatively quickly, but we know from the comprehension question data that about half
the items were assigned inverse scope. If assigning inverse scope incurs a processing cost,
why is there no evidence of this cost at the ellipsis sentence in the ambiguous inverse-scope
condition in Experiment 12?

One possibility is that the cost of assigning inverse scope in the ambiguous
conjoined sentence shows up farther downstream. However, there were no significant
differences across the four conditions in reading times for the sentence following the
conjoined quantified and ellipsis sentence (all \( F < 1 \)), nor was there a difference in
response latencies to the comprehension question between the surface-scope and inverse-
scope ambiguous condition (\( F < 1 \)).

There might be a grammatical account of this anomalous result. It is possible that
conjoining the quantified sentence with the ellipsis sentence allows the quantifier scope
relations to be encoded at LF in a different way from when each is a separate sentence.
Rather than a structure where the universal quantifier in each conjunct is QRed to the top
of its clause, as shown in (141), perhaps the correct representation for the conjoined
sentence includes across-the-board (ATB) movement of the universal, as depicted in (142).

(141)

```
and

every cliff

a climber

scaled every cliff

every cliff

the coach

scaled every cliff
```

(142)
Such a structure would not violate GSE, since the universal is QRed above the existential in the first conjunct.

The proposal is that, because an inverse-scope configuration is achieved by means of ATB movement, assigning inverse scope to the conjoined quantified and ellipsis sentence does not incur the same processing cost that arises when two separate QR operations are necessary. But we do observe a processing cost when the quantified clause is unambiguous inverse scope. Perhaps an ATB construction is not possible for the unambiguous sentence because of the presence of different. Beck (2000) proposed that “Q-bound different” must be in the scope of every in order to be properly bound; perhaps this requirement is even more rigorous, such that different needs to be adjacent to every to be properly bound. If sentence (143) has an LF with an ATB structure, then different cannot be raised to an LF position adjacent to every without violating the coordinate structure constraint (CSC) (Ross, 1967).

(143) A different climber scaled every cliff and the coach did, too.

The ATB structure is thus impossible when only one conjunct contains Q-bound different.

Independent evidence that Q-bound different must be adjacent to the universal to be correctly bound comes from a sentence like (144):

(144) Every climber scaled a different cliff and the water tower.

The first conjunct of (144), every climber scaled a different cliff, is ambiguous between the Q-bound, one-climber-per-cliff interpretation and the interpretation where the cliff is
different from some other cliff previously mentioned in the discourse. The Q-bound interpretation is possible because different can be moved to a position adjacent to every climber at LF. But when the second conjunct is added, only the non-Q-bound reading is possible. Extracting different would violate the CSC, so the Q-bound reading is impossible.

To summarize, this proposal suggests that the quantified clause and ellipsis clause are read relatively quickly when the quantified clause is ambiguous because of the option for low-cost ATB movement of the quantified phrase every cliff. But when the quantified clause is unambiguous inverse scope, the presence of different makes ATB movement impossible, and two costly operations of QR must occur. It seems unlikely, however, that covert ATB movement should be so easy for the processor while covert QR movement is costly. This proposal should be tested by investigating the processing of other constructions that include ATB movement, such as extraction from coordinate structures and either/or disjunctive structures.

The full pattern of results for Experiment 12, including the reading times for the quantified and ellipsis clauses and the responses to the comprehension questions, is not consistent with either GSE or PSE. Both GSE and PSE predicted slow reading times for the ellipsis sentence in both inverse-scope conditions. PSE also predicted slow reading times for inverse-scope quantified sentences. These predictions are borne out only for the unambiguous inverse-scope condition in Experiment 12. Neither principle can account for the relatively fast reading times in the ambiguous inverse-scope condition, given that the comprehension questions revealed that half of the items in this condition received inverse-scope interpretations.

The comprehension question data, on the other hand, are consistent with PSE but not with GSE. GSE predicted that the inverse-scope interpretation would be eliminated by the ellipsis sentence, and it was reasonable to expect that conjoining the ellipsis sentence with the quantified sentence would make the quantified sentence more susceptible to the influence of the ellipsis sentence. But the data show that perceivers chose the inverse-scope response just as often as they did in Experiment 7, where the ellipsis sentence and quantified sentence were separated by a sentence boundary, and indeed just as often as in Experiment 2 where the ellipsis sentence was not included. This result is consistent with PSE, which predicted that the inverse-scope interpretation would be dispreferred but still available. Thus, while neither hypothesis can account fully for the observed data in Experiment 12, the results provide stronger support for the principle of Processing Scope Economy than for Grammatical Scope Economy. One thing that we can conclude from Experiment 12 is that reanalysis of an inverse-scope quantified sentence to the surface-
scope interpretation was not assisted by conjoining the ellipsis sentence with the quantified sentence. As I remarked above, the ellipsis sentence is not very effective at triggering reanalysis of an inverse-scope quantified sentence, contrary to the prediction of GSE.

Experiment 13

Experiment 13 is a further replication and extension of Experiment 7, designed to determine the time course of interpretive decisions in the quantified and ellipsis sentences by presenting the target sentences one word at a time rather than in entire sentences. The results confirm the availability of the inverse-scope interpretation even after the ellipsis sentence has been read and processed.

Experiment 13 Method

Stimuli

The stimuli used in Experiment 13 were the same as those used in Experiment 7. The difference between the two experiments was in the presentation of the stimuli. In Experiment 13 the contexts were presented in regions of two or three words and the target sentences were presented one word at a time. Example (145) illustrates the presentation of an example stimulus item, where the pipe symbols (|) divide the presentation regions.

(145) The city | police department | had been | receiving | negative | publicity. The chief | denied that | brutality was | a problem | on the force. However, | when Internal Affairs | investigated, | they discovered that violence | was rampant | throughout the force. | Apparently, a | junior | constable | had | hit | every | suspect. The | senior | sergeant | had, | too. Eventually, | the police chief | was forced | to resign.

A comprehension question followed each item, just as in Experiment 7. The items were arranged in four lists as described above, including the same 24 fillers items used in Experiment 7, as well as 24 additional filler items that were also followed by comprehension questions.
Procedure

The procedure was identical to that used in Experiment 7.

Participants

Forty-nine Northwestern University students, 18 men and 31 women between the ages of 17 and 22 years, participated in Experiment 13 for course credit. The data from nine participants were excluded from the analysis: five who had already participated in a related experiment, two who self-reported as having dyslexia, one who self-reported with Attention Deficit Disorder, and one for whom the computer recorded no data due to an error.

Experiment 13 Results

Residual reading times for the target regions in Experiment 13 were calculated in two ways. To make it possible to compare the results of Experiment 13 directly to those of Experiments 7 and 12, mean residual reading times were calculated for the entire quantified sentence and the entire ellipsis sentence. Since the quantified and ellipsis sentences were presented word-by-word, it was also possible to calculate mean residual reading times for each word in the two sentences. The whole-sentence reading times are depicted in Figure 17 and Figure 18.
Figure 17. Mean residual reading time for entire quantified sentence, Experiment 13.
Figure 18. Mean residual reading time for entire ellipsis sentence, Experiment 13.
When residual reading times were calculated over the entire quantified sentence, there was no evidence for an effect of context (F1, F2<1, n.s.) or of ambiguity (F1, \( F(1,23)<1, n.s. \)), or of an interaction between context and ambiguity (F1, F2<1, n.s.). For the residual reading times calculated over the entire ellipsis sentence there was a significant main effect of context (F1(1,39)=9.1, \( p=0.005 \); F2(1,23)=3.3, \( p=0.08 \)) and a significant interaction between context and ambiguity (F1(1,39)=7.6, \( p=0.008 \); F2(1,23)=3.4, \( p=0.08 \)). There was little evidence for a main effect of ambiguity (F1(1,39)=3.3, \( p=0.08 \), F2(1,23)=1, \( p=0.3 \)). In short, the ellipsis sentence was read more quickly when the quantified sentence was unambiguous surface scope than in the other three conditions. This result is slightly different from that observed in Experiment 7 where the ellipsis sentence was read more slowly in the two inverse-scope conditions than in the two surface-scope conditions.

The following table (Table 15) provides the proportions of responses to the comprehension questions in Experiment 13. Comparing the two ambiguous conditions, there was a significant effect of context (t(178)=2.7, \( p=0.008 \); t(46)=1.9, \( p=0.06 \)). Comparing the two inverse-scope conditions, there was a significant effect of ambiguity (t(178)=6.7, \( p<0.0001 \); t(46)=5.2, \( p<0.0001 \)). Once again, the proportions of surface-scope and inverse-scope responses to the comprehension questions mirror closely the proportions observed in Experiment 7. Of particular note is the fact that perceivers chose the inverse-scope response Several for 41% of ambiguous quantified sentences in the inverse-scope context and 31% in the surface-scope context, in spite of the presence of the following VP-ellipsis sentence.
Table 15.

Experiment 13 responses to comprehension question

<table>
<thead>
<tr>
<th>Condition</th>
<th>Surface-scope response One</th>
<th>Inverse-scope response Several</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous surface-biased</td>
<td>0.69</td>
<td>0.31</td>
</tr>
<tr>
<td>unambiguous inverse scope</td>
<td>0.20</td>
<td>0.80</td>
</tr>
<tr>
<td>ambiguous inverse-biased</td>
<td>0.59</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Correct response   Incorrect response

| unambiguous surface scope | 0.95          | 0.05                          |

Figure 19 depicts the reading time data for each word in the quantified sentence, which reveal several important findings. I performed statistical analysis at three points in the sentence: at the adjective in the subject noun phrase (the adjective in the unambiguous inverse-scope condition was always different), at the verb, and at the last word of the sentence, the object.

The adjective different in the unambiguous inverse-scope condition was read significantly more quickly than the adjective in the other three conditions. There was a significant main effect of context ($F(1,39)=13.6$, $p<0.001$; $F_2(1,23)=8.4$, $p=0.008$), a significant main effect of ambiguity ($F(1,39)=10.1$, $p=0.003$; $F_2(1,23)=3.3$, $p=0.08$), and a significant interaction between context and ambiguity ($F(1,39)=12.4$, $p=0.001$; $F_2(1,23)=4.7$, $p=0.04$).
The verb was read more quickly in the two ambiguous conditions than in the two unambiguous conditions. There was a significant main effect of ambiguity ($F_1(1,39)=11.8$, $p=0.001$; $F_2(1,23)=5.9$, $p=0.02$). There was no evidence for an effect of context ($F_1(1,39)=1.9$, n.s.; $F_2(1,23)<1$, n.s.) or of an interaction between ambiguity and context ($F_1, F_2<1$, n.s.).

The object of the verb, i.e., the last word in the sentence, was read more slowly in the two inverse-scope conditions than in the surface-scope conditions. There was a significant main effect of context ($F_1(1,39)=7.3$, $p=0.01$; $F_2(1,23)=6.7$, $p=0.02$). There was no evidence for an effect of ambiguity ($F_1, F_2<1$, n.s.) or of an interaction between ambiguity and context ($F_1, F_2<1$, n.s.).

![Figure 19. Mean residual reading time for each word in quantified sentence, Experiment 13.](image-url)
Figure 20 shows the word-by-word breakdown of reading times for the ellipsis sentence. I performed statistical analysis at three locations: at the first word in the sentence, which was always the, at the head noun in the subject noun phrase, and at the last word too.

The first word of the ellipsis sentence, the, was read most slowly when the quantified sentence was unambiguous inverse scope. There was a significant main effect of context ($F_{1}(1,39)=7.2, p=0.01; F_{2}(1,23)=5.8, p=0.02$) and a nearly significant interaction between context and ambiguity ($F_{1}(1,39)=4.5, p=0.04; F_{2}(1,23)=3, p=0.09$). There was no evidence for a main effect of ambiguity ($F_{1}, F_{2}<1$, n.s.).

The head of the subject noun phrase was read more slowly in the inverse-scope conditions than in the surface-scope conditions. There was a significant main effect of context ($F_{1}(1,39)=7.2, p=0.01; F_{2}(1,23)=5.6, p=0.03$). There was no evidence for an effect of ambiguity ($F_{1}, F_{2}<1$, n.s.) or of an interaction between ambiguity and context ($F_{1}(1,39)=1.8, n.s.; F_{2}<1, n.s.$).
Figure 20. Mean residual reading time for each word in ellipsis sentence, Experiment 13.
The last word in the ellipsis sentence, *too*, was read more quickly when the quantified sentence was unambiguous surface-scope, but this difference was significant only in the analysis by subjects. The effect of context approached significance by subjects (\(F_1(1,39)=3.9, p=0.06; F_2<1, \text{n.s.}\)). The effect of ambiguity was significant only by subjects (\(F_1(1,39)=8.9, p=0.005; F_2(1,23)=2.7, p=0.12\)). There was no evidence of an interaction between the two factors (\(F_1, F_2<1, \text{n.s.}\)).

**Experiment 13 Discussion**

The word-by-word reading times for the quantified sentence replicate the finding of Experiment 7 and confirm the prediction of the PSE principle that assigning inverse scope to a quantified sentence is more costly than assigning surface scope. The results of Experiment 13 provide the additional information that this cost is incurred only at the end of the sentence, a finding that is consistent with the well-attested phenomenon of sentence wrap-up effects (Just & Carpenter, 1980). We can also conclude from the word-by-word data that the cost of inverse scope is not due to the presence of what Beck (2000) calls “Q-bound different” since *different* in the unambiguous inverse-scope condition was actually read more quickly than the adjectives that occupied the same position in the other three positions. In other words, *different* is processed relatively quickly, but the cost of inverse scope catches up by the end of the sentence.

In the reading times for the ellipsis sentence we observe once more the pattern that the ellipsis sentence was read more slowly in the inverse-scope conditions. The difference in Experiment 13 from Experiment 7 is that the ellipsis sentence is also read slowly when the quantified sentence was ambiguous surface scope. This result could be a symptom of the number of items (31%) to which perceivers assigned the inverse-scope interpretation, consuming processing resources, but if so, then why was the effect not observed in previous experiments where perceivers assigned inverse scope on some non-trivial proportion of surface-scope trials? The result might instead be an indicator of a delay in interpreting the quantified sentence, such that the cost persisted into the ellipsis sentence, such as was observed in Experiments 4 and 5.

Experiment 13 also provides still further confirmation that the inverse-scope interpretation is still available to perceivers even after they have read the ellipsis sentence. Participants chose the inverse-scope response *Several* to the comprehension question on 31% of trials in the ambiguous surface-scope condition and 41% of trials in the ambiguous inverse-scope condition. Once again, the comprehension question data conflict with the predictions of GSE but support PSE.
General Discussion

At the outset of this chapter I presented Fox’s (2000) grammatical principle of Scope Economy and considered its consequences for real-time sentence comprehension. I conjectured that if the processor encounters a sentence that violates Grammatical Scope Economy, it will attempt reanalysis in order to satisfy Scope Economy. In the constructions that I consider in this chapter, that is, pairs of doubly quantified and ellipsis sentences such as (146)–(147),

(146) An experienced climber scaled every cliff.

(147) The shop owner did, too.

the only parse that satisfies both Grammatical Scope Economy and the Parallelism requirement on VP-ellipsis is one where both the quantified sentence and the ellipsis sentence have surface-scope configurations. If the quantified sentence has already been assigned inverse scope, then the ungrammatical ellipsis sentence will prompt the processor to revise both the quantified sentence and the ellipsis sentence in order to give them both surface scope. GSE therefore predicted slow reading times for ellipsis sentences in inverse-scope contexts, and a high proportion of surface-scope responses to the comprehension questions.

I also proposed an alternative to GSE whereby apparent Scope Economy effects originate in the processor rather than the grammar. The principle of Processing Scope Economy, which was supported by the results of Chapter 2, postulates that an inverse-scope configuration is dispreferred because it is linguistically more complex than surface scope, and that the processor will therefore always prefer to construct a surface-scope configuration. According to PSE assigning inverse scope is possible, but incurs a processing cost. PSE predicted slow reading times for quantified and ellipsis sentences in inverse-scope contexts because of the relative complexity of an inverse-scope representation, but predicted similar proportions of comprehension question responses for items containing ellipsis sentences as for items without. The two principles thus made similar predictions about the reading times, but differed crucially in their predictions about the responses to the comprehension questions in inverse-scope contexts. GSE predicted that the ellipsis sentence would prompt reanalysis in inverse-scope contexts leading to a high proportion of surface-scope responses. PSE, in contrast, predicted just as many inverse-scope responses following an ellipsis sentence as when the ellipsis sentence was absent,
since the inverse-scope interpretation is grammatical even when the ellipsis sentence is present.

There were some differences in the reading time results of the three self-paced reading experiments and one pilot self-paced reading experiment presented in this chapter. In general, the unambiguous inverse-scope quantified sentence was always read more slowly than the unambiguous surface-scope quantified sentence in all four experiments. Likewise, the ellipsis sentence was always read quickly following an unambiguous surface-scope quantified sentence and slowly following an unambiguous inverse-scope quantified sentence. This pattern is consistent with the predictions of both hypotheses. The pattern of data differed somewhat across the experiments for the two ambiguous conditions: Experiment 7 showed that an ambiguous quantified sentence was read more slowly in inverse-scope contexts than in surface-scope contexts, while in Experiment 12 the ambiguous quantified sentence in both contexts was read as quickly as the unambiguous surface-scope quantified sentence, and in Experiment 13 the ambiguous quantified sentence in both contexts was read as slowly as the unambiguous inverse-scope quantified sentence. Further investigation is needed to determine the cause of these differences across the experiments, since the differences do not appear to be related in an obvious way to the differences in presentation of the stimuli (conjoining the two target sentences in Experiment 12 and presenting the target sentences one word at a time in Experiment 13).

Notwithstanding these irregularities in the reading time data, the comprehension question data revealed a clear and consistent pattern across all three experiments and the pilot experiment. Participants behaved in the expected way in the unambiguous conditions, choosing the surface-scope response on nearly every trial in the surface-scope condition and the inverse-scope response for most trials in the inverse-scope condition. In the conditions where the quantified sentence was ambiguous, participants selected about two-thirds surface-scope responses and one-third inverse-scope responses following the surface-scope contexts, and about half surface-scope and half inverse-scope responses following the inverse-scope contexts. In short, the presence of a singly quantified ellipsis sentence did not restrict the doubly quantified antecedent sentence to the surface-scope interpretation. The inverse-scope interpretation was still available to perceivers who had read and processed the ellipsis sentence.

I presented several possible explanations for this pattern of results. First, I considered the possibility that the stimulus items used in Experiments 7, 12 and 13 were not actually subject to Grammatical Scope Economy because they used eventive, stage-level predicates that contain implicit event quantifiers. If the implicit quantifiers allowed the
inverse-scope configuration for the ellipsis sentence by providing a second, truth-conditionally distinct interpretation of the sentence, GSE should still constrain the LF of non-eventive sentences that contain only a single quantifier. According to GSE, a Scope Economy violation in a stative ellipsis sentence should trigger reanalysis, with the result that only the surface-scope interpretation is available for the quantified sentence. Pilot Experiment 11 tested this hypothesis with four stimulus items that used stative, individual-level predicates. The results of the pilot experiment showed that perceivers frequently selected the inverse-scope interpretation response in both ambiguous conditions. The stative ellipsis sentence did not eliminate the inverse-scope interpretation of the quantified sentence, contrary to the prediction of GSE.

I also considered the possibility that the singly quantified ellipsis sentence did prompt reanalysis of the quantified sentence, as GSE predicted, but that this reanalysis failed because of the inaccessibility of the quantified sentence. It is well established that the syntactic representation of a sentence decays from memory once the processor carries out the sentence-final inferencing known as sentence wrap-up. In Experiment 12 I manipulated the availability of the quantified sentence for reanalysis by including it in the same sentence with the ellipsis sentence, so that its syntactic representation would still be accessible when the ellipsis sentence was encountered. This manipulation did not, however, facilitate reanalysis of the quantified sentence, since perceivers chose the inverse-scope response on just as many trials as in Experiment 7 where the ellipsis sentence was separated from the quantified sentence by a sentence boundary. We also had reason to suspect that the sentence boundary was not sufficient to prevent reanalysis of an inverse-scope quantified sentence, since Experiment 5A in Chapter 2 showed that a presupposition violation could trigger reanalysis to surface scope fairly effectively. Experiment 12 showed once again that the ellipsis sentence did not require a surface-scope interpretation for the quantified sentence, contrary to the prediction of GSE.

It is clear, therefore, that where the predictions of the two hypotheses differ, that is, in the responses to the comprehension questions in inverse-scope contexts, the results of all three self-paced reading experiments as well as the pilot self-paced reading experiment conflict with the predictions of the principle of Grammatical Scope Economy and support the predictions of the Processing Scope Economy principle. In addition to its empirical advantages, PSE also offers several theoretical advantages. While the present form of the hypothesis is specific to sentences with scope ambiguities, PSE may well be one instantiation of a more general processing principle that prefers simpler structures to more complex ones.
Recall that the Grammatical Scope Economy principle had a significant theoretical drawback associated with it, since it required that the syntactic well-formedness of a sentence’s LF be subject to aspects of the sentence’s meaning, thus jeopardizing the modularity of the grammar. If we can eliminate GSE as a principle governing LF configurations, then the autonomy of syntax is restored. In fact, the entire notion of economy in grammatical representations is somewhat suspect, since there is no inherent reason for a complex structure to be any less grammatical than a simpler one. The sentence processing mechanism, on the other hand, needs to be concerned with economy in parsing, since it has only limited cognitive resources available during on-line comprehension. Localizing economy effects in the processor’s inherent parsing preferences makes sense from the point of view both of grammar and of processing.

PSE also allows us to remain agnostic about the nature of the grammar of quantification. The principle proposes that inverse-scope interpretations are dispreferred because they are more complex, but, unlike GSE, does not require a commitment to a Quantifier Raising account of quantifier interpretation. Under a flexible types approach (Hendriks, 1987; Partee, 1986; Partee & Rooth, 1983), for example, the inverse-scope interpretation would still be the more complex representation in terms of its linguistic structure, so PSE would still hold if this turns out to be the more correct account.

The weight of the experimental evidence presented in this chapter together with the theoretical advantages of PSE lead me to reject the existence of Scope Economy as a grammatical constraint on QR operations. I conclude, therefore, that the principle of Processing Scope Economy forms some part of the architecture of the human sentence processing mechanism.

One puzzle still remains, though. I began this chapter with an observation of the apparent disambiguating effect of a singly quantified ellipsis sentence on its doubly quantified antecedent sentence. If Scope Economy is not part of the grammar, then how can we account for this apparent disambiguation, which is widely assumed to be true? If we accept PSE, as I have proposed, then the actual prediction is that the combination of a singular definite ellipsis sentence with an inverse-scope quantified sentence, as in (148), is not ungrammatical but simply difficult to process.

(148) Upon the signal, a paratrooper jumped from every plane, and the squadron commander did too.
Recent work by Carlson (2002) has shown that processing is facilitated when items that have parallel or similar properties, whether grammatical or extra-grammatical, occupy parallel syntactic positions. If this principle applies to a sentence like (148), the parallel surface syntax between the ellipsis sentence and its antecedent leads the processor to prefer to interpret them with parallel mental models. Since the first conjunct involves multiple paratroopers jumping from planes, the processor’s preferred interpretation for the second conjunct is one with multiple jumpers. Since the ellipsis clause involves only one jumper, the mismatch in conceptual parallelism makes the sentence difficult to process.

There exist other, well-known cases of unprocessability masquerading as ungrammaticality; center-embedded sentences such as (149) are among the clearest examples.

(149) The man the boy the dog bit saw left.

There is no grammatical criterion by which sentence (149) is ill-formed: the constituent the boy the dog bit is an acceptable noun phrase containing a relative clause, as is the man the boy saw. And the man left is a perfectly well-formed sentence. The tree diagram in (150) illustrates a well-formed representation of the sentence.

(150)

Yet in spite of the fact that it is perfectly grammatical, sentence (149) gives the impression of profound unacceptability. While it may be generated by the grammar of English, it would surely never be uttered spontaneously nor understood without considerable effort. The sentence’s unacceptability stems not from ungrammaticality, but from its unprocessability. The processor can handle one center-embedded relative clause, as illustrated by the acceptability of “The man the boy saw left” but the introduction of a second relative clause that makes the boy into the boy the dog bit renders the sentence impossible.
Sentence (149) serves as a caution that the unacceptability of a sentence is not sufficient evidence for classifying it as ungrammatical.

The same logic applies to sentence (148). Although linguists have assumed that an ellipsis sentence following an inverse-scope quantified sentence is ungrammatical, the only thing that we can conclude from intuition is that it is only marginally acceptable. Even though intuition suggests that (148) is unacceptable, the evidence from four experiments showed that perceivers found sentences with the same structure as (148) acceptable as much as half the time. Presumably the factor that makes these sentences acceptable is simply a factor that facilitates processing. If the context supports a mental model where the two conjuncts have non-parallel conceptual representations, then it should be comparatively easy for the processor to assign them non-parallel interpretations.

In this chapter I have provided empirical evidence in favour of the principle of Processing Scope Economy and against the existence of Grammatical Scope Economy. This evidence led me to conclude that Scope Economy is not a component of the grammar of quantification, but that apparent Scope Economy effects arise as a result of the sentence processing mechanism's inherent preference for simpler scope configurations, combined with the facilitating effects of conceptual parallelism.
I have presented a set of experiments that investigated the real-time comprehension of doubly quantified sentences. Although there have been relatively few such studies before now, the data from these on-line processing experiments offer insight not only into the psycholinguistic question of how sentences containing quantifiers are processed and understood, but also into the linguistic question of how quantifiers are represented at the syntax-semantics interface.

The Role of LF Structure in Real-time Semantic Inferencing

Lyn Frazier’s (1999) book highlighted a gap in the current sentence processing literature. Many studies of sentence processing have concentrated on how sentence structure is parsed. Likewise, many studies have examined how conceptual, non-linguistic representations (e.g., mental models) are constructed. The gap is in how the mind interprets a parsed structure to arrive at a conceptual representation. Serial, syntax-first models of sentence processing (Ferreira & Clifton, 1986; Frazier, 1987, 1990; Frazier & Clifton, 1996; Frazier & Fodor, 1978) posit that a sentence’s interpretation is computed from the completely parsed syntactic structure. On the other hand, in a constraint-based model in which multiple analyses compete for activation (MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus et al., 2000; Trueswell & Tanenhaus, 1994) structural information (which may indeed be merely an artifact of probabilistic information such as co-occurrence frequencies) might be overridden by other sources of information such as contextual bias. Quantifier scope ambiguity offers a unique opportunity to examine the role of abstract linguistic structure in forming conceptual representations, since the two interpretations of a doubly quantified sentence arise from a single surface syntactic representation. Any differences in processing between the two interpretations should, all else being equal, indicate differences in their more abstract structure, that is, in their respective LFs.

LF Complexity Leads to Processing Difficulty

The questionnaires presented in Chapter 2 revealed that perceivers had a default preference for the surface-scope interpretation of a doubly quantified sentence, which is consistent with intuition and with other results reported in the literature (Ioup, 1975; Kurtzman & MacDonald, 1993; Tunstall, 1998; VanLehn, 1978). In the first questionnaire, perceivers chose the surface-scope interpretation about 80% of the time for
a doubly quantified sentence presented in isolation. The context questionnaire showed that supporting discourse contexts could boost the low preference for the inverse-scope interpretation from 20% to slightly more than 50%.

Two self-paced reading experiments presented the doubly quantified sentences in isolation. The first found that a sentence following the quantified sentence that disambiguated to inverse scope was read more slowly than a surface-scope disambiguating sentence, for both surface configurations of quantifiers (\(a\ldots every\) and \(every\ldots a\)). The second showed that perceivers read the quantified sentence significantly more slowly when they assigned an inverse-scope interpretation to it than when they assigned the surface-scope interpretation. These results again confirmed the preference for the surface-scope interpretation and showed that assigning the inverse-scope interpretation incurred a processing cost.

The results of the experiments that presented the quantified sentences in context found that this processing cost was incurred even when the quantified sentence followed a context that reliably supported the inverse-scope interpretation, showing that contextual information was not sufficient to mitigate the cost, contrary to the predictions of the principle of parsimony (Altmann & Steedman, 1988; Crain & Steedman, 1985). This processing cost was found even in the baseline condition where the quantified sentence was unambiguous inverse scope. This last finding indicated that the processing cost cannot be attributed to competition between alternative analyses, as Kurtzman & MacDonald (1993) suggest.

**LF Structure is Crucial for Semantic Inferencing**

Taken together, the experiments presented in Chapter 2 showed that perceivers experienced processing difficulty in assigning the inverse-scope interpretation to a doubly quantified sentence, even when the inverse-scope interpretation was referentially simpler than the surface-scope interpretation (in \(every\ldots a\) sentences), even when the inverse-scope interpretation was supported by the preceding context, and even when there was no competing surface-scope analysis available (in the unambiguous condition). The results support an account of quantifier scope ambiguity wherein the inverse-scope interpretation has a more complex syntactic representation, such as the QR account (May, 1977, 1985), and a model of sentence comprehension in which syntactic structure makes a major contribution to interpretive decisions, such as the garden-path model or the Late Assignment of Syntax (LAST) model (Townsend & Bever, 2001).
The results are less consistent, on the other hand, with parallel, constraint-based models of sentence processing (MacDonald, 1994; Spivey & Tanenhaus, 1998; Tanenhaus et al., 2000; Trueswell, 1996; Trueswell & Tanenhaus, 1994). The slow reading times for the unambiguous inverse-scope quantified sentence are particularly troublesome for a model that considers multiple analyses in parallel (e.g., Kurtzman & MacDonald (1993)), since such a model predicts that processing difficulty is a result of two or more analyses competing with each other, but in this case processing difficulty arises even when there is no competing analysis.

The principle of parsimony (Altmann & Steedman, 1988; Crain & Steedman, 1985), which has been proposed as one constraint in a parallel model, predicts a cost for assigning the referentially more complex of competing analyses. For an a...every sentence like (151), for example, the surface-scope interpretation is referentially simpler because it involves only a single climber, while the inverse-scope interpretation involves multiple climbers.

(151) A climber scaled every cliff.

The principle thus predicts that the inverse-scope interpretation of an a...every sentence is more difficult to assign. However, Experiments 4 and 5 showed that disambiguating to the inverse-scope interpretation was also slow for every...a sentences even though the inverse-scope interpretation is referentially simpler than the surface-scope interpretation. Furthermore, the inverse-scope interpretation of an a...every sentence incurred a processing cost in Experiments 5 and 7 even when the quantified sentence was embedded in a context that set up the multiple referents beforehand.

In sum, while parallel processing theories can account for certain pieces of the data presented in Chapter 2, the entire set of results favours a theory in which syntactic structure plays a major role.

The Role of Semantics in Determining Well-formedness of LF Structure

Fox’s (2000) grammatical principle of Scope Economy theory proposed that the syntactic representation of a quantified sentence, specifically, the LF operation of Quantifier Raising (QR), is constrained by a sentence’s truth-conditions. This proposal undermines the modularity of grammar by requiring that the syntactic component consult
the semantics to determine well-formedness. Part of the evidence for the proposal was the interaction of doubly quantified sentences with ellipsis sentences such as the examples in (152)–(153).

(152) A paratrooper jumped from every plane.

(153) The squadron sergeant did, too.

Many people find pairs of sentences such as (152)–(153) acceptable only on the surface-scope interpretation of (152), where a single paratrooper jumps from each plane in succession, although this intuitive judgment is disputed. In other words, the ellipsis sentence (153) is ungrammatical on the inverse-scope interpretation of (152). Fox’s analysis of these grammaticality judgments predicts that, in real time, the inverse-scope interpretation of (152) should be unavailable to perceivers who have read and processed sentence (153). If the on-line comprehension data conflict with the introspective data, this should motivate a reexamination of the facts about sentences like (152)–(153) and of the theory based on them.

Inverse-scope Interpretation is Available Following Ellipsis Sentence

The experiments reported in Chapter 3 presented pairs of quantified and ellipsis sentences like (152)–(153) embedded in context paragraphs that supported either the surface-scope or the inverse-scope interpretation of the quantified sentence. After reading each paragraph, perceivers answered a question that indicated which interpretation they had assigned. The results of a questionnaire experiment showed that perceivers assigned the inverse-scope interpretation when they had read the ellipsis sentence just as often as they did when the ellipsis sentence was not present, that is, on about 50% of trials. This result suggested that the ellipsis sentence did not disambiguate the quantified sentence, contrary to the predictions of the Grammatical Scope Economy theory.

In the self-paced reading experiments, perceivers experienced processing difficulty, manifested as slower reading times, at the ellipsis sentence following the quantified sentence in an inverse-scope context, but still selected the inverse-scope response to the comprehension question about half the time. This same pattern of results was found when the quantified and ellipsis sentence were conjoined by and to form single sentence, and for sentences with individual-level, non-eventive predicates.
Scope Economy Reflects a Processing Preference

The Grammatical Scope Economy proposal was based on the intuition that an ellipsis sentence disambiguates a preceding doubly quantified sentence, eliminating the inverse-scope interpretation. However, the results of the questionnaire and self-paced reading experiments indicate that the inverse-scope interpretation was still available to perceivers following the ellipsis sentence, both in a time-constrained task (the self-paced reading task) and in a task that allowed more time for introspection (the questionnaire). Clearly the ellipsis sentence did not have the predicted disambiguating effect.

Based on the Chapter 2 experiments, I proposed that the sentence processing mechanism includes a principle that prefers simpler scope configurations over more complex ones. The principle of Processing Scope Economy is, of course, a particular instantiation of the more general principle that the processor prefers to construct the simplest syntactic representation necessary (Frazier, 1999; Frazier & Fodor, 1978; Pritchett & Whitman, 1995; Tunstall, 1998). PSE accounts for the observed processing difficulty at the ellipsis sentence following an inverse-scope quantified sentence, since the more complex inverse-scope configuration that Parallelism requires for the ellipsis sentence is dispreferred and consumes processing resources. At the same time, PSE leaves the dispreferred inverse-scope interpretation available to be assigned, since it is merely dispreferred and not ungrammatical.

This processing account offers a lesson about inferring ungrammaticality from a judgment of unacceptability. As long as the inverse-scope interpretation of a doubly quantified sentence was considered ungrammatical in the presence of an ellipsis sentence, then a grammatical explanation was required. Unfortunately, the grammatical solution that Fox proposed had the theoretically undesirable consequence of threatening the autonomy of the syntactic component of the grammar. But the processing data have revealed that the ellipsis sentence does not render the inverse-scope interpretation ungrammatical; indeed, the inverse-scope interpretation was not even any less frequent following the ellipsis sentence than when the ellipsis sentence was not present. PSE is independently motivated for many other kinds of complex syntactic constructions and it accounts for the real-time comprehension data. Apparently, an inverse-scope quantified sentence followed by an ellipsis sentence is not ungrammatical but simply difficult to process, because of the processor’s inherent preferences: for simpler structural representations, and for conceptual parallelism between elements that have parallel syntax. On this processing account, there is no need for the syntactic well-formedness of quantified
sentences to be subject to their semantics, since the inverse-scope configuration is available (but dispreferred) even if it does not lead to different truth-conditions. The data presented in Chapter 3 thus illustrate the value of obtaining objective measures of a sentence’s acceptability in cases where the grammaticality intuitions are in dispute.

Future Research

In Chapter 2 I argue that the processing cost associated with assigning the inverse-scope interpretation is a result of the greater complexity of the inverse-scope LF. However, the only quantifier-determiners that I tested were a and every. An obvious next step is to seek to confirm this hypothesis using other quantifiers. It will be particularly interesting to investigate sentences with each such as (154), since each has a stronger tendency than every to take wide scope (1983).

(154) A climber scaled each cliff.

If the cost of inverse scope is indeed a structural cost, then we should observe it even when the properties of the quantifiers favour an inverse-scope configuration, just as we did when the presence of different required inverse scope.

Likewise, the cost of inverse scope should be observable when an ellipsis sentence with an indefinite subject follows an inverse-scope quantified sentence, as in (155)–(156).

(155) A boy climbed each tree.

(156) A girl did, too.

PSE predicts that the slow reading times observed for a definite ellipsis sentence following an inverse-scope quantified sentence were not due to the ungrammaticality of the ellipsis sentence, but simply to the difficulty of constructing the inverse-scope configuration that Parallelism required for the ellipsis sentence. We should therefore observe a comparable processing cost for an indefinite ellipsis sentence such as (156) following an inverse-scope quantified sentence. If, however, the processing difficulty for the definite ellipsis sentence is a result of ungrammaticality, then we should observe no such difficulty for the indefinite, which is not predicted to be subject to Grammatical Scope Economy.
Another obvious follow-up is to extend Pilot Experiment 11, which used stimuli with individual-level predicates, into a full-scale experiment in order to obtain statistically significant results. That pilot was designed to eliminate the possibility that the ellipsis sentences’ failure to disambiguate the doubly quantified sentences in Experiment 7 was a result of existential quantifiers over events implicit in the stage-level predicates. The pattern of results in the pilot experiment mirrored the patterns observed in Experiments 7, 12 and 13, but because the pilot included only four stimulus items, it was not possible to test for significance.

One flaw that all of the self-paced reading experiments reported here share is that the quantified sentence was always followed by a sentence that could potentially disambiguate it. It turned out that the ellipsis sentence did not reliably disambiguate the quantified sentence, but the possibility exists that it could have triggered reanalysis of the quantified sentence for some perceivers or some items. For more compelling evidence of the cost of inverse scope, it would be better to design an experiment that questioned perceivers’ interpretation of the quantified sentence without any following disambiguating material.

Conclusions

If there is a moral to this story, it has two complementary elements. The first is an exhortation to theoretical linguists of the need to attend to psycholinguistic considerations, since a processing problem might be masquerading as a grammar problem. The second is a lesson to psycholinguists to consider abstract linguistic structure, since it plays a key role in real-time comprehension, in addition to the contributions of probabilistic, contextual and conceptual information.

The conclusions of this study represent two facets of a single theme, that of the integrity of syntax. In Chapter 2 I argued for the crucial role of syntactic structure in online interpretive decisions made by the human sentence processing mechanism, and in Chapter 3 I showed that an apparent syntactic principle is actually the result of a processor that operates in this way. Interpretive decisions are undergirded by abstract linguistic structure, not trivially by surface syntax but by covert LF representations. In the particular cases that I consider in this dissertation, the scope of quantified expressions is established at LF. Furthermore, LF is the input to the semantics and is not constrained by the results of semantic computation. In the case of quantified and ellipsis sentences, what had
appeared to be an instance of semantics governing syntactic well-formedness turned out to be a symptom of the processor’s operating principles, which are sensitive to the complexity of LF representations.


