Decompositionality and Identity

by

Shoichi Takahashi

B.A., Tamagawa University, 1996
M.A., Tamagawa University, 1998

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2006

© Shoichi Takahashi 2006. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute
publicly paper and electronic copies of this thesis documents in whole or
in part in any medium now known or hereafter created.

Author ........................... 

Department of Linguistics and Philosophy
January 31, 2006

Certified by ...........................

Danny Fox
Associate Professor of Linguistics
Thesis Supervisor

Accepted by ...........................

Irene Heim
Professor of Linguistics
Chair of the Linguistics Program
Decompositionality and Identity

by

Shoichi Takahashi

Submitted to the Department of Linguistics and Philosophy
on January 31, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The Scope of quantificational phrases (QPs) is often not represented in surface structures, in the sense that a constituent that is, on the surface, sister to a QP is not necessarily interpreted as its argument semantically. QPs can take wider or narrower scope than their surface position (Scope Widening and Scope Narrowing). Based on the fact that when an elided constituent contains a variable that takes part in a particular type of binding dependency, the biggest deletable constituent must be elided, Chapter 2 argues that the presence of this maximization effect in the context of Scope Widening is straightforwardly explained by syntactic approaches, in which QPs undergo movement to widen their scope, rather than by semantic approaches, which adopt type shifting rules to deal with Scope Widening. Chapter 3 defends the copy theory of movement, which captures the effects of Scope Narrowing in a simple way, by developing a new theory of the counter-cyclic merger, which accounts for the facts that seem challenging to this theory. Those facts have been taken to indicate that movement sometimes leaves a trace. I propose that restrictors of determiners can merge with them counter-cyclically. Copies of determiners without their restrictors turn into syntactic objects that receive the same interpretation as that assigned to a trace. This is done by the independently motivated procedure that converts movement chains into interpretable objects. Chapter 3 demonstrates that together with other properties of grammar, this approach explains the challenging facts, and yet is compatible with the copy theory. Chapter 4 investigates the nature of the scope assignment mechanism by examining the fact that the scopal possibilities of comparative QPs (e.g., more than three books) are more restricted than other QPs. It demonstrates that this fact is captured by the interplay of the decompositional approach, in which comparative QPs are decomposed into two QPs (e.g., -er than three and many books), and independently motivated constraints on the scope assignment (e.g., locality and Scope Economy in Fox 2000). Thus, their peculiar scopal behavior can be taken as additional evidence for the core properties of the scope assignment mechanism.

Thesis Supervisor: Danny Fox
Title: Associate Professor
Acknowledgments

I owe a tremendous intellectual debt to the chair of my thesis committee, Danny Fox. In the fall semester of my first year at MIT, I was very fond of sitting down in a comfortable black leather chair in the second floor corridor in E39. One day, Danny talked to me when I was there, and we had a conversation about quantifiers, which is the origin of Chapter 4 of this dissertation. Since then, Danny kindly spent countless hours to discuss various syntactic and semantic issues with me, which inspired and had a great influence on me for the past four and a half years. I am also deeply grateful to him for his generosity in allowing me to include our collaborative work on ellipsis in this dissertation as Chapter 2. I can never thank Danny enough for his constant intellectual guidance, support, encouragement, and patience.

I would also like to express my profound gratitude to other members of my thesis committee, Irene Heim, Kyle Johnson, and David Pesetsky. Irene taught Introduction to Semantics with Danny in my first year. Their lectures contained a lot of wonderful discussion and illustration about what studying semantics teaches us about the nature of language and led me to pursue the 5 year program specialization in semantics, though I had only studied syntax at that point. At every stage of my MIT life, I have been greatly inspired by a countless number of meetings with her and by her work. I asked Kyle to join my thesis committee at the last minute. Nonetheless, he generously complied with my request. He kindly traveled to MIT from Amherst and provided me many insightful and inspiring comments. David was always very generous in clarifying my vague thoughts and pointed out accurately what would be the next step that I should take to make headway on my projects. His talent for organizing ideas as well as his helpful and invaluable comments improved the dissertation enormously. I am deeply grateful to have had the opportunity to work with all of the members of my thesis committee.

I also want to thank other professors at the department for providing me the peerless wonderful environment for studying linguistics and undertaking research: Adam Albright, Noam Chomsky, Michel DeGraff, Kai von Fintel, Suzanne Flynn, Edward Gibson, Michael Glanzberg, Morris Halle, James Harris, Sabine Iatridou, Michael
I had the good fortune to start my life at MIT with the class of ling-01: Allison Adler, Pranav Anand, Justin Fitzpatrick, Valentine Hacquard, Andrés Salanova, Mary Ann Walter. I often indulge in our sweet reminiscences of our first year days, when we worked together every day and every night. I really miss those days. I also feel regret at getting to throw a ling-01 dinner less often than now, but I am pretty sure that we will continue to do it even after we graduate. Without their friendship, boundless kindness and support, my last four years and half at MIT would have been much less fun and more painful. Thanks also go to the extended ling-01 members for their friendship: Diana Fitzpatrick, Silva Maria Amélia Reis, Nikki Vamosi, and Ben Walter.

I also want to express my gratitude to my colleagues, Ken Hiraiwa and Sachiko Kato, with whom I lived in a pleasant house in Arlington for two and a half years. Our conversations about various topics including linguistics often lasted all through the night. 37 Alfred Rd. is a place associated with full of memories of my happy MIT life.

I have been lucky enough to share offices with my great officemates. Anikó Csirmaz and Youngjoo Lee were my officemates in our sentimental E39 for one and a half years and Sarah Hulsey and Ivona Kučerová, in our new building, the Ray and Maria Stata Center, for two years. I would like to thank them for their friendship, constant support, and providing me the comfortable and relaxed offices. I will also miss a spectacular view of downtown Boston lying spread out in front of my desk.

The following people also deserve special mention. I benefited from discussion with Martin Hackl, Jon Nissenbaum, and Uli Sauerland. Discussion with them as well as their writings inspired me at all stages of working on the projects reported in the dissertation. Rajesh Bhatt and Christopher Kennedy were visiting professors in my second year. They generously spent time to discuss various issues of comparatives and meeting with them was very fruitful. Pauline Jacobson gave a course of variable-free semantics at Harvard University in the fall of 2004, which greatly inspired me and had an impact on my works. Pauline also provided Danny and me with her helpful comments on an earlier version of Chapter 2. Sigrid Beck gave me detailed written comments on an earlier version of Chapter 4, which improved it substantially. Nobuko Hasegawa,
Kazuko Inoue, and Akira Watanabe introduced me to formal linguistics when I was a graduate student at Kanda University of International Studies in Japan. Their encouragement and support continued even after I came to MIT. Nobuko was a visiting scholar in my fourth year, and it was very lucky that I could talk with her about linguistics and other things at MIT.

Special thanks to Feng-fan Hsieh and Hironobu Kasai for their friendship. Hironobu is a friend of longstanding since we were graduate students at Kanda University. I was very fortunate to continue to have him in the neighborhood of MIT (i.e., at Harvard). I had a countless number of lunches and dinners with Feng-fan, who took me to lots of Chinese and Taiwanese restaurants, and introduced various sorts of delicious foods to me. I am also always impressed by his incredibly fluent Japanese. Their presence extremely enriched my MIT life.

I want to thank the following people who I interacted with in the past four and a half years. All of them contributed to the dissertation or my life at MIT in all sorts of ways (e.g., giving judgments, comments, encouragement, fun, or various other support): Jun Abe, Marta Abrusan, Karlos Arregi, Asaf Bachrach, Mark Baltin, Teal Doggett Bissell, Cedric Boeckx, John Bowers, Michael Brody, Seth Cable, Markéta Ceplová, Jessica Coon, Bridget Copley, Maria Cristina Cuervo, Paul Elbourne, Marcelo Ferreira, Elissa Flagg, Kazuma Fujimaki, Koji Fujita, Jon Gajewski, Gillian Gallagher, Maria Giavazzi, Martina Gracanin-Yuksek, Andrea Gualmini, Elena Guerzoni, Daniel Harbour, Daniel Hardt, David Hill, Franny Hsiao, Michela Ippolito, Takako Iseda, Shinichiro Ishihara, Yukiko Kambara, Noboru Kamiya, Masaomi Kato, Roni Katzir, Ruriko Kawashima, Ezra Keshet, Sunny Kim, Hisatsugu Kitahara, Heejeong Ko, Masatoshi Koizumi, Sachie Kotani, Haruo Kubozono, Masakazu Kuno, Idan Landau, Howard Lasnik, Winfried Lechner, Julie Legate, Zhiqiang Li, Vivian Lin, A noop Mahajan, Hideki Maki, Yoko Maki, Giorgio Magri, Roger Martin, Tatjana Marvin, Lisa Matthewson, Ora Matushansky, Tara McAllister, Martha McGinnis, Jason Merchant, Luisa Meroni, Edson Miyamoto, Mitsue Motomura, Alan Munn, Kimiko Nakanishi, Lance Nathan, Andrew Nevins, Masashi Nomura, Rebecca Norris, Junko Ochi, Masao Ochi, Naoko Okura, Liina Pylkkänen, Andrea Rackowski, Tanya Reinhart, Charles Reiss, Joseph Sabbagh, Ken Safir, Carson Schütze, Junri Shimada, Rajvinder Singh, Penka
Stateva, Tamina Stephenson, Linnaea Stockall, Koji Sugisaki, Shogo Suzuki, Ana Szabolcsi, Hidekazu Tanaka, Christopher Tancredi, Satoshi Tomioka, Sumiko Tonosaki, Yukiko Ueda, Michael Wagner, Iris Wu, Cristina Ximenes, Masashi Yamada, Tomohiro Yano, Kie Zuraw, and one anonymous reviewer for *Natural Language Semantics*. I fear that I didn’t list many people who have also made much contribution to the dissertation and my life. I hope that they know how much I appreciate their support.

I want to thank the staffs in the headquarters of the department for their generous help: Dan Giblin, Christine Graham, Mary Grenham, Stefanie Hanlon, Chris Naylor, and Jennifer Purdy.

Parts of Chapter 2 were presented with Danny Fox at SALT XV (UCLA, 2005) and the 28th GLOW Colloquium (University of Geneva, 2005). I also presented parts of Chapter 4 at NELS 33 (MIT, 2002), the workshop at the 20th National Conference of the English Linguistics of Japan (Aoyama Gakuin University, 2002), CUNY Syntax Supper (CUNY, 2004), the colloquium at Tohoku University (2004), and Generative Lyceum (Kwansei Gakuin University, 2004). Thanks go to audiences at these talks. An earlier version of Chapter 4 will appear in the following form: Takahashi, Shoichi. To appear. More than two quantifiers. To appear in *Natural Language Semantics*. © 2006 Springer Science and Business Media. I am greatly grateful to the publisher for its kind permission to reproduce copyrighted material in this dissertation.

During my first three years at MIT, I received financial support from Syoyu Club. I am very grateful to them for their support and encouragement.

I want to express my sincere gratitude to Yoko Hattori for her love, faith in me, boundless support and patience. Finally, I wish to thank my parents, Masami and Kazuko, and my sister, Mika, for their love and support. This dissertation is dedicated to my family.
# Contents

1 Introduction ............................................. 14
   1.1 Overview of the Dissertation. ....................... 14
   1.2 Approaches to Scope Widening and Scope Narrowing. .... 20
      1.2.1 Syntactic Approaches. .......................... 20
      1.2.2 Semantic Approaches. .......................... 26

2 MaxElide and the Re-binding Problem ................... 34
   2.1 Introduction. ......................................... 34
      2.1.1 The Goal. ....................................... 34
      2.1.2 The Proposal in a Nutshell. ..................... 35
   2.2 The Re-binding Problem. ............................... 37
      2.2.1 Evidence for the Sag-Williams Position. .......... 37
      2.2.2 Evidence against the Sag-Williams Position. .... 38
   2.3 The Re-binding Generalization. ......................... 39
      2.3.1 MaxElide. ....................................... 39
      2.3.2 Circumvention of MaxElide. ...................... 40
   2.4 The Proposal. .......................................... 42
      2.4.1 An Outline. ...................................... 42
2.4.2 A Theory of the Parallelism Condition ................. 42
2.4.3 The Proposal ........................................ 45
2.4.4 A Constraint on the Assignment of Variable Names ........ 47
2.5 Further Evidence ........................................ 48
2.5.1 Re-binding as a Consequence of Covert Movement ........ 48
2.5.2 The Effects of MaxElide in Sluicing ....................... 51
2.5.3 The Position of Intervening Focus ........................ 52
2.5.4 Syntactic Licensing Conditions .......................... 53
2.6 Evidence for Variables ..................................... 54
2.7 Further Issues ............................................. 56
2.7.1 Focused Traces ......................................... 56
2.7.2 The A/A’-distinction .................................... 57
2.8 Residual Issues ............................................. 59
2.8.1 Jacobson’s Problem ...................................... 59
2.8.2 Hardt’s Problem ......................................... 60
2.8.3 A Note on Antecedent-Contained Deletion ................. 61
2.9 Conclusion .................................................. 62

3 Wholesale Late Merge ........................................ 64
3.1 Introduction .................................................. 64
3.2 The Scope Narrowing Puzzle ................................. 66
3.2.1 Scope Narrowing in A-movement and A’-movement ........ 66
3.2.2 Condition C Bleeding in A-movement ...................... 70
3.3 The Proposal .................................. 73
  3.3.1 An Outline. ................................ 73
  3.3.2 The First Component: Late Merge. ................. 74
  3.3.3 The Second Component: A Late Merge Analysis of
          Extraposition. ................................ 77
  3.3.4 The Third Component: Trace Conversion. .......... 80
  3.3.5 Wholesale Late Merge. .......................... 84

3.4 Further Evidence from Ellipsis ......................... 90
  3.4.1 An Outline. ................................ 90
  3.4.2 A Late Merge Analysis of ACD. ..................... 90
  3.4.3 The Lexical Content Effect in Ellipsis. ............. 95
  3.4.4 The Complement/Adjunct Asymmetry. ................. 100

3.5 A Wholesale Late Merge Approach to the ACD Puzzle ........ 103
  3.5.1 An Outline. ................................ 103
  3.5.2 The ACD Puzzle. ................................ 104
  3.5.3 The First Component: A Raising Structure of
          Relative Clauses. ............................... 108
  3.5.4 The Second Component: Wholesale Late Merge. ...... 111
  3.5.5 The Third Component: Null Operator Movement. .... 114
  3.5.6 Further Issues. ................................ 118

3.6 A Note on Case. .................................. 124

3.7 Conclusion ....................................... 127
4 More Than Two Quantifiers

4.1 Introduction .............................................................. 128
  4.1.1 The Goal ........................................................... 128
  4.1.2 The Proposal in a Nutshell .................................... 128

4.2 The Comparative QP Scope Puzzle .................................. 131

4.3 The Plot ........................................................................ 134
  4.3.1 An Outline ........................................................... 134
  4.3.2 The First Component: A Decompositional Approach to
      Comparative QPs ..................................................... 134
  4.3.3 The Second Component: A Locality Condition on
      SSOs ..................................................................... 138
  4.3.4 The Third Component: An Intervention Constraint ....... 142
  4.3.5 An Account of the Comparative QP Scope
      Generalization ......................................................... 146

4.4 Wide Scope of Comparative QPs ..................................... 147

4.5 Implications for the Architecture of Grammar ................. 152

4.6 Towards Explaining the Heim-Kennedy Constraint ............ 156
  4.6.1 An Outline ........................................................... 156
  4.6.2 Scopal Commutativity: Monotone Increasing
      Comparative QPs ..................................................... 157
  4.6.3 Scopal Commutativity: Monotone Decreasing
      Comparative QPs ..................................................... 162
  4.6.4 Cross-Polar Anomaly .............................................. 168
4.6.5 Interim Conclusion. .............................. 172

4.7 The Heim-Kennedy Constraint or Scope Economy? ............... 172

4.7.1 An Outline. ................................... 172

4.7.2 Monotone Increasing Comparative QPs in
Subject Position. .................................. 173

4.7.3 Monotone Decreasing Comparative QPs in
Subject Position. .............................. 178

4.8 Further Issues. ...................................... 181

4.8.1 More on the Heim-Kennedy Constraint. ................ 181

4.8.2 More on Intervention Constraints. .................... 185

4.8.3 A Puzzle of QR. ................................. 188

4.9 Conclusion and Outlook. .................................. 191

5 Conclusions and Outlook ................................. 194
Chapter 1

Introduction

1.1 Overview of the Dissertation

This dissertation explores various issues pertaining to the nature of linguistic representations that serve as inputs to the semantic computation and approaches those issues specifically by scrutinizing the scopal behavior of quantificational expressions under diverse circumstances. To set the stage for formulating the questions that this dissertation investigates, let me first introduce two scopal phenomena to which much attention has been paid in syntactic and semantic studies of quantificational expressions. First, quantificational phrases (QPs) can take wider scope than their surface position, which I call Scope Widening.¹ For instance, the sentence in (1) illustrates the effects of Scope Widening.

(1) Some fifth year student read every book in the library.

(∃>∀) (∀>∃)

This sentence can be truthfully uttered even if there is no single fifth year student that read every book in the library as long as every book was read by a (possibly different) fifth year student. This interpretation is obtained when the object QP outscopes the

¹ A QP α takes its scope at a position β when a constituent that is sister to β is interpreted as an argument of α semantically.
subject QP. Second, it is also well-known that QPs can take narrower scope than the position where they are pronounced. One representative case that demonstrates this phenomenon is illustrated in (2) (May 1977, 1985, among many others). The sentence in (2) has a reading which is true if most of the lottery tickets were sold to Boston residents, so that it is likely that one of the residents who bought some wins the lottery. This interpretation is derived if the subject QP takes narrow scope relative to the predicate likely. In what follows, this scopal behavior of QPs is referred to as Scope Narrowing:\(^2\)

\[(2)\quad \text{Someone living in Boston is likely to win the lottery.}\]

\[(\exists \text{living in Boston}) (\text{likely} > \exists)\]

It is clear that the scope of QPs is not overtly indicated in a surface representation when the QPs exhibit the effects of Scope Widening and Scope Narrowing. However, the information about their scope needs to be embodied in a linguistic representation that serves as an input to the semantic component in some way or other. Given this, I formulate the following research questions, to which this dissertation makes an attempt to provide possible answers:

I. How is the information about the scope of QPs embodied in the linguistic representations that serve as inputs to the semantic computation?

II. What is the nature of the mechanism for generating those linguistic representations?

To set the stage for exploring these questions, the dissertation begins with discussing the previously proposed approaches to Scope Widening and Scope Narrowing. Section 1.2 will be devoted to introducing two major approaches to these scopal phenomena, which I will call syntactic and semantic approaches. The essential claim made in syntactic approaches is that both Scope Widening and Scope Narrowing are direct consequences of movement operations that apply to QPs. In syntactic approaches,

---

\(^2\) Scope Narrowing is sometimes called reconstruction in the literature.
the effects of Scope Widening are produced by applying upward movement to QPs, which is known as Quantifier Raising (QR). Similarly, downward movement of QPs, which is called Quantifier Lowering (QL), is one of the syntactic approaches to the effects of Scope Narrowing. Thus, the scope of QPs is structurally represented in syntactic representations, which are sometimes called Logical Forms (LFs). As we will see below, LF representations that involve movement chains are straightforwardly interpreted by assuming that movement establishes a variable binding dependency, which relates a moved constituent and its original position. In these respects, semantic approaches are strikingly different from syntactic ones. In semantic approaches, the effects of Scope Widening and Scope Narrowing are derived by applying semantic rules (i.e., type shifting rules) to linguistic representations that do not involve movement of QPs. Consequently, movement and, hence, variable binding need not to be postulated in grammar at least for the purpose of dealing with these scopal effects.

One of the primary goals of Chapter 2 is to demonstrate that when QPs exhibit the effects of Scope Widening, the LF representations from which the effects are derived involve a variable binding dependency. This claim is made on the basis of the fact that in certain contexts of ellipsis, the biggest deletable constituent must be deleted (Merchant to appear). Chapter 2 argues that the contexts where the maximization effects are observed can be characterized straightforwardly by making reference to the presence of variables within an ellipsis site that partake in a binding dependency of a particular sort, which we will call Re-binding. We will see that Re-binding has the consequence of making the constituent that satisfies the licensing condition on ellipsis relatively big. We will propose an analysis that accounts for the maximization effects by making crucial use of this consequence of Re-binding. Particularly important for the purpose of comparing syntactic and semantic approaches is the fact that the maximization effects are observed in the context where QPs exhibit the effects of Scope Widening. This suggests that the operation responsible for producing the effects of Scope Widening leads to the creation of a Re-binding configuration. This is expected in syntactic approaches where the scope

3 Alternatively, the effects of Scope Narrowing can be derived as consequences of the copy theory of movement, as we will discuss in 1.2.1.
4 Chapter 2 is a slightly modified version of Takahashi and Fox (to appear).
of QPs is determined by movement that generates a variable binding dependency. As we will discuss, if the effects of Scope Widening were instead captured by applications of type shifting rules, as in semantic approaches, it would be less clear how to characterize the contexts where the maximization effects are observed. Consequently, if the arguments in Chapter 2 are successful, they can be taken as evidence in favor of syntactic approaches.

In syntactic approaches, wide scope of the object QP relative to the subject QP in (3)a is derived from an LF representation close to (3)b.

(3)  
\begin{enumerate}
  \item Some fifth year student read every book.
  \item \([\text{every book}]_{1}\text{ [some fifth year student read }t_{1}]\)
\end{enumerate}

Chapter 3 looks into properties of variable binding dependencies more closely, especially focusing on the issue of how the tail of the dependencies is represented (i.e., the position represented as \(t_{1}\) in (3)b). It has been extensively argued, since Chomsky (1993, 1995), that movement is a merger operation of a single constituent in different positions, as illustrated in (4).

(4) \([\text{every book}]\text{ [some fifth year student read }\text{every book}]]\]

This view of movement is now known as the copy theory of movement. The copy theory of movement is theoretically simple, and it has also been supported by various empirical observations. Furthermore, the copy theory of movement has a consequence for dealing with the effects of Scope Narrowing. Capitalizing on the fact that there is more than one copy of a moved QP in a structure, it is possible to derive the effects of Scope Narrowing by deleting the higher copy, and choosing a lower one to be interpreted in the semantic component, as illustrated in (5). (The doubly strikeout material is intended to represent constituents invisible to the semantic component.)

(5) \([\text{someone living in Boston}],\text{ is likely to [someone living in Boston]},\text{ win the lottery}]\)
However, several facts have been pointed out in the literature as challenges to this view of movement. They seem to indicate that movement sometimes leaves a contentless trace, as in (3)b, instead of a copy of a moved element, like in (4) and (5). Chapter 3 develops a new theory of movement, which accounts for the challenging facts in a way consistent with the copy theory of movement. As a first step, I suggest extending Lebeaux’s (1988) idea that adjuncts can be merged counter-cyclically. Chapter 3 argues that the counter-cyclic merger is also applicable to constituents other than adjuncts. More specifically, it claims that a determiner can move alone, and a restrictor constituent of the determiner, which is an argument of the determiner, can merge with the determiner counter-cyclically. It is certainly the case that a chain created by movement of a determiner alone is not an interpretable syntactic object. However, the interpretability issue is more general in that the movement chain of the QP in (4) is not interpretable, either. To derive interpretable syntactic objects from those uninterpretable movement chains, I adopt the procedure proposed in Fox (1999a, 2002), which converts copies that are not a head of a chain into definite descriptions (see also Sauerland 1998, 2001, 2004; cf., Fox 2003). We will see that if this procedure applies to a copy of a determiner that does not involve its restrictor constituent (i.e., a copy left behind by movement of a determiner alone), it converts the copy into a syntactic object that receives the same interpretation as that assigned to a trace. Chapter 3 argues for one constraint on the application on movement of a determiner alone, and demonstrates that with this constraint, the proposed analysis can account for the impression that movement sometimes leaves a trace in a manner compatible with the copy theory of movement. If the discussion in Chapter 3 is successful, we can adopt the simplest theory of movement, and pursue the copy theory based approach to Scope Narrowing, as depicted in (5).

Contrary to what one might expect based on cases such as (3)a, scopal interaction between two QPs are not totally free, as has sometimes been pointed out in the past literature (Beghelli 1995, Beghelli and Stowell 1997, Liu 1990, 1997, among many others). The main empirical concern of Chapter 4 is one case that demonstrates this
point: if so-called comparative QPs (CQPs), such as *more than three books*, are in object position, as in (6), wide scope of the object QP relative to the subject QP is unavailable.\(^5\)

(6) Some fifth year student read more than three books.

\((\exists>CQP) *(CQP>\exists)\)

Facts of this sort are unexpected within the framework that I have discussed so far (or within the semantic alternatives that I am aware of), and suggest that the mechanism for assigning scope to QPs should be constrained. Exploring the scopal behavior of comparative QPs, Chapter 4 argues for several constraints on such a mechanism, and explores how the mechanism works. In Chapter 4, I will pursue a decompositional approach to comparative QPs proposed in Hackl (2000). In this approach, comparative QPs are decomposed into two quantificational elements (i.e., the comparative operator *-er than three* and the DP *many books* in the case of *more than three books*). Therefore, there are indeed three QPs in the structure of (6). Let us consider here the three scopal relations of the subject QP in relation to the two other QPs, which are given in (7). (There are three more scopal relations that the three QPs could participate in. However, they are excluded for an unrelated reason. See Chapter 4 for details.) As we will see, wide scope of the comparative QP is produced from narrowest scope of the subject QP in (7)c. Since the scopal relations in (7)a and in (7)b end up producing truth-conditionally the same interpretation, the Scope Shifting Operation (SSO) that derives the scopal relation in (7)b is prohibited by Fox’s (2000) Scope Economy, which requires SSOs to produce a new interpretation.

(7) a. step I: some>-er>many (widest scope of the subject)

b. step II: -er>some>many (intermediate scope of the subject)

c. step III: -er>many>some (narrowest scope of the subject)

\(^5\) Chapter 4 is a slightly modified version of Takahashi (to appear).
I will argue that since the step II in (7)b is ruled out by Scope Economy, a longer instance of SSO is ruled out. Therefore, only the wide scope reading of the subject QP is available in (6). If this argument is successful, it can be taken as an argument in favor of a derivational approach to the mechanism that determines the scope of QPs.

I close this section by providing brief answers to the questions raised above, which are repeated below.

I. How is the information about the scope of QPs encoded in the linguistic representations that serve as inputs to the semantic computation?

II. What is the nature of the mechanism for generating those linguistic representations?

As for the first question, this dissertation argues that the scope of QPs is syntactically represented in LF representations. An answer to the second question is that the mechanism for generating LF representations is derivational in nature.

1.2 Approaches to Scope Widening and Scope Narrowing

1.2.1 Syntactic Approaches

Much research has been executed in attempting to understand the nature of the effects of Scope Widening and Scope Narrowing. In this subsection, I discuss one theory of these scopal phenomena. In these approaches, which I call syntactic approaches, both Scope Widening and Scope Narrowing are produced as direct consequences of syntactic movement, and the scope of QPs defined in fn. 1 is structurally embodied in LF representations that serve as inputs to the semantic component. It is argued that QPs

---

6 I will propose a constraint that requires the scopal relation in (7)b to be established before the scopal relation in (7)c is generated.
undergo movement from their argument position to a position where they take their scope, and this movement operation establishes a variable binding dependency, which relates the moved QPs to their argument positions. As we will see shortly, LF representations derived this way are interpreted by a set of semantic rules.

In order to illustrate syntactic approaches, let me begin the discussion with presenting one representative argument in favor of the claim that the scope assignment to QPs is accomplished by movement. It has been argued that SSOs that produce the effects of Scope Widening are governed by the same set of constraints as other movements, such as wh-movement. More specifically, island constraints come into force in these two types of operations. For instance, they are subject to the coordinate structure constraint (CSC), which prohibits extraction out of only one conjunct (Ross 1967). In cases of wh-movement, a violation of the CSC makes sentences unacceptable, as in (8). Similarly, the CSC diminishes the number of scopal possibilities in cases of scopal interaction. Thus, wide scope of the universal QP relative to the subject QP is absent in (9) because movement of the universal QP over the subject QP violates the CSC (see also Rodman 1976 for relevant facts).

(8) *[Which student], did the professor [[admire _t1] and [despise the Dean]]?

(9) Some professor [[admires every student] and [despises the Dean]].

(May 1985:59)

The parallelism between wh-movement and SSOs goes even further. The CSC is circumvented in wh-movement if there is a pronoun in the second conjunct, which is bound by a wh-phrase in the first conjunct, as illustrated in (10). Ruys (1993) argues that wide scope of a QP in the first conjunct relative to a subject QP becomes available in

---

7 One qualification is in order here. As is often discussed, scope of certain types of indefinites is not restricted by island constraints. However, it is claimed that the scope assignment of indefinites of this sort is accomplished by a linguistic tool that does not rely upon movement, for example, by choice functions (see Kratzer 1998, Reinhart 1997, Winter 1997, among others, and see also 2.5.2 for relevant discussion).
exactly the same environment, as illustrated in (11) (see also Fox 2000 and Lin 2001 for relevant discussion and further data).

(10) a. *Which student [[likes which professor] and [hates the dean]]?
b. Which student [[likes [which professor],] and [wants him, to be on his committee]]?

(Fox 2000:53 and Ruys 1993)

(11) a. A student [[likes every professor] and [hates the dean]].
   \( (\exists > \forall) \; * (\forall > \exists) \)
b. A student [[likes [every professor],] and [wants him, to be on his committee]].
   \( (\exists > \forall) \; (\forall > \exists) \)

(Fox 2000:52 and Ruys 1993)

The parallelism discussed above has been taken as good ground for the claim that the SSOs that derive the effects of Scope Widening are movements of QPs.

Based on the fact of this sort (and others), May (1977, 1985) develops one version of a syntactic approach in which the effects of Scope Widening are produced by QR. As illustrated in (12)b, the object QP undergoes QR to its scope position and leaves a trace in its original position.

---

8 I cannot discuss other forms of syntactic approaches proposed in the past literature here. However, one remark should be made on a particular syntactic approach that does not assume QR. Hornstein (1995) and Kitahara (1996) argue that the effects of Scope Widening are derived as consequences of DP movement for reasons of Case assignment (see also Pica and Snyder 1995). However, various empirical problems have been raised against approaches along this line, which suggests that they should not be viewed as a general theory of the mechanism for assigning scope to QPs (see Fox 2000:46 and Kennedy 1997, among others).
(12)  a. Some fifth year student read every book in the library.

\[ \exists \forall \rightarrow \forall \exists \]

b. \[
[\text{every book in the library}] \quad [\text{some fifth year student read it}]
\]

Furthermore, May argues that in order for QPs to have narrower scope than their surface position, they undergo QL, as indicated in the representation in (13)b.

(13)  a. Someone living in Boston is likely to win the lottery.

\[ \exists \rightarrow \text{likely} \rightarrow \exists \]

b. \[
[\quad \text{is likely to} \quad [\text{someone living in Boston] win the lottery}]
\]

May’s QR approach to Scope Widening fits rather nicely with the widely adopted syntactic theory in which movement is always upward. However, his treatment of Scope Narrowing is less convincing in this respect since it resorts to downward movement, which is empirically not attested in any evident form. For this reason (and others, some of which we will discuss in the subsequent chapters), an alternative approach to Scope Narrowing has been developed within the so-called copy theory of movement framework (Chomsky 1993, 1995). In this theory, movement is considered to be a merging process of a single constituent in different syntactic positions. As a consequence of this theory, the scope of a QP can be determined simply by choosing which copy of the QP to bear quantificational force (the copy theory approach to Scope Narrowing). The (slightly simplified) LF representation in (14)a postulated for (13)a involves two copies of the indefinite subject as a consequence of its movement. The copy theory approach derives narrow scope of the subject QP relative to the predicate likely by deleting the upper copy and utilizing the lower one for the purpose of the semantic computation, as shown in (14)b.
It has been claimed that the copy theory approach to Scope Narrowing is theoretically desirable in that it can do away with downward movement (see Chapter 3 for further discussion).

Before leaving this subsection, I illustrate how the semantic component interprets LF representations which embody the scope of QPs by making essential use of movement dependencies. Throughout this dissertation, I assume Heim and Kratzer (1998), which proposes a set of semantic rules to interpret movement chains. Heim and Kratzer suggest that movement introduces a numerical index right below a moved element, as illustrated in (15). The introduced numerical index serves as a variable binder that binds co-indexed variables within a constituent sister to the index:

9 It is not necessary to assume that numerical indices must be present at LF representations in order to interpret movement dependencies in the semantic component (for example, see Fox 2003 for an alternative approach). See Tomioka 2001 for relevant discussion, and also 2.7.2 for a fact that might be taken as an argument in favor of the presence of variable binders in LF representations.

10 For perspicuity, I assume the trace theory of movement here. As we will discuss in 3.3.4, we need an additional rule to interpret movement dependencies within the copy theory of movement framework (Fox 2002 and Sauerland 1998, 2001, 2004).
The constituent labeled “Σ” in (15) is interpreted as a λ-predicate by Predicate Abstraction in (16). Indexed traces are interpreted by Trace Rule in (17). Since these semantic rules crucially refer to indices, we need to assume another semantic tool, namely, variable assignments. The definition of variable assignments is provided in (18), and variable assignments are modified in the course of semantic computation, as illustrated in (19).

(16) Predicate Abstraction (PA)

Let α be a branching node with daughters β and γ, where β is the numerical index $i$.

Then, for any variable assignment $g$, \( \llbracket \alpha \rrbracket^g = \lambda x \in \mathbb{D}_e. \llbracket \gamma \rrbracket^{x/i} \).

(Heim and Kratzer 1998:186)

(17) Traces Rule

If α is a trace, $g$ is a variable assignment, and $i \in \text{dom}(g)$, then \( \llbracket \alpha \rrbracket^g = g(i) \).

(Heim and Kratzer 1998:116)

(18) A variable assignment is a function from the set of natural numbers into $\mathbb{D}_e$.

(adapted from Heim and Kratzer 1998:111)

(19) Let $g$ be a variable assignment, $i$ be a natural number, and $x \in \mathbb{D}_e$.

Then $g^{x/i}$ is the unique assignment which fulfills the following conditions:

a. \( \text{dom}(g^{x/i}) = \text{dom}(g) \cup \{i\} \),

b. \( g^{x/i}(i) = x \), and

c. for every $j \in \text{dom}(g^{x/i})$ such that $j \neq i$: \( g^{x/i}(j) = g(j) \).

(Heim and Kratzer 1998:112)

Given these semantic rules, the semantic component can derive the meaning of the LF representation in (15) by Functional Application in (20).
If \( \alpha \) is a branching node and \( \{ \beta, \gamma \} \) the set of its daughters, then, for any assignment function \( g \), \( \alpha \) is in the domain of \( \{\}^g \) if both \( \beta \) and \( \gamma \) are, and \( \{\}^g \) is a function whose domain contains \( \{\}^g \).

In this case, \( \{\}^g = \{\}^g (\{\}^g) \).

(Heim and Kratzer 1998:105)

The abridged calculation of the denotation of (15) is illustrated below:

\[
\begin{align*}
\text{(21) } & \quad \text{[[TP [every book] [\exists 7 some fifth year student read t_7]]]}^g \\
= & \quad \text{[[every]}^g (\text{book})^g (\text{[7 [some fifth year student] read t_7]}^g) \quad \text{(FA)} \\
= & \quad \text{[[every]}^g (\text{book})^g (\lambda x. \text{[some]}^g (\text{fifth year student})^g (\text{x}^{1/7})) \\
& \quad \text{([[read]}^g (\text{[t_7]}^g x^{1/7})) \quad \text{(FA & PA)} \\
= & \quad \text{[[every]}^g (\lambda z. \text{z is a book}) (\lambda x. \exists y[y is a fifth year student and y read x])} \\
& \quad \text{(Lexical entries & \(\lambda\)-conversion)} \\
= & \quad \forall x[x \text{ is a book } \rightarrow \exists y[y \text{ is a fifth year student and y read x}] \\
& \quad \text{(Lexical entry & \(\lambda\)-conversion)}
\end{align*}
\]

In this subsection, we have seen that syntactic approaches can derive both the effects of Scope Widening and Scope Narrowing as consequences of the syntactic movement of QPs. We have also seen that LF representations that involve movement chains are straightforwardly interpreted in the semantic component by a set of semantic rules. In the next subsection, I will introduce alternative approaches to scopal phenomena, in which movement is not adopted as a linguistic tool to assign scope to QPs.

### 1.2.2 Semantic Approaches

In this subsection, we discuss alternative approaches to Scope Widening and Scope Narrowing. These approaches, which I call *semantic approaches*, are strikingly different
from syntactic ones in that semantic approaches do not employ movement of QPs in order to derive the effects of Scope Widening and Scope Narrowing. Instead, these effects are produced by applying type shifting rules to syntactic representations which do not involve movement of QPs. Consequently, movement of QPs and hence, variable binding dependencies need not to be postulated in grammar, at least for the purpose of assigning scope to QPs. It has sometimes been argued that semantic approaches simplify the architecture of grammar in the sense that they can dispense with some of the assumptions that syntactic approaches are committed to (see Jacobson 2002, among many others).

To explicate these claims, this subsection discusses one particular semantic approach developed in Jacobson (1999, 2000). Jacobson advocates a semantic approach within a framework in which variables play no essential role in grammar (variable-free semantics) and demonstrates that some of the facts that have been used as arguments in favor of the postulation of variables and LFs can be captured within this framework (most notably, Jacobson 1992a, 1999, 2000, to appear). Since the architecture of grammar assumed in variable-free semantics is strikingly different from the one postulated in syntactic approaches, an illustration of how the effects of Scope Widening and Scope Narrowing are captured within the variable-free semantics framework would serve as a good point of departure of comparing the two types of approaches.

To set the stage for discussing how Jacobson's approach derives the effects of Scope Widening and Scope Narrowing, let me begin with Jacobson's treatment of pronouns. As mentioned above, since grammar does not make any use of variables within this framework, a question arises about a semantic interpretation of pronouns since they are often assumed to be analyzed as variables. Jacobson (1999, 2000) claims that pronouns denote the identity function on individuals (assuming an array of possible presuppositions about gender, number, and person):

\[ \text{It is possible to postulate empty elements in syntactic representations, which accomplish the same result as applications of type shifting rules. The terms "syntactic approaches" and "semantic approaches" are used in the text for expository purposes, and they can be replaced with "movement approaches" and "non-movement approaches".} \]
(22)  \[ [\text{she}] = \lambda x. x \]

Under the assumption that pronouns are of type \( (e, e) \), as in (22), this raises a question about how pronouns combine with predicates whose argument slots are of type \( e \). Jacobson argues for a semantic rule to put them together, namely, the unary rule \( g \) in (23). The \( g \) rule passes up argument slots and keeps them unsaturated. Below is the calculation of the denotation of the sentence in (24)a where the subject is a “free” pronoun.\(^{12,13}\)

(23)  \textit{The g rule:}

\begin{align*}
\text{If } f \text{ is a function of type } (a, b) \text{ then } g_c(f) \text{ is a function of type } (c, (c, a), (c, b)), \\
\text{where } g_c(f) = \lambda V_c. f(V(C)). \\
\end{align*}

(adapted from Jacobson 2000:104)

(24)  \begin{align*}
\text{a. She left.} \\
\text{b. left; } \lambda x. x \text{ left} \\
& \quad \rightarrow g, \\
& \quad \lambda f_{(c, e)}. [\lambda y. \text{left}(f(y))] \\
\text{c. she; } \lambda x. x \\
\text{d. she left; } \lambda f_{(c, e)}. [\lambda y. \text{left}(f(y))](\lambda x. x) \\
& = \lambda y. y \text{ left}
\end{align*}

As is clear from (24)d, the sentence in (24)a does not denote a proposition in Jacobson’s approach, but rather denotes the set of individuals who left, which is semantically identical to the predicate \textit{left}. Jacobson claims that this is not problematic since this

\(^{12}\) For brevity, I will not discuss any syntactic aspect of Jacobson’s system.\(^{13}\) As Jacobson discusses, the \( g \) rule can be considered as a unary version of the function composition operator:

\begin{align*}
\text{(i)} & \quad \text{For any functions } h, f, g(h)(f) = h \circ f
\end{align*}

See Dowty (1988), Oehrle (1991), and Steedman (1987, 1990), among others, for an analysis of various constructions that can be accounted for by making use of function composition.
function applies to a contextually salient individual, and a propositional meaning can be obtained.

In addition to the g rule, we need to introduce another semantic rule in order to interpret QPs without postulating their movement. Since QPs are second order predicates whose semantic type is \((e, t), t\), they are interpretable only in a position sister to an element of type \((e, t)\). Therefore, QPs are not interpretable in object position of transitive predicates. In syntactic approaches, this type mismatch is resolved by movement of QPs, which leaves behind a variable of type e. It is clear that this strategy to solve type mismatch cannot be employed in Jacobson’s approach. Thus, Jacobson (1992b) embraces Partee and Rooth’s (1983) idea that predicates are listed as having their lowest possible type in the lexicon, but there is a rule which lifts the type of an argument slot from type e to \((e, t), t\) (see also Montague 1973).

With these semantic rules, we are now ready to illustrate how the effects of Scope Widening are produced by applying type shifting rules to linguistic representations that do not involve movement of QPs. The semantic computation of the object narrow scope reading in (25) is presented in (26). The argument lift rule first applies to the object argument slot in (26)a, and the rest of the calculation is carried out by applications in (26)b and (26)c.

(25) Some student read every book.

\((\exists \forall) (\forall \exists)\)
(26) Narrow Scope of the Object QP

a. read; \( \lambda v. [\lambda z. z \text{ read } v] \)
   \[ \rightarrow \text{ argument lift} \]
   \( \lambda P_{\langle e, 0, 0 \rangle}. [\lambda x. [P(\lambda y. x \text{ read } y)]] \)

b. read every book;
   \( \lambda P_{\langle e, 0, 0 \rangle}. [\lambda x. [P(\lambda y. x \text{ read } y)]](\lambda Q_{\langle e, 0 \rangle}. [\forall z [z \text{ is a book } \rightarrow Q(z)]])) \)
   \( = \lambda x. \forall z [z \text{ is a book } \rightarrow x \text{ read } z] \)

c. some student read every book;
   \( \lambda P_{\langle e, 0 \rangle}. \exists v [v \text{ is a student and } P(v)](\lambda x. \forall z [z \text{ is a book } \rightarrow x \text{ read } z]) \)
   \( = \exists v [v \text{ is a student and } \forall z [z \text{ is a book } \rightarrow v \text{ read } z]] \)

A detailed derivation of wide scope of the object QP (i.e., Scope Widening) is given in (27). First, the \( g \) rule applies in order to pass up the object argument slot in (27)a. Second, the unfilled object argument slot of type \( e \) is lifted to type \( \langle e, t \rangle, t \) in (27)b.\(^{14} \)

(27) Wide Scope of the Object QP

a. some student; \( \lambda P_{\langle e, 0 \rangle}. \exists v [v \text{ is a student and } P(v)] \)
   \[ \rightarrow \ g_e \]
   \( \lambda V_{\langle e, 0, 0 \rangle}. [\lambda x. \exists v [v \text{ is a student and } V(x)(v)]] \)

b. some student read;
   \( \lambda V_{\langle e, 0, 0 \rangle}. [\lambda x. \exists v [v \text{ is a student and } V(x)(v)]](\lambda y. [\lambda z. z \text{ read } y]) \)
   \( = \lambda x. \exists v [v \text{ is a student and } v \text{ read } x] \)
   \[ \rightarrow \text{ argument lift} \]
   \( \lambda P_{\langle e, 0, 0 \rangle}. P(\lambda x. [\exists v [v \text{ is a student and } v \text{ read } x]]) \)

c. some student read every book;
   \( \lambda P_{\langle e, 0, 0 \rangle}. [P(\lambda x. [\exists v [v \text{ is a student and } v \text{ read } x]])](\lambda Q_{\langle e, 0 \rangle}. \forall y [y \text{ is a book } \rightarrow Q(y)]) \)
   \( = \forall y [y \text{ is a book } \rightarrow \exists v [v \text{ is a student and } v \text{ read } y]] \)

\(^{14}\) There are other proposals to derive the effects of Scope Widening without movement of QPs. See Hendricks (1993) and Steedman (2000), among many others.
We have seen that the two interpretations of the sentence in (25) are produced without movement of QPs in Jacobson’s approach.

Now, I illustrate how this approach derives the effects of Scope Narrowing. Let us first consider the second sentence in (28) in which the pronoun is interpreted as a variable bound by the QP. Pronouns are interpreted as variables bound by a QP only when they are within the scopal domain of the QP. In syntactic approaches, the constituent that contains the pronoun is interpreted inside the scopal domain of the QP, as a consequence of movement.

(28)  [Not every boy], loves his, father, but his, mother, [every boy], loves.

However, Jacobson’s approach produces the effect of Scope Narrowing in (28) without making reference to movement. Before discussing how it does, we need to introduce a semantic rule that produces the effects of binding in variable-free semantics. Jacobson (1999, 2000) proposes that the effects of binding are obtained by applying the unary rule \( z \) to a predicate whose argument is a binder of pronouns:

(29)  \textit{The }z\textit{ rule:}

Let \( g \) be a function of type \((a, (e, b))\).

Then \( z(g) \) is a function of type \( ((e, a), (e, b)) \) such that \( z(g) = \lambda f \{ f(a) \} = \lambda x. g(f(x))(x) \).

(30)  Every man loves his mother.
Let me now explain how Jacobson's approach yields the effect of Scope Narrowing in (28). The \( g \) rule applies to pass up an object argument slot in (32)a, and the \( z \)-loves in (32)b produces the effect of the binding relation between the subject QP and the pronoun.\(^{15}\)

\(^{15}\)Cresti (1995) and Rullmann (1995) propose an alternative analysis of Scope Narrowing. Cresti and Rullmann claim that QPs can leave a trace of type \(<(e, t), t>\). If a QP leaves a trace of this type, the resulting structure would receive the same interpretation as that assigned to the structure without movement of the QP. Since in LF representations QPs do not occupy a position whose sister is interpreted as their argument in the semantic component and the effects of Scope Narrowing are generated by semantic rules, this approach can be taken as one of semantic approaches to Scope Narrowing. See Fox (1999a) and Romero (1998b) for arguments against this approach. Those arguments can also be taken as counterarguments to Jacobson's approach to Scope Narrowing.
(32) ... his [every boy] loves.

a. every boy; \( \lambda e \cdot \forall y [\text{y is a boy} \rightarrow Q(y)] \)

\[ \Rightarrow \lambda e \cdot \forall y [\text{y is a boy} \rightarrow Q(y)] \]

\[ = \lambda e \cdot \forall y [\text{y is a boy} \rightarrow Q(y)] \]

b. \( \lambda x \cdot [\text{x loves y}] \)

\[ \Rightarrow \lambda e \cdot [\text{x loves f(x)}] \]

c. every boy loves;

\[ \lambda e \cdot \forall y [\text{y is a boy} \rightarrow V(C)(y)](\lambda x \cdot \text{loves f(x)}) \]

\[ = \lambda e \cdot \forall y [\text{y is a boy} \rightarrow V(C)(y)] \]

d. his mother every boy loves

\[ \lambda e \cdot \forall y [\text{y is a boy} \rightarrow y \text{ loves C}(y)](\lambda x \cdot \text{the mother of x}) \]

\[ = \forall y [\text{y is a boy} \rightarrow y \text{ loves the mother of y}] \]

By specifically discussing Jacobson’s approach, we have seen that there are ways to derive the effects of Scope Widening and Scope Narrowing by applying a set of type shifting rules to representations without movement.
Chapter 2

MaxElide and the Re-binding Problem

2.1 Introduction

2.1.1 The Goal

In Chapter 1, we discussed two families of approaches to the effects of Scope Widening, namely, syntactic and semantic approaches. The current chapter argues for a syntactic approach on the basis of a maximization effect observed in ellipsis environments. As we will discuss, variables and variable names play an essential role in characterizing the situations where the maximization effect is visible. Particularly important for our purposes is the fact that effects of this sort are also found in ellipsis in which a QP shows the effect of Scope Widening (see 2.5.1). This chapter takes this fact as an argument for syntactic approaches, in which the scope of QPs is determined by movement and movement establishes a variable binding dependency.

An earlier version of this chapter will be published as Takahashi and Fox (to appear). However, any inadequacies in the content that is not included in Takahashi and Fox (to appear) are attributable only to me.

34
2.1.2 The Proposal in a Nutshell

It is well-known that ellipsis constructions involving bound variables are sometimes ambiguous.\(^1\) For instance, the sentence uttered by Speaker B in (1) can be understood either as a statement about John (Speaker B1, strict identity) or as a statement about Bill alone (Speaker B2, sloppy identity). (Angle brackets are used to mark the elided constituent (EC) and focused material is written in capital letters.)

(1) Speaker A: John admires his professor.
   Speaker B: BILL also does <admire his professor>.
   1. Bill admires John’s professor. (strict identity)
   2. Bill admires Bill’s professor. (sloppy identity)

One might suggest that sloppy identity results from a configuration in which a variable is free inside EC and is bound by an antecedent outside EC, as illustrated in (2).

(2) John \([vP \text{admires his professor}]\). Bill also does \(<EC \text{admire his professor}>\).

Such an analysis, however, raises questions about the semantic licensing conditions on ellipsis (henceforth, the parallelism condition). Suppose we adopt the rather simple condition proposed by Sag (1976) and Williams (1977), namely, that ellipsis of EC requires semantic identity with an antecedent constituent (AC). If sloppy identity involves the representation in (2), it is not obvious that the parallelism condition is met.\(^2\)

For this and other reasons (to which we will return), Sag and Williams argue that there is an alternative analysis of sloppy identity. They claim that the relevant variable is not free within EC in sloppy identity, but is instead bound by a λ-operator internal to EC.

---

\(^1\) As will be clear in section 2.6, the proposal in this chapter can be taken as an argument for a framework in which variables are postulated.

\(^2\) Specifically, if we assume, following Heim (1997), that the variables in (2) must bear different names, the Sag-Williams parallelism condition will not be satisfied: EC will not be semantically identical to AC (under every assignment function). Heim’s assumption is stated below in (28).
Specifically, they assume Partee’s (1975) Derived VP Rule, which introduces a λ-operator at the VP level. (Predicate Abstraction introduced in Chapter 1 can be employed to provide the same result.) Their analysis of sloppy identity is illustrated in (3). We will call such variable binding structures, in which all the variables are bound internal to the elided constituent, Internal-binding:

(3) **Internal-binding**

| John: \[AC \lambda x. \text{x admires x’s professor}\] |
| Bill also does \(<EC \lambda y. \text{y admires y’s professor}\> |

It is easy to see that Internal-binding in (3) allows EC and AC to be semantically identical. The more general prediction that follows from the proposals made by Sag and Williams is that the variable binding structure in (4), which we will call Re-binding, is never allowed.

(4) **Re-binding**

| Antecedent Clause: \[\ldots [XP_x \ldots [AC \ldots x \ldots]]\] |
| Ellipsis Clause: \[\ldots [YP_y \ldots <EC \ldots y \ldots>]\]

In Re-binding, there are variables that are free inside EC and AC, and there are binders outside EC and AC that bind the variables. The prediction that Re-binding is never possible follows from the simple parallelism condition introduced above (together with the assumption mentioned in fn., 2, and (28) below), and Sag and Williams support it by a variety of empirical observations. Counter-evidence, however, has been accumulated over the years (Evans 1988, Fiengo and May 1994, Jacobson 1992a, Lappin 1984, Merchant to appear, and Schuyler 2001, among others).

We argue that the empirical discrepancy in the past literature results from the fact that Re-binding is allowed, but is constrained. Building on a proposal made by Merchant (to appear), we claim that in Re-binding contexts, ellipsis must target the largest deletable
constituent (MaxElide). (See also Fiengo and May 1994:106-7 and Kennedy 2002 for relevant discussion.) The puzzling fact to be discussed is that the effects of MaxElide are observable only in Re-binding environments.

One could accommodate this fact by stipulation, namely, by a direct restriction of MaxElide to the relevant environments. We argue instead that MaxElide applies in those syntactic domains that are relevant for the evaluation of the parallelism condition on ellipsis. As we will see, the relevant syntactic domains have to be relatively big in Re-binding configurations. In other contexts, they can be as small as ECs themselves, thereby leading to the impression that MaxElide is not active.

The difference in the size of the relevant syntactic domains is a direct consequence of a Rooth-type theory of the parallelism condition (Rooth 1992b) when embedded in a system that makes use of variables and variable names. To the extent that our account is successful, it might provide an argument for such a system.

2.2 The Re-binding Problem

2.2.1 Evidence for the Sag-Williams Position

The Sag-Williams parallelism condition is not only conceptually natural, but is empirically supported, as well. The condition entails that sloppy identity is possible only when there is a way to analyze the relevant structure as one of Internal-binding (as in (3)). The Derived VP Rule makes such an analysis possible only when the understood antecedent of the relevant variable is the sister of EC. Sag and Williams claim that this is a good result. In (5)b, in which the embedded VP is deleted, sloppy identity is not possible, since it cannot result from an Internal-binding structure (in contrast to (5)a).

(5)  a. John said Mary hit him, and BILL also did $<\lambda x. \ x \ say \ Mary \ hit \ x>$.
    b. *John said Mary hit him, and BILL also $\lambda x. \ x \ said \ she \ did <\lambda y. \ y \ hit \ x>$.

(adapted from Sag 1976:131)
Sloppy identity in (5)b requires a Re-binding configuration: the variable x must be free within EC and can only be bound from a matrix position, either by the matrix subject or by the λ-operator introduced by the Derived VP Rule at the matrix VP level. The contrast in (6) makes the same point.

\[
(6) \quad \begin{align*}
\text{a. John is proud that there are pictures of him there, and BILL is } & \lambda x. \ x \\
& \text{proud that there are pictures of x there}, \text{ too.}
\end{align*}
\]

\[
\begin{align*}
\text{b. } & *\text{John is proud that there are pictures of him there, and BILL is } \lambda x. \ x \\
& \text{proud that there are } <\text{pictures of x there}, \text{ too.}
\end{align*}
\]

(Williams 1977:122)

The Sag-Williams claim receives further empirical support from the contrast in (7). The unacceptability of (7)b suggests that Re-binding as a consequence of a movement operation is not permitted, either (Merchant to appear, Sag 1976, Schuyler 2001, and Williams 1977).

\[
(7) \quad \begin{align*}
\text{a. John knows which professor we invited, but he is not allowed to reveal} \\
& \text{which one } <\lambda x. \text{ we invited } x>.
\end{align*}
\]

\[
\begin{align*}
\text{b. } & *\text{John knows which professor we invited, but he is not allowed to reveal} \\
& \text{which one } \lambda x. \text{ we did } <\lambda y. \text{ y invite } x>.
\end{align*}
\]

2.2.2 Evidence against the Sag-Williams Position

Contrary to the Sag-Williams position, it has been observed that Re-binding is sometimes allowed in ellipsis. In (8), sloppy identity is acceptable, just like in (5)a and (6)a, even though, in this case, it cannot result from an Internal-binding structure. (See Jacobson 1992a for relevant observations.)

\[
(8) \quad \text{John argued that Mary hit him, but BILL } \lambda x. \ x \text{ DENIED that she did } <\lambda y. \ y \text{ hit } x>.
\]
Similarly, Re-binding structures created by movement are also sometimes allowed (Evans 1988 and Fiengo and May 1994, among others):

(9) Mary doesn’t know who we can invite, but she can tell you who λx. we can NOT λy. y invite x.

What is the difference between (5)-(7) and (8)-(9) that might account for the fact that Re-binding is allowed only in the latter? A possibility that suggests itself is that the distinguishing property of (8) and (9) is the existence of focused material between the re-binder and the re-bound variable (Intervening Focus). In (8), the matrix verb, denied, is focused, and in (9), negation is. One might suggest that this Intervening Focus allows for Re-binding. In (5)-(7), by contrast, there is no Intervening Focus, and Re-binding is impossible. But, why should Intervening Focus be required for Re-binding? We will suggest a possible answer in the next section.

2.3 The Re-binding Generalization

2.3.1 MaxElide

The constraint in (10), suggested in Fiengo and May (1994), Kennedy (2002), and Merchant (to appear), might be relevant to our concerns.

(10) MaxElide (first pass):
   Elide the biggest deletable constituent.

In all of the unacceptable Re-binding cases, there is a bigger constituent that could have been deleted. In (7)b, repeated here as (11)a, the matrix VP could have been deleted, as illustrated in (7)a, repeated as (11)b. However, the situation is different in the acceptable
cases. In these cases, Intervening Focus blocks deletion of a bigger constituent, as illustrated in (11)c, where negation is focused.

(11) a. *John knows which professor we invited, but he is not allowed to reveal which one \( \lambda x. \text{we did } \langle \lambda y. \text{y invite x} \rangle \).
   b. John knows which professor we invited, but he is not allowed to reveal which one \( \lambda x. \text{we invited x} \).
   c. Mary doesn’t know who we can invite, but she can tell you who \( \lambda x. \text{we can NOT } \langle \lambda y. \text{y invite x} \rangle \).

2.3.2 Circumvention of MaxElide

Unfortunately, there are many counter-examples to MaxElide, as formulated in (10). One such counter-example is (12)a, which is acceptable despite the possibility of deleting a bigger constituent, (12)b.

(12) a. John said Mary likes Peter. BILL also said she does \( \langle \text{like Peter} \rangle \).
   b. John said Mary likes Peter. BILL also did \( \langle \text{say Mary likes Peter} \rangle \).

In response to the contrast between (11)a and (11)b, Merchant (to appear) restricts the application of MaxElide to cases in which EC contains an A’-trace:

(13) MaxElide (adapted from Merchant to appear):
    Elide the biggest deletable constituent if EC contains an A’-trace.

If MaxElide is to help us capture the Re-binding facts, it has to be further restricted, in order to accommodate cases of sloppy identity such as (5) and (6). The sentences in (5)b and (6)b should be ruled out by MaxElide, just like (11)a, due to the fact that deletion of a bigger constituent is possible ((5)a and (6)a). Thus, MaxElide might be modified as follows:
(14) MaxElide (minor modification):

Elide the biggest deletable constituent if EC contains a variable that is free within EC.

However, MaxElide needs to be modified further because the presence of a variable free within EC does not always require ellipsis of the biggest deletable constituent. In both (15)a and (15)b, EC contains a variable which is free within EC. Despite this fact, the possibility of deleting the bigger VP in (15)a does not exclude (15)b, in which the smaller VP is elided, unlike in Re-binding cases (e.g., (5)b, (6)b and (11)a).

(15) a. I know which puppy λx. you said Mary would adopt x and FRED did <say she would adopt x>, too.

b. I know which puppy λx. you said Mary would adopt x and FRED said she would <adopt x>, too.

The difference between (15) and cases of Re-binding is that in the former, a single element binds the variables which are free within AC and EC (i.e., which puppy in (15)). Configurations in which a single binder of this sort exists, configurations of Co-binding, are schematically represented in (16). The relevant distinction is that in Re-binding, unlike in Co-binding, the variables in EC and AC are bound by distinct binders, as illustrated in (4) above:

(16) Co-binding

\[ [... [XP_x ... [... x ...] and \langle EC ... x ... \rangle] \]

The circumvention of MaxElide in Co-binding environments suggests that it is necessary to further restrict MaxElide to Re-binding contexts:
(17) a. MaxElide (final descriptive version, to be modified):
    Elide the biggest deletable constituent in Re-binding configurations.

b. Re-binding:
    A structure in which EC dominates a variable that is free within EC and is
    bound by a binder YP outside EC, and there is no variable in AC bound by
    YP.

What remains puzzling is why such a complicated principle should be part of grammar.

2.4 The Proposal

2.4.1 An Outline

We argue that a complicated principle such as (17) does not exist. What exists, instead,
is the simple constraint that prohibits ellipsis of small constituents under all circumstances, i.e., something close to (10). However, we claim that this constraint, MaxElide, applies to particular syntactic constituents – to those constituents that are subject to the parallelism condition on ellipsis. As we will see shortly, these constituents have to be relatively big in Re-binding environments. In other contexts, they can be as small as ECs themselves, thereby leading to the impression that MaxElide is not active.

2.4.2 A Theory of the Parallelism Condition

As discussed above, the Sag-Williams parallelism condition always applies to EC and AC and requires that they be semantically identical. In Re-binding, which is represented again in (18), the presence of free variables within these constituents ensures that the condition will not be satisfied.
Indeed, any constituent in (18) will not be semantically identical to any constituent in the antecedent clause (besides the special case where $[[XP]] = [[YP]]$, in which $[[ZP]]$ might be identical to $[[WP]]$). However, $WP$, where the variable $y$ is bound, might bear a certain relationship to $ZP$ in the antecedent clause, a possibility that various researchers capitalize on. (See in particular Fiengo and May 1994 and Rooth 1992b.) Specifically, $WP$ might be semantically identical to $ZP$, modulo focus marked material. In Rooth's (1992a, b, 1996) terms, $ZP$ could be a member of the focus value of $WP$. Rooth (1992b) claims that this relationship, which might hold between constituents bigger than $EC$ and $AC$, licenses ellipsis of $EC$ in (18). What is going to be crucial is that this relationship, which might hold between $WP$ and $ZP$, cannot hold between any constituent smaller than $WP$ and any corresponding constituent in the antecedent clause. A condition on ellipsis along the lines of Rooth's (1992b) is formulated below:

(19) For ellipsis of $EC$ to be licensed, there must exist a constituent, which reflexively dominates $EC$, and satisfies the parallelism condition in (20).

We will call such a constituent a *Parallelism Domain*, or a *PD*:

\[ \text{XP reflexively dominates YP if XP dominates YP or XP = YP.} \]
Parallelism

PD satisfies the parallelism condition if PD is \textit{semantically identical to another constituent AC, modulo focus marked constituents.}

PD is semantically identical to AC modulo focus marked constituents, if there is a focus alternative to PD, PD_{Alt}, such that for every assignment function, g,

$$[[PD_{Alt}]]^g = [[AC]]^g.$$  

PD_{Alt} is an alternative to PD if PD_{Alt} can be derived from PD by replacing focus marked constituents with their alternatives.\(^4\)

To illustrate how a parallelism condition of this sort works, let us consider the contrast between (21) and (22), which indicates that Re-binding of a variable \(x\) in an elided constituent is possible only if a parallel variable in the antecedent is bound from a syntactically parallel position (the parallel dependency generalization is extensively discussed in May and Fiengo 1994 and Rooth 1992b).

(21) Johni asked his father to water hisi plants.
    BILL\(_j\) asked his BROTHER to <EC water hisj plants>.

(22) Johni asked his father to water hisi plants.
    *BILL asked [his BROTHER\(_j\)] to <EC water hisj plants>.

In (21), the matrix TP dominating EC can be taken as a PD because there is a focus alternative to this PD, PD_{Alt} (PD_{Alt} = John\(_i\) asked his father to water his\(_i\) plants), and PD_{Alt} and AC are semantically identical, modulo the focus marked material. Indeed, the matrix TP is the only constituent that can be a PD in (21), due to the presence of the re-bound variable. In contrast to (21), the parallelism condition is not met in (22) because the

\(^4\) Parallelism in (20) is satisfied if AC and PD are semantically identical. This is different from Rooth’s (1992a, 1996) Focus Interpretation Principle. We hope that our modification will not be problematic for other areas where the theory of focus interpretation is implemented.
matrix TPs are not semantically identical, modulo focus marked material, due to the fact that the variables are bound by binders in syntactically non-parallel positions. Notice that it would be difficult to capture the parallel dependency generalization within the Sag-Williams framework because there is no difference in the EC itself between (21) and (22). Thus, it can be taken as an argument for a parallelism condition of the above sort (Fiengo and May 1994 and Rooth 1992b).

2.4.3 The Proposal

Unlike the Sag-Williams parallelism condition, the size of constituents to which the parallelism condition applies can vary in a Rooth-type theory. Furthermore, the range of possible variation depends on the presence of a re-bound variable in EC and on the position from which it is bound. If there is such a variable, the PD must be a constituent that dominates the re-binder, i.e., it must be bigger than EC, as in (21). However, if there is no such variable, EC itself can be taken as the PD. We capitalize on this difference to capture the fact that MaxElide effects are observable only in Re-binding environments. Specifically, we propose that MaxElide, as defined in (23), applies in Parallelism Domains:

(23) MaxElide (our proposal):

Elide the biggest deletable constituent reflexively dominated by PD.

As mentioned above, if there is no re-bound variable involved, EC itself can be a PD. Thus, MaxElide is trivially satisfied both in (24)a and in (24)b, which have previously been taken to involve MaxElide circumvention. Notice that there is no bigger deletable constituent bigger than EC within the PD since in these cases, PD = EC:

5 Again, the assumption mentioned in fn. 2 enters into the picture: the re-bound variables must bear different names here. See 2.4.4 for further discussion.
(24)  a. John said Mary likes Peter. BILL also said she does \[\text{PD} \langle \text{EC}\ like\ Peter\rangle].
    b. John said Mary likes Peter. BILL also did \[\text{PD} \langle \text{EC}\ say\ Mary\ likes\ Peter\rangle].

In Re-binding contexts repeated here as (25)a, the embedded CP is the smallest possible PD for reasons stated more generally with respect to (18).

(25)  a. *John knows which professor we invited, but he is not allowed to reveal \[\text{PD}\ \langle\text{which one},\ we\ invited\ t\rangle].
    b. John knows which professor we invited, but he is not allowed to reveal \[\text{PD}\ \langle\text{which one},\ <\text{EC}\ we\ invited\ t}\rangle].
    c. Mary doesn't know who we can invite, but she can tell you \[\text{PD}\ \langle\text{who},\ we\ can\ NOT\ <\text{EC}\ invite\ t}\rangle].

MaxElide demands deletion of the biggest deletable constituent within this domain, which is the embedded TP.\(^6\) Thus, VP-ellipsis in (25)a is ruled out by MaxElide because it involves deletion of a smaller constituent than the embedded TP. As expected, the sentence in (25)b is acceptable since the embedded TP is deleted.\(^7\) Like in (25)a and (25)b, the embedded CP is the smallest PD in (25)c where Intervening Focus is involved. Since focus marked material cannot be deleted, the elided VP is the biggest deletable constituent in this sentence.

Consider now instances of Co-binding, repeated here in (26).

(26)  a. I know which puppy\(_1\) you said Mary would \[\text{AC}\ adopt\ t\] and FRED said she would \[\text{PD} \langle \text{EC}\ adopt\ t\rangle], too.
    b. I know which puppy\(_1\) you \[\text{AC}\ said\ Mary\ would\ adopt\ t\] and FRED did \[\text{PD} \langle \text{EC}\ say\ she\ would\ adopt\ t\rangle], too.

---

\(^6\) The embedded CP cannot be deleted for syntactic reasons. See 2.5.4 for relevant discussion.

\(^7\) In our analysis, the choice of PD determines the size of a constituent to be deleted. Sauerland (1998, 1999) claims that the choice of PD determines how much of a constituent should be focused. See also Sauerland (2000a, to appear) for relevant discussion.
In Co-binding, just like Re-binding, variables are free within EC and AC. However, in Co-binding, all variables are bound by a single binder. Thus, they share the same variable name. For this reason, the elided embedded VP can be semantically identical to AC in (26)a and, hence, can serve as a Parallelism Domain. Therefore, MaxElide is obeyed in (26)a, just like in (24)a. The possibility of deleting the bigger VP in (26)b does not preclude (26)a being grammatical since, in the latter, MaxElide applies to the embedded VP and, within this constituent, deletion is indeed maximal. 8

2.4.4 A Constraint on the Assignment of Variable Names

As we have seen in the last subsection, identity of variable names plays a crucial role in the account of the Co-binding facts. Given this, one might suggest that a structure analogous to Co-binding could be postulated in Re-binding by assigning the same variable name to the variables within AC and EC, as illustrated in (27). In this representation, EC and AC would be semantically identical because they involve variables with the same name, like in Co-binding. Consequently, the embedded VP would be a possible target for deletion, contrary to fact:

(27) *John knows which professor₁ we [invited t₁], but he is not allowed to reveal which one₁ we did [PD <EC invite t₁>].

For this reason, we are committed, like Sag and Williams, to the claim that identical variable names cannot be assigned to variables bound by distinct binders, as stated in (28). 9 Thus, there is no way to regard EC itself as a PD in Re-binding configurations:

---

8 One attractive implementation of our idea relies on the assumption that deletion can apply at the course of the derivation. In the non-Re-binding cases, a constituent could be elided before a bigger deletable constituent is created. In the Re-binding context, deletion cannot apply until a re-binder is introduced into the derivation, since the parallelism condition is not met before that stage of the derivation.

9 Kennedy (2004) proposes an alternative constraint on the assignment on variable names, which allows the representation in (27) and is thus inconsistent with our proposal.
(28) No Meaningless Co-indexation

If an LF contains an occurrence of a variable \( v \) that is bound by a node \( \alpha \), then all occurrences of \( v \) in this LF must be bound by the same node \( \alpha \).

(Heim 1997:202; see also Sag 1976:180)

Independently of the issue of MaxElide, No Meaningless Co-indexation is necessary in order to capture the parallel dependency generalization discussed in 2.4.2: Re-binding of a variable \( x \) in an elided constituent is possible only if the parallel variable in the antecedent is bound from a syntactically parallel position. If we did not assume this constraint, there would be no straightforward way to prohibit representations such as (29). Since EC itself is semantically identical to AC in (29), sloppy identity would be allowed, contrary to fact.

(29) Johni asked his father to [water hisi plants].

*BILL asked [his BROTHER]i to <EC water hisi plants>.

2.5 Further Evidence

2.5.1 Re-binding as a Consequence of Covert Movement

In the previous sections, we have discussed the effects of MaxElide when Re-binding results from overt \( wh \)-movement (i.e., the contrast between (25)a and (25)b). However, we should expect to find the same effects when a Re-binding configuration is created by other types of movement. We will see evidence that this expectation is borne out in the case of covert movement, specifically, Quantifier Raising (QR).

To set the stage, we first present a well-known argument that QR can create a Re-binding configuration. Hirschbühler (1982) discovered that an object can take wide scope over the subject in VP-ellipsis. This is illustrated in (30).
(30) A doctor treated every patient. A NURSE did <treat every patient>, too.

\[ (\exists \forall) (\forall \exists) \]

(Hirschbühler 1982)

Object wide scope is derived from the representations in (31) in which the object quantifier occupies a structurally higher position than the subject.

(31) Antecedent Clause: \([\text{every patient}_1 [\text{a doctor} [\text{treated} \text{t}_1]]] \]

Ellipsis Clause: \([_{\text{PD}} \text{every patient}_2 [\text{a nurse} <_{\text{EC}} \text{treated} \text{t}_2>]] \]

These representations involve variables that are bound by two distinct constituents (the object quantifiers). If the object of the elided verb is outside EC, (30) involves Re-binding.\(^{10}\) This means that the Parallelism Domain must include the re-binder, as depicted in (31). Since the subject, \(a \text{ nurse} \), in the ellipsis clause is focus marked in (30), the elided VP is the biggest deletable constituent. Thus, the Re-binding configuration that results from object wide scope is permitted in (30).\(^{11}\)

Given this discussion, we expect to find cases where object wide scope is ruled out by MaxElide. More specifically, we correctly predict object wide scope to be impossible in (32)a (observed by Williams 2003).

\(^{10}\) For theory internal reasons (consistency with the Sag-Williams theory), Hirschbühler suggests a Co-binding analysis of object wide scope, in which an object quantifier (e.g., \(\text{every patient} \) in (30)) undergoes across-the-board movement to a position structurally higher than the subjects of the two clauses. However, he also presents very strong counter-evidence to this analysis.

\(^{11}\) Sag (1976) and Williams (1977) claim that object wide scope is impossible in ellipsis on the basis of sentences like (i). They take this fact as further evidence for their position that Re-binding is never allowed:

(i) A doctor treated every patient. MARY did <treat every patient>, too.

\[ (\exists \forall) * (\forall \exists) \]

The contrast between (i) and (30) suggests that Re-binding as a consequence of covert movement is only allowed under certain circumstances (Fox 1995a, 2000 and Tomioka 1997, among others.). Chapter 4 discusses this issue in detail.
(32)  a. At least one doctor tried to get me to arrest every patient, and at least one nurse tried to get me to <arrest every patient>, as well.
   \((\exists > \forall) \neq (\forall > \exists)\)

   b. At least one doctor tried to get me to arrest every patient, and at least one nurse did <try to get me to arrest every patient>, as well.
   \((\exists > \forall) \neq (\forall > \exists)\)

   (Williams 2003)

As illustrated in (33), which is the representation of object wide scope for the ellipsis clause in (32)a and (32)b, there is only one possible PD, namely, the entire ellipsis clause. Since there is a possibility of deleting the bigger VP, (32)b, ellipsis of the smaller VP is impossible in (32)a: 12

\[
(33) \quad [PD \text{ every patient}, [at least one nurse [VP tried to get me to [VP arrest t]]]]
\]

Williams' (2003) interpretation of the facts is different. On the basis of the unavailability of object wide scope in (32)a, he concludes that Re-binding is in general impossible. 13 However, our alternative is corroborated by the fact that object wide scope is permitted even when the smaller VP is elided as long as Intervening Focus prevents ellipsis of the matrix VP, (34). We take this fact as further evidence for our claim that Re-binding is allowed, but is constrained by MaxElide:

---

12 The availability of subject wide scope in (32)a follows straightforwardly. Since the object quantifier adjoins to the most embedded VP by QR in this case, subject wide scope does not involve Re-binding. Thus, the most embedded VP (or TP, see 2.7.2) can be chosen as a PD and be targeted for deletion without violating MaxElide.

13 More specifically, he claims that in (30), the subject quantifier can undergo reconstruction into a VP-internal position, which is below the QRed object. Consequently, there is a way to analyze object wide scope with an Internal-binding structure. In contrast, a control predicate, try, is used in the matrix position in (32), which prohibits a subject from reconstructing into the complement clause.
(34) A doctor tried to arrest every patient, and a NURSE MANAGED to <arrest every patient>.

(∃∀)(∀∃)

2.5.2 The Effects of MaxElide in Sluicing

The effects of MaxElide are also observable in the sentences in (35), which do not involve (overt) movement in the antecedent clause (Merchant to appear and Schuyler 2001). Deletion of TP as in (35)a will be referred to as Sluicing here, following Ross (1969).

(35) a. I know we invited some professor, but I can’t remember which professor <we invited>.

b. *I know we invited some professor, but I can’t remember which professor we did <invite>.

The effects of MaxElide in (35) can be explained in the way advocated above since Sluicing involves Re-binding.

For concreteness, I assume Fox and Lasnik’s (2003) analysis of Sluicing, which adopts Reinhart’s (1997) choice function approach to indefinites and wh-phrases.\(^{14}\) Reinhart claims that both indefinites and wh-phrases induce choice function variables and their scope is assigned by existential closure.\(^{15}\) Given this approach, an indefinite remains in situ, and its choice function variable is bound by the existential quantifier, as illustrated in the antecedent clause in (36). Fox and Lasnik argue that the wh-phrase in the ellipsis clause introduces the same type of dependency. In the ellipsis clause in (36),

\(^{14}\) An alternative approach to Sluicing is that an indefinite undergoes QR in the antecedent clause, which establishes a dependency parallel to the one of wh-movement in the ellipsis clause (Chung, Ladusaw and McCloskey 1995 and Merchant to appear). Since this approach also assumes that Sluicing involves Re-binding, the argument in this subsection could be made within this framework, as well.

\(^{15}\) A function \(f\) of type \(\langle(e, t), e\rangle\) is a choice function \(CH\) iff for every non-empty set \(P\), \(f(P) \subseteq P\).
only the wh-word *which* remains in the derived position, and it functions as an existential quantifier over choice functions. This existential quantifier binds a choice function variable in the original position of the *wh*-phrase, as shown in (36).16

(36) Antecedent Clause: I know [\(\exists f \ CH(f) \ [\text{we invited } f(\text{professor})]\)]

Ellipsis Clause: I can’t remember [\(\text{PD which } g \ CH(g) [\text{we invited } g(\text{professor})]\)]

It is clear that the structures in (36) constitute the Re-binding configuration, due to the presence of the re-bound choice function variables. As a consequence, the entire embedded clause in the ellipsis clause in (36) counts as the smallest possible PD where the choice function variable is bound. Therefore, VP-ellipsis in (35)b is ruled out by MaxElide, due to the possibility of Sluicing in (35)a. This analysis of the contrast between Sluicing and VP-ellipsis above is corroborated by the grammatical VP-ellipsis in (37), which obeys MaxElide because of the presence of Intervening Focus (Merchant to appear and Schuyler 2001).

(37) It’s clear that they could invite some professor, but I don’t know which professor they ever WOULD <invite>.

2.5.3 The Position of Intervening Focus

We have argued that Intervening Focus is required for Re-binding. If there is Intervening Focus, MaxElide is satisfied as a consequence of the fact that ellipsis of any constituent dominating the Intervening Focus is blocked. However, it is crucial to notice that for Re-binding to be allowed, the *biggest* deletable constituent up to Intervening Focus has to be

16 Fox and Lasnik (2003) claim that the structure like (36) is derived by the general mechanism of reconstruction (see Chapter 3) within the copy theory of movement framework. As Fox and Lasnik discuss, *wh*-movement must take place in one fell swoop fashion in order to make the binding dependencies in the two clauses parallel (the parallel dependency generalization). See Fox and Lasnik (2003) for its consequences.
elided. Even in cases where Intervening Focus is involved, Re-binding should be impossible if the biggest deletable constituent is not elided.

This is indeed what we find in (38). Here, negation, not, is the relevant focus marked material. In (38), there are two deletable constituents, namely, the intermediate VP headed by agree and the most embedded VP headed by adopt. Ellipsis of the most embedded VP in (38)a is excluded by MaxElide, due to the possibility of deleting the intermediate VP, as demonstrated in (38)b.

(38) a. *I don’t know which puppy you should agree to adopt, but I know [PD which one1 you should NOT [VP agree to <EC adopt t1>]].
   b. I don’t know which puppy you should agree to adopt, but I know [PD which one1 you should NOT <EC agree to adopt t1>].

As illustrated in (39), deletion of the most embedded VP is permitted if Intervening Focus (refuse in (39)) occupies a structurally low position so that its existence prohibits ellipsis of any constituent bigger than EC.

(39) I don’t know which puppy you should agree to adopt, but I know [PD which one1 you should REFUSE to <EC adopt t1>].

2.5.4 Syntactic Licensing Conditions

As we saw above, the presence of focus marked material diminishes the number of candidates for ellipsis. (No dominating constituents can be deleted.) We suggest that certain syntactic licensing conditions on ellipsis have the same effect. As illustrated in (40)a, EC cannot be a constituent headed by the perfective auxiliary, have.

(40) a. *I don’t know which puppy you should agree to adopt, but I know [PD which one1 you should REFUSE to <EC adopt t1>].
(40)  a. *Sally might have eaten rutabagas, but Holly should NOT <have eaten rutabagas>.  
     (Johnson 2001:442)
     b. Sally might have eaten rutabagas, but Holly should NOT have <eaten rutabagas>.

The same is also true in Re-binding, (41)a. If the constituent headed by have could have counted as a candidate that is relevant for the evaluation of MaxElide, (41)b would have been ruled out since a smaller constituent is elided. Thus, the facts in (41) suggest that MaxElide is a principle that demands deletion of the biggest constituent that obeys other conditions of grammar, among them, the syntactic licensing conditions on ellipsis and the requirement that focus be overtly realized.

(41)  a. *I know which dish Sally should have eaten, but I don’t know [PD which one₁ she should NOT <have eaten t₁>].
     b. I know which dish Sally should have eaten, but I don’t know [PD which one₁ she should NOT have <eaten t₁>].

2.6 Evidence for Variables

In Re-binding, the Parallelism Domain must dominate the re-binder. Thus, the Parallelism Domain is necessarily bigger than EC in this configuration. In contrast, we have suggested that EC itself can be a Parallelism Domain in other contexts. In order to differentiate Re-binding from other cases, we have made crucial use of variables and variable names.

However, as discussed in 1.2.2, variables and variable names do not play any role in variable-free semantics (Jacobson 1992a, 1999, 2000, to appear, among others). Since pronouns are assumed to be identity functions, a pronoun does not make any contribution to the meaning of an elided VP in a sloppy identity construction such as (42) (see 1.2.2 for details). Therefore, the meaning of the embedded VP in (42) is the same as the meaning of the verb hit. Under this assumption, EC itself could be taken as the PD in
(42) (because it is semantically identical to the embedded VP in the antecedent clause). If this is a possible option, MaxElide would be satisfied in (42) and the sentence would be acceptable, contrary to fact.

(42) *John said Mary \[AC \text{hit him}\], and BILL also said she did \[PD <EC \text{hit}>\].

In 2.5.1, we have discussed the fact that object wide scope becomes impossible if the biggest deletable constituent is not elided (e.g., (32)a). We have argued that object wide scope in this case is ruled out by MaxElide, assuming syntactic approaches to Scope Widening in which the scope of QPs is determined by movement and assuming that movement establishes a variable binding dependency. Notice however that the effects of Scope Widening are produced by applications of type shifting rules in semantic approaches, including variable-free semantics, as discussed in 1.2.2. Since the structure from which object wide scope is derived would not involve Re-binding in semantic approaches, it would be less clear why the effects of MaxElide are visible in the context of Scope Widening.\(^\text{17}\)

Is there a formulation of MaxElide that will have the desired consequences within variable-free semantics and semantics approaches to Scope Widening? We do not know. (See Jacobson 2004a for a potentially relevant proposal.) If the answer is negative, we can take our results to argue in favor of systems that postulate variables and variable names and syntactic approaches to Scope Widening.

---

\(^\text{17}\) If movement traces were postulated within the variable-free semantics framework, they would also be semantically vacuous.
2.7 Further Issues

2.7.1 Focused Traces

Schuyler (2001) provides potential counter-examples to MaxElide. The sentences in (43) are problematic because they are acceptable even though there is a possibility of deleting a bigger constituent than EC (i.e., the embedded TP in (43)a and the matrix VP in (43)b).

\[(43)\]

a. ?I don’t know which puppy you should adopt, but I know which KITTEN, 
   [TP you should <adopt t1>].

b. The blue papers I think Pete signed, and the GREEN ones, I [VP think he 
   did <sign t1>], too.

(Schuyler 2001:8, 13)

Two properties are common to the potentially problematic cases: the re-binder is a phrase that has moved out of EC, and the re-binder bears contrastive focus. These properties seem crucial, as suggested by facts already discussed. The ungrammatical sentences in (44)a and (44)b lack the first and the second property, respectively.

\[(44)\]

a. *John, said Mary hit him, and BILL also said she did <hit him>.

b. *John knows which professor we invited, but he is not allowed to reveal 
   which one, we did <invite t1>.

We suggest, following Sauerland (1998) and Selkirk (1995), among others, that if an element in a head of a movement chain is focused (such as \textit{which kitten} in (43)a and \textit{the green ones} in (43)b), focus is optionally present within the trace. Together with the assumption that movement takes place successive-cyclically (e.g., moved phrases adjoin to every VP and CP), (43)b could involve the following representation:
(45) \[ \ldots [CP \text{which } KITTEN_1 [TP \text{you}_2 \text{should } [VP \text{which } KITTEN_1 [VP \text{you adopt which kitten}_2]]]] \]

Ellipsis of a larger constituent than EC is prohibited in (43) because the focused trace in the VP-adjoined position functions as Intervening Focus.\(^{18,19}\)

### 2.7.2 The A/A'-distinction

Lasnik (2001) suggests that when a subject is extracted by wh-movement, the possibility of deleting TP does not preclude VP-ellipsis, as illustrated in (46). Based on this fact, Merchant (to appear) claims that MaxElide applies only to cases in which EC involves A'-traces, as was stated above in (13):

(46) a. Someone solved the problem.
     1. Who <[TP t\text{A'} [VP t\text{A} solved the problem]]>?
     2. Who did <[VP t\text{A} solve the problem]>?

\(^{18}\) We suggest that focus marking on traces is optional. Thus, the trace in the complement of VP does not have to be focused, and the VP can be deleted. We also need to say that traces are interpreted in a way that would make sense of focus marking. See Fox (1999a, 2000, 2002) and Sauerland (1998, 2004). Chapter 3 discusses this issue in detail.

\(^{19}\) Romero (1998a) argues that a wh-phrase must be focused in Sluicing configurations if its corresponding constituent is an indefinite in an antecedent clause, as shown in (i)a. The fact that VP-ellipsis in (i)b is ungrammatical seems to indicate that this wh-phrase cannot leave a focused trace, which blocks deletion of TP.

(i) a. I know John met someone, but I don’t know WHO.
     b. *I know John met someone, but I don’t know WHO he did.

Assuming that the wh-phrase who is decomposed into the wh-part WH and an indefinite at LF, we suggest that the wh-part receives focus, and that since the wh-part is interpreted only in the Spec of CP, there is no constituent that functions as Intervening Focus within a Parallelism Domain. The LF representation postulated for (i) is given below:

(ii) I know [\exists f CH(f) [TP John met f(person)]]
     I don’t know [WH g CH(g) [TP he met g(person)]]

It is crucial that some material in the restrictor constituent of the determiner is focused in (43).
Unlike Merchant’s, our formulation of MaxElide in (23) predicts (46)b2 to be unacceptable if the subject wh-phrase is in the Spec of CP: the PD would be the entire CP. If, however, we assume that wh-movement of a subject to the Spec of CP is optional, both (46)b1 and (46)b2 could be explained in a way consistent with our formulation of MaxElide. If the subject moves to the Spec of CP, (46)b1 is derived. If wh-movement does not take place, the derivation ends up with VP-ellipsis, (46)b2.

Raising constructions like (47) seem to be more appropriate for investigating whether there is an A/A' distinction in the application of MaxElide.

(47) a. I know that John is likely to win the election, but I am not allowed to reveal that he is <likely to win the election>.
   b. I know that John is likely to win the election, but I am not allowed to reveal that he is likely to <win the election>.

We consulted with eleven native speakers about the sentences in (47). All of them found (47)a acceptable. In contrast, we found variation on the judgment on (47)b. Four of our informants found it acceptable, four, slightly degraded (one question mark), and three, ungrammatical. We assume that A-movement takes place successive cyclically and that movement introduces a λ-operator right below a moved constituent, as indicated in (48). We suggest that our informants who found (47)b acceptable could be taking the λ-predicate at the embedded TP level in (48) (λy) as a PD. Since the biggest deletable constituent within the PD is elided in (47)b, it is grammatical.

(48) I am not allowed to reveal that [he [λx. [is likely x [PD λy. [to y win the
election]]]]]

One might wonder about the ramifications of these assumptions for those ungrammatical cases that have been ruled out by MaxElide. More specifically, do these assumptions lead to the prediction that MaxElide is not violated in those cases (hence, incorrectly predicting that they should be grammatical)? The answer to this question is negative. To illustrate this point, let us consider one representative case that has been excluded by
MaxElide in (49)a. As illustrated in the LF representation in (49)b postulated for (49)a, (49)a involves two types of movement, namely, movement of the subject and wh-movement.

(49) a. *John knows which professor we invited, but he is not allowed to reveal which one; we did <invite _t>.
   b. ... [CP which one [TP λx. [TP we λy. [VP x λz. [VP y invite z]]]]]

Even if we assume that wh-movement takes place successive cyclically and that movement introduces a λ-operator, the λ-predicate at the TP level (λx) is the smallest possible Parallelism Domain in (49)b, due to the presence of traces of A-movement of the subject (y) and wh-movement (x and z). Thus, (49)a is once again ruled out by MaxElide.

### 2.8 Residual Issues

#### 2.8.1 Jacobson’s Problem

Pauline Jacobson (personal communication) points out that the effects of MaxElide are observable also in cases where Re-binding is not involved. The sentence in (50)a is unacceptable and this might be attributed to the possibility of deleting a bigger constituent than EC, as illustrated in (50)b. Notice that (50)a can be analyzed as an Internal-binding structure since all the variables are bound internal to the elided constituent. Since the Parallelism Domain could have been as small as EC itself in (50)a, MaxElide should have been satisfied.
a. *John loves every woman who skis, and BILL also loves every woman who does \(<\lambda x. \text{x ski}\>.

b. John loves every woman who skis, and BILL also does \(<\lambda x. \text{x love every woman who skis}\>.

(Pauline Jacobson, personal communication)

We do not have a solution to Jacobson’s puzzle. We hope that the correct solution lies in constraints on the choice of Parallelism Domains. Since a complex NP intervenes between two Parallelism Domains, PD1 and PD2, as illustrated in (51), we would like to explore constraints that refer to the presence of islands. However, detailed investigation of this issue needs to be left for future research.\(^{20}\)

\[
(51) [\text{PD1 BILL also love } [\text{island every woman [PD2 who skis]]}]
\]

### 2.8.2 Hardt’s Problem

Daniel Hardt (personal communication) observes a case of sloppy identity, which suggests that MaxElide is sometimes circumvented (see also Hardt 2004, 2005).

(52) Nearly every boy said Mary hit him. But, BILL \(\lambda x. \text{didn’t say she did } \text{hit } x\>.

(Daniel Hardt personal communication)

\(^{20}\) A related issue is the following facts observed by Merchant (to appear). Given the discussion so far, the smallest PD would be the embedded CP in (i). However, the matrix VP, which is larger than the smallest PD, needs to be deleted:

(i) a. Ben knows who she invited, but Charlie doesn’t \(<\text{know who she invited}\>.

b. ??Ben knows who she invited, but Charlie doesn’t know who \(<\text{she invited}\>.

c. ??Ben knows who she invited, but Charlie doesn’t know who she did \(<\text{invite}\>.

(Merchant to appear:9)

One possible approach to this problem would be to constrain the types of phrases that can be taken as Parallelism Domains. For example, CPs cannot be PDs, but VPs (and some other phrases) can. If constraints of this sort were postulated, the sentence in (i)a would be the only one that obeys MaxElide.
What is the difference between (52) and the Sag-Williams case repeated below?

(53) *John said Mary hit him, and BILL λx. also said she did \( \text{hit } x \).

One possible characterization of the facts is that when the subject in the ellipsis clause, which is a binder of a re-bound pronoun, is a member of the set of individuals which the quantificational element in the corresponding subject position in the antecedent clause quantifiers over, MaxElide is circumvented.\(^{21}\) We will need to explore an analysis that solves Hardt’s fact on another occasion.

### 2.8.3 A Note on Antecedent-Contained Deletion

There is a debate as to whether antecedent-contained deletion (ACD) involves Re-binding. (See Heim 1997 for discussion, and see also Chapter 3 for a Re-binding analysis of ACD) Sentences that bear on this debate in the context of MaxElide would be ones such as (54). The interpretation relevant for our purposes is object wide scope. In the representation that gives this interpretation, a PD would become relatively big so that the effects of MaxElide should be visible (see 2.5.1 for details). If (54)a is degraded under the wide scope reading of the object, a Re-binding analysis would be supported.

(54) a. At least one student tried to solve every problem that Prof. SMITH tried to \(<\text{solve}>\).

b. At least one student tried to solve every problem that Prof. SMITH did \(<\text{try to solve}>\).

---

\(^{21}\) Hardt (2004, 2005) proposes an alternative characterization: in cases where a constituent smaller than the biggest deletable constituent is deleted, if strict identity is possible, sloppy identity is prohibited. As far as we can tell, this formulation raises an issue concerning the effects of Intervening Focus. For instance, sloppy identity is possible in (8), repeated here as (i), contrary to what Hardt’s alternative characterization shall lead us to expect.

(i) John argued that Mary hit him, but BILL DENIED that she did \(<\text{hit him}>\).
Five out of ten native speakers with whom we consulted (54) found (54)a acceptable, two, slightly degraded, and three, unacceptable.  

This seems to suggest that ACD does not involve Re-binding. Heim (1997) proposes one non-Re-binding analysis of ACD (however, see Jacobson 1998 for evidence against Heim's approach). We will need to explore an analysis of ACD that is consistent with the absence of the effects of MaxElide in (54)a on another occasion.

2.9 Conclusion

We have claimed that MaxElide is a constraint that forces deletion of the biggest deletable constituent under all circumstances. We also claimed that this constraint applies to those constituents which are subject to the parallelism condition (Parallelism Domains). Parallelism Domains must be bigger than ECs in Re-binding contexts, but in other contexts, they can be the ECs themselves. This explains the fact that the effects of MaxElide are observable only in Re-binding contexts. The difference in the size of Parallelism Domains is a corollary of context-sensitive parallelism conditions (such as Rooth's 1992b) in a system that postulates variables and variable names. Thus, we have suggested that our results can be taken as evidence for such systems.

Particularly important for the comparison between syntactic approaches and semantic approaches to Scope Widening is the fact that the effects of MaxElide are observed in cases in which QPs exhibit the effects of Scope Widening, discussed in 2.5.1. This fact can be straightforwardly explained by the syntactic approaches, in which the scope of QPs is determined by their movement, which establishes a variable binding dependency. Thus, this chapter concludes that the presence of the effects of MaxElide in the context of Scope Widening is an argument in favor of the syntactic approaches.

---

22 For three speakers, the elided constituent in (54)b must be interpreted as solve, instead of try to solve.
Chapter 3

Wholesale Late Merge

3.1 Introduction

In Chapter 2, I argued for syntactic approaches to the effects of Scope Widening in which the scope of QPs is determined by movement in syntax which establishes a variable binding dependency. When combined with the copy theory of movement, syntactic approaches have a consequence for the treatment of the effects of Scope Narrowing, as discussed in 1.2.2. In this particular view of movement, these effects can be derived by deleting a higher copy of a QP and making a lower one involving quantificational force accessible to the semantic component. As opposed to the traditional trace theory of movement, the copy theory of movement is theoretically simple in that movement can be viewed as a merger operation of a single constituent in different positions (Chomsky 1993, 1995). It is empirically well-motivated, as well (Fox 1999a, Romero 1998b, and Sauerland 2004, among many others). However, some challenging facts to the copy theory of movement have often been pointed out, which seem to indicate that movement sometimes leaves a contentless trace, instead of a copy of a moved constituent. The most notable fact is that A-movement always bleeds Condition C, which is strikingly different from A’-movement, which bleeds Condition C only in limited cases (Chomsky 1993 and...
Lebeaux 1988, among others). Taking up this puzzle as a point of departure, this chapter develops a new theory of counter-cyclic merger of lexical material (Late Merge).

I suggest that the first step towards a possible solution to this challenging fact in A-movement is to elaborate Lebeaux’s (1988) analysis of the limited Condition C bleeding in wh-movement: wh-movement bleeds Condition C when a relevant R-expression is included in an adjunct which modifies a moved wh-phrase. Lebeaux claims that adjuncts can merge with a wh-phrase at any point of the derivation because they are optional in nature. As a consequence, adjuncts can Late Merge with a wh-phrase after wh-movement, and this opens the possibility of a derivation in which Condition C is not violated at any stage. Although Lebeaux’s original idea of Late Merge does not capture the asymmetry in the Condition C bleeding between A-movement and A’-movement, I will elaborate this theory and argue for an alternative view of Late Merge. The primary proposal in this chapter is that Late Merge is applicable to a wider range of constituents than considered so far. Capitalizing on the implication that results from Fox’s (2002) proposal of Late Merge, and its extension by Bhatt and Pancheva (2004, 2005), I specifically argue that restrictor constituents of determiners can merge with determiners counter-cyclically (Wholesale Late Merge). In other words, it is possible for determiners alone to undergo movement. Chains formed by movement of this sort are converted into interpretable syntactic objects by the general procedure that applies to copies, such as Fox’s (1999a, 2002) Trace Conversion (see also Sauerland 2001, 2004). As we will see, if determiners do not appear with a restrictor constituent in their complement, copies of determiners manipulated by Trace Conversion turn into syntactic objects that end up producing the same interpretation as that traditionally assigned to traces (contentless copies). I will demonstrate that the interplay of these syntactic operations (together with one constraint on the application of Wholesale Late Merge) leads us to the impression that movement sometimes leaves a trace. If the arguments in this chapter are successful, we will be able to maintain the claim that movement obligatorily leaves a copy, and pursue the copy theory approach to the effects of Scope Narrowing.

This chapter is organized as follows. After reviewing some arguments in favor of the copy theory of movement, the next section discusses the fact that A-movement always bleeds Condition C as a challenge to this view of movement. The first three
subsections in section 3.3 are devoted to introducing three essential components of the proposed analysis, each of which are independently motivated. They include Lebeaux’s (1988) Late Merge, the Late Merge approaches to Extraposition proposed in Fox and Nissenbaum (1999) and further studied in Bhatt and Pancheva (2004, 2005), and Fox’s (1999a, 2002) Trace Conversion. In 3.3.5, I argue that these independently motivated components of grammar are consistent with determiners merging with their restrictor constituent counter-cyclically (i.e., Wholesale Late Merge), and that Wholesale Late Merge is a possible solution to the Condition C asymmetry between A-movement and A’-movement. Further evidence from ellipsis favoring Wholesale Late Merge will be provided in section 3.4. Section 3.5 examines some puzzling ACD facts discussed in Wold (1995). Those ACD facts seem to suggest that QR sometimes leaves a trace, which is reminiscent of the above-mentioned property of A-movement. Taking this similarity as motivation, I will extend the Wholesale Late Merge approach to Wold’s ACD puzzle. Although the Wholesale Late Merge approach does not address the puzzle completely, I suggest that it is the first step towards solving the puzzle. As will be clear below, I make crucial use of a certain property of Case assignment to constrain the application of Wholesale Late Merge. In section 3.6, I suggest that the Case and agreement system developed in Pesetsky and Torrego (2001, 2004, 2005, to appear) fits well with the proposals in this chapter.

3.2 The Scope Narrowing Puzzle

3.2.1 Scope Narrowing in A-movement and A’-movement

In order to set the stage for introducing the challenges to the copy theory of movement, this subsection will establish that the effects of Scope Narrowing are observed both in A-movement and A’-movement and review some arguments favoring the claim that they are straightforwardly explained within the copy theory of movement framework.
First, let us consider some cases that demonstrate the effects of Scope Narrowing in A-movement. As illustrated in (1), subject QPs in the raising construction can take narrow scope relative to the predicates that they c-command at the surface syntactic representation.\(^1\)

(1) a. Someone from New York seems to be very likely to win the lottery.
   (∃>seem>likely) (seem>∃>likely) (seem>likely>∃)

b. Many soldiers seem to be very likely to die in the battle.
   (many>seem>likely) (seem>many>likely) (seem>likely>many)

(Fox 1999a:160)

The presence of the effects of Scope Narrowing in A-movement receives further support from the fact that pronouns can be bound by QPs whose scope is superficially narrower than the subjects containing the pronouns, as in (2). Pronouns are interpreted as bound variables only when they are within scope of their binders at LF, as discussed in 1.2.2. Thus, the facts in (2) can be taken as another instance of Scope Narrowing.

(2) a. [Someone from his class] seems to [every professor] to be a genius.

b. [His father] seems to [every boy] to be a genius.

(Fox 1999a:161)

Notice that it is widely argued that both in (1) and in (2), the subject is originally generated in a position lower than the elements that can take wide scope over it in LF, and it undergoes overt movement to its surface position. For instance, it is argued that the movement dependency illustrated in (3) is involved in the derivation of (2)b.

---

\(^1\) The facts discussed in this section are mostly taken from Fox (1999a). The data that demonstrate the claims discussed here can be found elsewhere in the literature, as well. See Heycock (1995), Lebeaux (1998), May (1985), Romero (1998b), and Sportiche (2005), among many others, for relevant facts and discussion.
(3)  [[his, father] seems to [every boy], [ __ to be [ __ a genius]]]

As opposed to the cases in (2) where the relevant movement of the subject is involved, variable binding is not allowed in (4), in which it is argued that the relevant movement is absent, as illustrated in the syntactic representation in (5).²

(4)  a.  ??[Someone from his, class] shouted to [every professor], to be careful.
     b.  ??[His, father] wrote to [every boy], to be a genius.
     
     (Fox 1999a: 161)

(5)  [[[his, father] [ __ wrote to [every boy]],] [PRO to be a genius]]

The generalization that emerges from the facts above is stated in (6), and this generalization has been argued for by various sorts of evidence in the literature (Fox 1999a, Lebeaux 1998, May 1985, and Sauerland and Elbourne 2002, among many others).

(6)  The Correlation between Movement and Interpretation

A constituent β can be interpreted at a position γ, only if β partakes in a movement dependency that includes γ as a chain member.

The correlation in (6) can be seen in the A'-movement context, as well. In (7), pronouns inside the wh-phrases can be bound by the QPs that are structurally lower than the wh-phrases at the surface syntactic representation. Notice that in (7), the wh-phrases are base-generated in the position below the QPs, as shown in (8).

² In (5), QR of the QP over the pronoun does not help establish a variable binding dependency because it induces a violation of weak crossover. This is also the case in wh-movement cases in (9).
(7)  a.  [Which of his students] did [every professor], talk to?
b.  [Which of his students] did [no professor], talk to?
c.  [Which student of his] did you think [every professor], talked to?
d.  [Which student of his] did you think [no professor], talked to?

(Fox 1999a: 172)

(8)  [[which student of his] did you think [no professor], talked to ___ ]

Once again, variable binding is impossible in (9) where a wh-phrase containing a pronoun is not originally generated in a position lower than its binder. (See (5) for relevant syntactic representation.)

(9)  a.  *[Which of his students] talked to [every professor],?
b.  *[Which of his students] talked to [no professor],?
c.  *[Which student of his] did you think talked to [every professor],?
d.  *[Which student of his] did you think talked to [no professor],?

(Fox 1999a: 172)

The correlation in (6), which holds both in A-movement and in A’-movement, is particularly illuminating in exploring the theory of movement. It has often been taken as an empirical argument for the copy theory of movement. The correlation in (6) is straightforwardly captured by capitalizing on consequences of this view of movement, in which movement is considered to be a merging process of a single constituent in different syntactic positions (i.e., Remerge). In this view, the scope of a constituent is determined simply by choosing which copy of the constituent will bear quantificational force.3 For example, the LF representation in (10) serves as an input of narrowest scope of the subject QP in (1)a, and, as one can easily see, other scopal relations can be produced by

3 A QP α takes scope at a position β when a constituent that is sister to β is interpreted as an argument of α in the semantic component.
the same mechanism. 4 (The doubly strikeout material is intended to represent constituents invisible to the semantic component.)

(10) [(someone from New York) seems [(someone from New York) to be very likely [someone from New York] to win the lottery]]

Given the correlation between movement and interpretation, it seems reasonable to draw the conclusion that A-movement and A’-movement share the same property in this respect: they can leave a copy of a moved constituent. However, this conclusion needs to be reassessed once the facts introduced in the next subsection are taken into consideration.

3.2.2 Condition C Bleeding in A-movement

While various arguments have been proposed that support the copy theory approach to the effects of Scope Narrowing (including the correlation in (6)), it is also true that some

---

4 The presence of the effects of Scope Narrowing in A’-movement is uncontroversial. However, it has sometimes been suggested that it is less obvious in A-movement. Most notably, Lasnik (1999a) provides facts such as (i), building on an observation from Chomsky (1995), and argues that Scope Narrowing is never possible in A-movement.

(i) a. Every Mersenne number was proved not to be prime.
   (∀>not) *(not>∀)
   b. No large Mersenne number was proven to be prime.
   (no>prove) *(prove>no)
   c. No one is certain to solve the problem.
   (no>certain) *(certain>no)

(Lasnik 1999a:199, 205)

I cannot offer an account of the absence of the effects of Scope Narrowing in cases such as (i). However, I suggest that a key to this issue lies in exploring constraints on the process sometimes called total reconstruction. The effects of total reconstruction can be found in A-movement where entire constituents can be interpreted in a position lower than their surface position, as exemplified in (10). In this respect, wh-movement is different from A-movement. In wh-movement, a wh-operator takes its scope in its derived position (partial reconstruction). Chapter 4 will discuss several constraints on total reconstruction. See also Aoun (1982), Boeckx (2001), von Fintel and Iatridou (2003), Nevins and Anand (2003), and Wurmbrand and Bobaljik (1999) for relevant discussion and observation.
challenges to this view have often been pointed out in the past literature. This subsection takes up one such challenge.

To set the stage for discussing this challenge, let us first consider the fact that A'-movement does not bleed Condition C (Freidin 1986, Lebeaux 1988, and van Riemsdijk and Williams 1981).5 This fact is exemplified by wh-movement in (11), in which the pronoun cannot be co-referential with the R-expression John in the fronted wh-phrase.

(11) a. *\[Which pictures of John\] does he, like?
   (Lebeaux 1988:144, 146)
   b. *[Which corner of Johni's room] did he, sitting in?
      (David Pesetsky personal communication)

This fact naturally follows from the conclusion reached in the last subsection: A'-movement leaves a copy of a moved constituent. Under the copy theory of movement, there is a copy of the moved wh-phrase in the c-command domain of the pronoun, which violates Condition C, as shown in (12).

(12) a. *[[which pictures of Johni] does [he, like [which pictures of Johni]]]
   b. *[[which corner of Johni's room] did [he, sitting in [which corner of Johni's room]]]

As illustrated in (13), covert A'-movement such as QR does not bleed Condition C, either. This is also expected under the copy theory of movement, as illustrated in (14).

(13) *A different person told him, about every argument that Johni is a genius.
    (Fox 1999a:192)

(14) *[[every argument that Johni is a genius] [a different person told him, about
        [every argument that Johni is a genius]]]

5 One qualification is in order here. This is true only when relevant R-expressions are within argument constituents of wh-phrases (see references just cited and discussion below).
Together with the correlation between movement and interpretation in (6), the fact that A’-movement does not bleed Condition C in the above cases has been taken as further empirical evidence for the copy theory of movement (Chomsky 1993, 1995, Fox 1999a, 2002, Fox and Nissenbaum 1999, and Sauerland 1998, 2004, among many others). At this stage, it seems reasonable to draw a conclusion stronger than the one in the last subsection: A’-movement obligatorily leaves a copy.

Let me now introduce a challenge to the copy theory of movement. A puzzling fact is that, as opposed to A’-movement, A-movement does bleed Condition C. This is illustrated in (15).

(15) a. The claim that Johni was asleep seems to himi to be correct.  
(Chomsky 1993:37)
b. Every argument that Johni is a genius seems to himi to be flawless.  
(Fox 1999a:192)
c. Johni’s mother seems to himi to be wonderful.
d. Pictures of Johni seem to himi to be great.  
(Lebeaux 1998:23-24)

Within the copy theory of movement framework, we would postulate the LF representation in (16) for the sentence in (15)a. However, it is clear that this is not the right result since, in this LF representation, Condition C is violated, contrary to fact:

(16) [[the claim that Johni was asleep] seems to himi [[the claim that Johni was asleep] to be correct]]

---

6 Jacobson (2004b), Kuno (2004), and Lasnik (1999a) cast doubt on the facts in (11). A relevant argument can, however, be constructed by employing covert A’-movement, as in (13). The judgment on covert movement cases is clearer than overt movement cases. See also Safir (1999) for related facts.
Rather, the LF representation that we want to derive is something close to the following one where the (relevant) tail of the A-movement chain is represented by a contentless trace:

(17) \[\text{[[the claim that John was asleep]] seems to him, [t₁ to be correct]]\]

In the previous subsection, we have already seen evidence that suggests that A-movement can also leave a copy. Together with this evidence, the Condition C bleeding in A-movement suggests that A-movement *optionally* leaves a trace. The puzzling asymmetry between A-movement and A*-movement is summarized as follows (Fox 1999a):

(18)  
   a. A-movement *optionally* leaves a trace.
   b. A*-movement *obligatorily* leaves a copy.

Building on the independently motivated ideas about movement and the merger operation, I will instead argue that there is a way to yield an LF representation that is identical to (17) from the perspective of Condition C and yet is consistent with the copy theory of movement. In addition, I will demonstrate that this option is available in A-movement cases such as (15), but not in A*-movement cases such as (11) and (13).

3.3 The Proposal

3.3.1 An Outline

In this section, I will propose a solution to the asymmetry between the two types of movement discussed in the previous section. The proposed account makes crucial use of three components of grammar, each of which is independently motivated. After introducing these components in the first three subsections, I will argue for a theory of
movement which captures the asymmetry in (18) within the copy theory of movement framework.

3.3.2 The First Component: Late Merge

I have discussed evidence suggesting that A’-movement obligatorily leaves a copy, and much evidence in favor of this claim has been accumulated in the literature. However, it has also been pointed out that there is one phenomenon which could be regarded as an exception to this claim. As we have already seen, the sentences in (19)a and (20)a suggest that wh-movement does not bleed Condition C. Notice that in these cases, the relevant R-expression is included in an argument constituent of a moved wh-phrase. From the perspective of the copy theory of movement, one puzzling fact is that if the relevant R-expression is dominated by adjuncts that modify a wh-phrase (e.g., relative clauses), wh-movement appears to bleed Condition C, as has been pointed out by Freidin (1986), Lebeaux (1988), and van Riemsdijk and Williams (1981), among others. This argument/adjunct asymmetry in Condition C bleeding is illustrated in (19) and (20).7

7 We might also expect the same asymmetry to be found in covert A’-movement, or QR. However, the contrast is less clear in QR. It has been suggested that there is variation on the judgment of adjunct cases such as (i)b (Fox 1995b, 2002 and Kennedy 1997).

(i) a. *A different person told him, about every argument that John is a genius.
   b. ??A different person told him, about every argument that John made.

However, Fiengo and May (1994) and Fox (1995b, 2000, 2002) observe that QR does bleed Condition C in ACD:

(ii) a. You sent him, the letter that John expected you would <send him>.
b. You introduced him, to everyone John, wanted you to <introduce him, to>.
c. I reported him, to every cop John was afraid I would <report him, to>.
   (Fox 2002:84)

See Fox (1995b, 2000) for one possible analysis that explains (i)b and (ii). I will discuss Fox’s (2002) account of the facts in (ii) in 3.4.2.
Lebeaux (1988) offers a way to account for the argument/adjunct asymmetry in Condition C bleeding in A'-movement in a way that is compatible with the copy theory of movement. (Note that while Lebeaux made his proposal prior to the copy theory of movement, Lebeaux’s idea can be stated within this framework. See Chomsky 1993, 1995 for this point.) Lebeaux’s proposal constitutes the first step towards our proposed analysis capturing the movement properties in (18).

The facts in (19) and (20) seem to indicate that when the relevant R-expression is within an adjunct of a moved wh-phrase, the R-expression is not c-commanded by the co-referential pronoun at any stage of the derivation. Lebeaux (1988) accounts for this by making essential use of the complementation property of lexical items and a novel theory of the merger operation. Following Chomsky’s (1981) idea that the subcategorization property of lexical items must be satisfied throughout the derivation (which is known as the Projection Principle), Lebeaux argues that argument constituents of a lexical item such as the sentential complement clause in (19)a and (20)a must merge with the lexical item that selects them prior to the application of movement, in order to meet its subcategorization requirement from the beginning of the derivation. Therefore, a violation of Condition C is inevitable in (19)a and (20)a, as shown in the LF representations below:

\[(21) \quad a. \quad *[[\text{which report that John, was incompetent\]} \text{ did } \he, \text{ submit } [\text{which report that John, was incompetent}]]\]

\[(21) \quad b. \quad *[[\text{which argument that John, is a genius\]} \text{ did } \he, \text{ believe } [\text{which argument that John, is a genius}]]\]

75
On the other hand, adjuncts are not constituents that are required by the selectional property of a lexical item. Capitalizing on this feature of adjuncts, Lebeaux claims that they can be introduced at any stage of the derivation. More specifically, adjuncts can be attached onto a wh-phrase that they modify after the application of movement (Late Merge). Consequently, Condition C is not violated in (19)b and (20)b:

\[
\begin{align*}
(22) & \quad \text{a. } \text{[[which report [that John, revised]] did [he, submit [which report]]]} \\
& \quad \text{b. } \text{[[which argument [that John, made]] did [he, believe [which argument]]]}
\end{align*}
\]

From these facts, Lebeaux draws the generalization that Condition C is obeyed only when a relevant R-expression is included in adjuncts that modify a moved constituent (Lebeaux’s Generalization).

Lebeaux’s Late Merge analysis is insightful in the sense that the phenomenon, which could be taken as an exception to the claim that A’-movement obligatorily leaves a copy, can be dealt with in a way that is completely compatible with the copy theory of movement, as Chomsky (1993, 1995) discusses. However, Lebeaux’s version of Late Merge is problematic for A-movement. As mentioned, A-movement always bleeds Condition C:

\[
\text{In (22), the two copies of the wh-phrase are represented in the structure, and only the higher one dominates the Late Merged adjunct. In the view of the copy theory of movement where movement creates multiple duplicates of a constituent, it is clear that Late Merged material is interpreted only in a position where it is merged. However, it is less clear how Late Merge is implemented within a Remerge view of movement where there is only one occurrence of a constituent, which can be interpreted in more than one position. I suggest that Late Merge could be viewed as a post-interpretation operation within the Remerge framework. Let us assume that syntactic structures get interpreted cyclically and, once a structure is interpreted, it never gets interpreted again. In (19)b and (20)b, a constituent that dominates a wh-phrase (e.g., vP) is first interpreted and the wh-phrase Remerges with a CP constituent. An adjunct merges with the wh-phrase and the constituent that has not been interpreted gets interpreted. Given these assumptions, even if there is only one occurrence of the wh-phrase in a structure and the adjunct is adjoined to this wh-phrase, the adjunct is not interpreted together with the constituent that dominates the coreferential pronoun. To achieve this result, it is crucial to assume that once a structure gets interpreted, it is never interpreted again. Otherwise, Late Merge would not help us avoid a Condition C violation in (22). See also Fitzpatrick and Groat (2005) for proposal of how the effects of Late Merge are derived within the Remerge framework.}
\]
(23)  a. Every argument that John is a genius seems to him to be flawless.
     b. Every argument that John made seems to him to be flawless.

Given Lebeaux’s idea discussed above, it is clear that Late Merge of a sentential complement is not an option in (23)a (i.e., it would violate the Projection Principle). Therefore, Lebeaux’s theory of Late Merge needs to be revised if it is to help us solve the problem that A-movement always bleeds Condition C. I suggest that the key lies in exploring a different approach to the driving force for merging argument constituents with their heads prior to the application of movement. This issue will be discussed in the following two subsections.

3.3.3 The Second Component: A Late Merge Analysis of Extraposition

In order to investigate the reason why Late Merge is not applicable to argument constituents in the above cases, it is instructive to examine Extraposition, because it is claimed that argument constituents can sometimes be Late Merged in this construction (Bhatt and Pancheva 2004, 2005). To discuss this issue, let us first consider adjunct Extraposition, which also displays the Condition C bleeding effect, as shown in (24) and (25).

(24)  a. ??/*I gave him a picture [from John’s collection] yesterday.
     b. I gave him, a picture yesterday [from John’s collection].
     (Fox and Nissenbaum 1999:139)

(25)  a. ??/*I gave him an argument [that supports John’s theory] yesterday.
     b. I gave him, an argument yesterday [that supports John’s theory].
     (Fox and Nissenbaum 1999:139)
The Condition C bleeding in adjunct Extraposition can be taken as another instance of Lebeaux’s Generalization. For this and other reasons, Fox and Nissenbaum (1999) develop an analysis of adjunct Extraposition that makes crucial use of Late Merge of adjuncts. Fox and Nissenbaum claim that a DP first undergoes rightward QR, moving across an adverbial and that this movement is followed by Late Merge of an adjunct with the QRed DP. The fact that the DP and the adjunct are superficially discontinuous in Extraposition is captured by assuming that the DP can be pronounced in the tail position of the QR chain (Bobaljik 1995, 2002, Groat and O’Neil 1996, and Pesetsky 1998). The relevant stages of the derivation are illustrated in (26).

(26)

The Condition C bleeding fact in adjunct Extraposition can be taken as further evidence in favor of Lebeaux’s Generalization, which in turn supports the idea that Late Merge is applicable only to adjuncts.9

---

9 Fox and Nissenbaum (1999) argue that complement Extraposition does not bleed Condition C, as in (i).

(i) a. ??/*I gave him, a picture yesterday [of John,‘s mother].
   b. ??/*I gave him, an argument yesterday [that this sentence supports John,‘s theory].
   (Fox and Nissenbaum 1999:139)

If complement Extraposition were derived in the same way as adjunct Extraposition, it would also bleed Condition C, contrary to fact. Based on this and other reasons, Fox and Nissenbaum suggest that complement Extraposition is produced by movement of a complement constituent. This analysis of complement Extraposition is consistent with Lebeaux’s proposal of Late Merge.
However, Bhatt and Pancheva (2004, 2005) provide empirical evidence that suggests that Lebeaux's Generalization needs to be reconsidered. It comes from Extraposition of a so-called comparative complement introduced by *than*, as in (27)b.

(27)  
\begin{itemize}
  \item a. I bought more books [than you read] yesterday.
  \item b. I bought more books yesterday [than you read].
\end{itemize}

Bhatt and Pancheva propose that Extraposition in this case also involves Late Merge of the superficially extraposed constituent. One of their arguments in favor of the Late Merge approach comes from the fact that Extraposition bleeds Condition C in this context, as well:

(28) I will tell him a sillier rumor (about Ann) tomorrow [than Mary told John].

(Bhatt and Pancheva 2004:19)

Extending Fox and Nissenbaum’s analysis of adjunct Extraposition, Bhatt and Pancheva claim that in Extraposition of the comparative complement, QR first applies to the comparative operator -*er*, and the comparative complement Late Merges with the moved comparative operator, as illustrated in (29) (but see Grosu and Horvath 2005 for interesting criticism). Here, the tail of the QR chain is overtly realized, just like in adjunct Extraposition:

(29) 
\[
\begin{array}{c}
\text{TP} \\
\text{I}_1 \quad T' \\
\text{will} \quad \text{vP} \\
\text{vP} \quad \text{vP} \\
\text{t}_1 \text{tell him a sillier [-er] rumor tomorrow} \quad [-\text{er} \text{than Mary told John}] \\
\text{QR} \quad \text{Late Merge}
\end{array}
\]
Bhatt and Pancheva’s analysis of Extraposition leads us to reassess Lebeaux’s Generalization because, as discussed in Bhatt and Pancheva, we have good reason to regard the comparative complement as an argument constituent of the comparative operator (Bresnan 1973, 1975, and Carlson 1977, among others). As a consequence, their approach implies that in addition to adjuncts, Late Merge is sometimes applicable to argument constituents.

Now, the goal is to understand what aspect of grammar distinguishes Bhatt and Pancheva’s case, in which Late Merge of arguments is possible, from Lebeaux’s case, in which it is not. Fox (2002) proposes a constraint on Late Merge, which is further developed by Bhatt and Pancheva (2004, 2005). The constraint is based on a proposal used for the interpretation of chains within the copy theory of movement framework, to which we now turn.10

3.3.4 The Third Component: Trace Conversion

A general issue that arises in the copy theory of movement is how movement chains are embodied in LF representations so that the semantic component can interpret them compositionally. This issue can be appreciated by looking at a movement dependency of an object QP. As I have already argued in Chapter 2, the effects of Scope Widening are accomplished by movement, and, in this approach, object wide scope would be derived from syntactic representations such as (30)b.

(30)  a. Some fifth year student read every book.
      b. [[[every book], [some fifth year student read [every book]]]]

---

10 Fox (2002) argues that the possibility of Late Merge in wh-movement and in adjunct Extraposition is captured by a condition that requires LF representations to be interpretable without resort to the Projection Principle. Together with an ancillary assumption which I will discuss in the next subsection, approaches of this sort imply that Lebeaux’s Generalization should be reassessed. For expository purposes, this subsection has discussed Extraposition of the comparative complement as empirical motivation to reassess Lebeaux’s Generalization.
As discussed in Chapter 1, the LF representation in (30)b is not interpretable for several reasons. For one, the lower copy of the object QP in (30)b cannot combine with the transitive predicate read, which demands an element of type e as its first argument. Fox (1999a, 2002) proposes that copies left behind by movement are converted into definite descriptions of type e, which solves this type mismatch (see also Sauerland 1998, 2001, 2004). As Fox discusses, the fact that wide scope of the object QP can expressed by the following paraphrase is also suggestive:

(31) For every book x, there is some fifth year student who read the book identical to x.

Fox’s mechanism for producing this result is stated in (32).

(32) Trace Conversion
   a. Variable Insertion: (Det) Pred → (Det) [Pred λy(y=x)]
   b. Determiner Replacement: (Det) [Pred λy(y=x)] → the [Pred λy(y=x)]

(Fox 2002:67)

Let me illustrate here how Trace Conversion converts the uninterpretable representation in (30)b into an interpretable representation. First, movement of the object QP introduces a λ-operator right below the moved element (see Chapter 1 for relevant discussion). The restrictor NP book combines with the predicate λy [y=x] introduced by Variable Insertion.\(^{11}\) This step establishes a variable binding dependency between the λ-operator

\(^{11}\) This is accomplished by Predicate Modification in (i).

(i) Predicate Modification
   If α is a branching node and \{β, γ\} the set of its daughters, then, for any assignment g, α is in the domain of \[\parallel x\parallel^α\] if both β and γ are, and \[\parallel β\parallel^α\] and \[\parallel γ\parallel^α\] are both of type \langle e, t\rangle. In this case, \[\parallel α\parallel^α = λx: x ∈ D\] and x is in the domain of \[\parallel β\parallel^α\] and \[\parallel γ\parallel^α\].

\[\parallel β\parallel^α (x) = \parallel γ\parallel^α (x) = 1.\]

(Heim and Kratzer 1998:126)

Following Fox (2002), I adopt the expression identical to x in order to illustrate the result of an application of Predicate Modification.

81
and the lower copy. Finally, the determiner *every* is replaced with the definite determiner in the lower copy.\(^\text{12}\) The resulting representation in (33) can be compositionally interpreted by the semantic rules discussed in Chapter 1, producing the appropriate interpretation.\(^\text{13}\)

\[
(33) \quad \text{[some fifth year student read every book]}
\]

\[
\rightarrow \text{movement & } \lambda\text{-abstraction}
\]

\[
[[\text{every book}] \lambda x. \text{[some fifth year student read [every book]]}]
\]

\[
\rightarrow \text{Variable Insertion}
\]

\[
[[\text{every book}] \lambda x. \text{[some fifth year student read [every book identical to } x]]]
\]

\[
\rightarrow \text{Determiner Replacement}
\]

\[
[[\text{every book}] \lambda x. \text{[some fifth year student read [the book identical to } x]]]
\]

We are now in a position to return to the question of when Late Merge is possible. A possible answer first suggested in Fox (2002) and further studied in Bhatt and Pancheva (2004, 2005) is that argument constituents can be Late Merged as long as the resulting structure is interpretable. Let us first see whether Extraposition of a

\(^{12}\) Following Chomsky (1993, 1995), I have argued that as opposed to the traditional trace theory, the copy theory of movement simplifies the theory of movement. However, as Fox (2003) argues, if Trace Conversion is viewed as a syntactic rule, the force of this argument becomes less strong because it is a rule that needs to be postulated only in the copy theory (see also Adger and Ramchand 2005 for relevant discussion). Fox (2003) suggests one possible approach to this issue. Both the copy theory and the trace theory need to employ a semantic rule that assigns a meaning to a constituent sister of a moved constituent. One such semantic rule is Predicate Abstraction, introduced in Chapter 1. Capitalizing on this property of grammar, Fox formulates a semantic rule of this sort in a particular way so that its application produces the same interpretation as Trace Conversion does, when taken together with the version of Predicate Abstraction, introduced in Chapter 1. Given an approach along these lines, Trace Conversion does not have to be regarded as an independent syntactic rule. For expository purposes, I continue to assume that Trace Conversion is a rule in syntax.

\(^{13}\) As Fox (2002) points out, Trace Conversion produces a correct result as long as determiners are conservative, as defined in (i).

\[
\text{(i) A determiner } D \text{ is conservative iff for any } P, Q, D(P)(Q) = D(P)(P \cap Q).
\]

Since all determiners in natural language are considered to be conservative, Trace Conversion does not make any incorrect prediction (but, see Bhatt and Pancheva 2004, 2005 for an interesting consequence in this respect).
comparative complement results in an interpretable structure. Bhatt and Pancheva (2004, 2005) crucially exploit the fact that the comparative operator is semantically identical to determiners in relevant aspects (see Chapter 4 for detailed semantics of the comparative operator). This property of the comparative operator is important for converting its movement chain into an interpretable object by Trace Conversion. Because the comparative operator is a determiner, its copy is replaced with the definite determiner by Determiner Replacement. In addition, Variable Insertion supplies a predicate of type \( \langle d, t \rangle \) to the comparative operator, which functions as its first argument.\(^{14,15}\) These steps are illustrated in (34)b, and make the movement chain of the comparative operator compositionally interpretable, as Bhatt and Pancheva discuss.

(34) a. John was taller last year than Bill was <tall>.
   b. [John was tall[-er] last year]

\[\rightarrow \text{movement, } \lambda\text{-abstraction, } \& \text{ Late Merge}\]
\[\rightarrow \text{Variable Insertion}\]
\[\rightarrow \text{Determiner Replacement}\]

What Fox suggests, and Bhatt and Pancheva demonstrate by investigating Extraposition in comparatives, is that Late Merge of an argument of a determiner in general does not cause an interpretability problem, due to Trace Conversion.

\[^{14}\text{I assume that the comparative operator denotes an ordering relation between sets of degrees. For brevity, I do not make a distinction between degrees and individuals here. See Chapter 4 for details.}\]
\[^{15}\text{For Bhatt and Pancheva's purpose (and ours, as well, which I will discuss shortly), Trace Conversion needs to be revised to accommodate the situation in which restrictor constituents of determiners are absent:}\]

(i) **Trace Conversion**
   a. Variable Insertion: \((\text{Det}) \rightarrow (\text{Det}) [(\text{Pred}) \lambda y(y=x)]\)
   b. Determiner Replacement: \((\text{Det}) [(\text{Pred}) \lambda y(y=x)] \rightarrow \text{the } [(\text{Pred}) \lambda y(y=x)]\)
Let us now check whether Trace Conversion is also able to derive an interpretable LF representation from a structure that is generated by Late Merge of an argument in Lebeaux's wh-movement case. Fox (2002) argues that it is not. Notice crucially that the Late Merged constituents in this case are arguments of NPs, unlike in the comparative case above. For the sake of our discussion, let us assume that NPs that take a sentential complement (e.g., argument in the expression the argument that John is a genius) is of type \((t, \langle e, t \rangle)\). Suppose that this sentential complement Late Merges with a DP after its movement, which produces the syntactic representation in (35). In (35), the lower copy of the NP argument is of type \((t, \langle e, t \rangle)\) because its first argument is not saturated. For this reason, it cannot combine either with which or with the definite determiner, which is the replacement of which by Trace Conversion. This situation is strikingly different from Extraposition of the comparative complement. In (35), neither Determiner Replacement nor Variable Insertion can fix the problem of interpretability. More generally, the structure that involves Late Merge of an argument constituent of an NP ends up being a semantically illegitimate representation (Bhatt and Pancheva 2004, 2005 and Fox 2002).

(35) \([\text{CP } [\text{which argument [that John is a genius]] } [\text{TP he believe [which argument]]}]\]

For this interpretability reason, Late Merge can be applied to an argument constituent of the comparative operator, but not to an argument constituent of an NP, which correctly captures the Condition C bleeding facts in the various cases discussed above.

### 3.3.5 Wholesale Late Merge

I will demonstrate here that together with one constraint on application of the counter-cyclic merger that I introduce shortly, the three components discussed above conspire to derive the puzzling asymmetry between A-movement and A'-movement repeated here:

(36) a. A-movement \textit{optionally} leaves a trace.

b. A'-movement \textit{obligatorily} leaves a copy.
In order to explain the properties of the two types of movements in (36), I capitalizing on an implication of Bhatt and Pancheva’s (2004, 2005) and Fox’s (2002) view of Late Merge: Late Merge is applicable as long as a resulting structure is interpretable. Specifically, I claim that all types of determiners (quantificational determiners as well as the comparative operator) can undergo movement alone, and their restrictor constituents can be attached onto the moved determiners counter-cyclically (Wholesale Late Merge), as has already been hinted at in the previous subsection.\(^{16,17}\) Let me first illustrate how this Wholesale Late Merge approach helps us explain the fact that A-movement always bleeds Condition C, as repeatedly illustrated in (37).

\[(37)\] a. The claim that John\(_i\) was asleep seems to him\(_i\) to be correct.  
   b. Every argument that John\(_i\) is a genius seems to him\(_i\) to be flawless.  
   c. John\(_i\)’s mother seems to him\(_i\) to be wonderful.  
   d. Pictures of John\(_i\) seem to him\(_i\) to be great.

I argue that the restrictor NPs containing the R-expression Wholesale Late Merge with the determiners in (37) after they move out of the c-command domain of the co-referential pronoun, as shown in (38).\(^{18}\) (I will specify shortly the exact syntactic position where the restrictor NP is introduced in (38).) The interplay of Determiner Replacement and Variable Insertion converts all copies of every, but the one in the head of the chain, into syntactic objects that receive the same interpretation as is assigned to traces (contentless copies). As is clear from the final outcome in (38), Condition C is not violated because there is no copy of the R-expression c-commanded by the co-referential pronoun.

\(^{16}\) Late Merge of this sort has already been suggested in Fox (2002: fn. 11) where it is attributed to Uli Sauerland. However, a potential counter-argument has also been pointed out there, which I will discuss below.

\(^{17}\) In a series of papers, Sportiche (1997, 1999, 2005) develops a theory where DPs are constructed in the reverse order to the Wholesale Late Merge approach. Sportiche claims that only NPs are introduced in their argument position and determiners are base-generated in a position higher than that.

\(^{18}\) Pesetsky and Torrego (2005) suggest an approach to the Condition C bleeding fact in A-movement, which is analogous to the analysis proposed here.
(38) [[every] flawless]
  \rightarrow \text{successive cyclic movement of a determiner}
  [[every] seems to him, [[every] to be [[every] flawless]]]
  \rightarrow \text{Wholesale Late Merge}
  [[every [argument that Johni is a genius]] seems to himi, [[every] to be [[every] flawless]]]
  \rightarrow \text{Trace Conversion}
  [[every [argument that Johni is a genius]] \lambda x. \text{seems to him, } [[x] \lambda y. \text{to be}}
  \rightarrow [[y] \text{flawless]]]

So far, the principle of grammar relevant for the applicability of the counter-cyclic merger operation is the interpretability condition on LF representations. I will now introduce another property of grammar that plays a role in the application of this operation.

The proposed approach offers a way to derive the property of A-movement that it optionally leaves a trace. However, this approach does not capture the property of A'-movement that it obligatorily leaves a copy because nothing so far prohibits the same processes from taking place in A'-movement. As a consequence, it would predict that Condition C is also always obviated by A'-movement, contrary to fact:

(39) [believe [which]]
  \rightarrow \text{movement of a determiner, Wholesale Late Merge, & Trace Conversion}
  [[which [argument that Johni is a genius]] \lambda x. \text{he, believe } [x]]]

I take this problem as an indication that Late Merge is further constrained. Here, I suggest a constraint that makes specific reference to the Late Merged NPs. It has been argued that DPs receive Case by entering an agreement relation with their Case assigning heads, and that this agreement relation is established only when Case assigning heads c-command DPs. Capitalizing on this claim, I argue that the NPs in the complement of determiners must be introduced within the c-command domain of a head that provides Case for them (see section 3.6 for further discussion on Case). This Case requirement on
NPs rules out the derivation in (39) because the NP is merged higher than the Case assigning head and forces the whole DP to merge with the predicate, as illustrated in (40). Here, the complement of the predicate is the only position where the object DP can be introduced into the derivation because it needs to receive Case from the functional head immediately above the predicate (i.e., $v$ in Chomsky’s 1995 framework).

$$
\text{(40)} \quad [_{vP} v \text{ believe [which argument that Johni is a genius]}]
$$

$$
\rightarrow \text{wh-movement}
$$

$$
[_{CP} \text{ [which argument that Johni is a genius]} \ [_{TP} \text{ hei believe [which argument that Johni is a genius]}]]
$$

The proposed solution now correctly predicts that Condition C cannot be obviated by A'-movement, and, more generally, derives the property of A'-movement in (36)b.

Let us now go back to A-movement and make sure that Wholesale Late Merge in the derivation in (38) obeys the Case requirement on NPs. I argue that the restrictor NP merges with the moved determiner at the matrix VP-adjoined position in this case, as shown in (41).\footnote{The availability of the matrix VP-adjoined position as an intermediate landing site in A-movement receives support from the scope fact observed in Sauerland (2003a). In (i), the universal QPs can take narrow scope relative to negation. At the same time, the pronouns in the experiencer position of the matrix predicate \textit{seem} can be bound by the universal QPs. This suggests that there is a copy of the subject QPs between the two elements, which I assume is the matrix VP-adjoined position, following Sauerland (2003a):}

\begin{enumerate}
\item[(i)]
\begin{enumerate}
\item \[\text{[Every childi doesn’t seem to hisi father to be smart.}\]
\quad (\forall > \text{not}) \ (\text{not} > \forall)
\item \[\text{[Every participanti didn’t seem to hisi coach to be in bad shape.}\]
\quad (\forall > \text{not}) \ (\text{not} > \forall)
\item \[\text{[All linguistsi didn’t seem to theiri employer to work hard.}\]
\quad (\forall > \text{not}) \ (\text{not} > \forall)
\end{enumerate}
\end{enumerate}

\begin{flushright}
(Sauerland 2003a:310-311)
\end{flushright}

See also Legate (2003) for relevant observation and discussion.
than the Case assigner of the NP, namely, a tense head, T. As a consequence, Wholesale Late Merge meets all the relevant requirements in this case:

\[
(41) \quad [\text{VP} \ [\text{every}] \ [\text{VP} \ \text{seems to him} \ [\text{TP} \ \text{every}] \ \text{to be } \ [\text{every} \ \text{flawless}]])
\]

\[\rightarrow\ \text{Wholesale Late Merge}\]

\[
[\text{VP} \ [\text{every}] \ [\text{argument that Johni is a genius}] \ [\text{VP} \ \text{seems to him} \ [\text{every}] \ \text{to be } \ [\text{every} \ \text{flawless}]])
\]

\[\rightarrow\ \text{movement of the DP to the Spec of TP}\]

\[
[\text{TP} \ [\text{every}] \ [\text{argument that Johni is a genius}] \ T \ [\text{VP} \ [\text{every}] \ [\text{argument that Johni is a genius}] \ [\text{VP} \ \text{seems to him} \ [\text{every}] \ \text{to be } \ [\text{every} \ \text{flawless}]])]
\]

Together with the Case constraint on NPs, the Wholesale Late Merge approach now derives both of the properties in (36) within the copy theory of movement framework.

As discussed above, Trace Conversion converts every copy of a determiner that does not appear with its restrictor NP into a syntactic object that receives an interpretation identical to the one that is assigned to traces. Thus, the copy theory approach to the effects of Scope Narrowing predicts that QPs cannot take their scope in positions occupied by copies of this sort. Given this, we expect that if a subject contains an R-expression, it cannot take scope in a position lower than a co-referential pronoun. As Fox (1999a) and Romero (1998b) shows, this is true.

---

20 Bhatt and Pancheva (2004:fn. 33) have already pointed out the possibility that a derivation like (39) would not be allowed by Case considerations. However, Bhatt and Pancheva do not explore its consequences (e.g., that Wholesale Late Merge would be allowed in A-movement, as in (41)).

21 Fox (2000) argues that the fact that Scope Narrowing feeds Condition C is demonstrated more clearly by employing creation verbs, which force Scope Narrowing of their argument QPs, due to their meaning:

(i) a. For these issues to be clarified, many more/new papers about his philosophy seem to Quine, to be needed.

b. #For these issues to be clarified, many more/new papers about Quine's philosophy seem to him, to be needed.

(Fox 2000:170)

Sportiche (2003, 2005) discusses facts that make a similar point: Scope Narrowing forced by idiom chunks feeds Condition C.

88
(42)  a.  [A student of Davidi’s] seems to him to be at the party.
   (∃>seem) *(seem>∃)
  
   b.  [A student of his] seems to Davidi to be at the party.
   (∃>seem) (seem>∃)

(43)  a.  [Someone from Davidi’s city] seems to him to be likely to win the lottery.
   (∃>seem) *(seem>∃)
  
   b.  [Someone from his city] seems to Davidi to be likely to win the lottery.
   (∃>seem) (seem>∃)

(Fox 1999a:197)

In (42)a and (43)a, Condition C conflicts with the requirement for deriving narrow scope of the subject QP that the restrictor NP must be introduced below seem.

To sum up the proposals in this subsection, argument constituents can be introduced counter-cyclically as long as the resulting structure is licensed by principles of grammar. I have discussed the two relevant principles: the interpretability condition on LF representations (Bhatt and Pancheva 2004, 2005 and Fox 2002) and the Case constraint on NPs. I have demonstrated that the proposed Wholesale Late Merge approach captures the above-mentioned properties of A-movement and A’-movement within the simplest version of the copy theory of movement framework. If this proposal is correct, the A/A’-distinction is not relevant for whether movement can leave a contentless copy. Instead, we can draw the following generalization:

(44)  The Contentless Copy Generalization

In a movement chain of a DP <α₁ ... αₙ>, where α₁ is the tail and αₙ is the highest position within the c-command domain of its Case assigning head, the chain members <α₁ ... αₙ₋₁> can be interpreted as syntactic objects that receive the same interpretation as that assigned to traces.

---

22 The Contentless Copy Generalization holds only for DPs in which a restrictor constituent of a determiner is an NP. I will discuss a special case where the restrictor is a CP. This generalization does not hold for this special case for a reason to which I will return in section 3.5.
In the reminder of this chapter, I will present additional arguments that support this generalization.

3.4 Further Evidence from Ellipsis

3.4.1 An Outline

In this section, I will discuss additional arguments in favor of the Wholesale Late Merge approach. The main empirical concerns are facts about ellipsis discussed in Kennedy (1994, 2004) and Sauerland (1998, 2004), among others. Those facts suggest that if a DP moves out of an elided constituent and the corresponding DP moves out of the antecedent constituent, copies left behind by these movements have an effect on licensing ellipsis of elided material in certain contexts. Sauerland (1998, 2004) takes these facts to be as evidence in favor of the copy theory of movement, but also provides a challenging fact to this theory. I will argue that this challenging fact can be dealt with by the Wholesale Late Merge analysis. Since the following discussion pertains to the analysis of ACD within the copy theory of movement framework, I first discuss Fox’s (2002) Late Merge approach, which explains ACD within this theory.

3.4.2 A Late Merge Analysis of ACD

As discussed in Chapter 2, ellipsis of an elided constituent (EC) is licensed only when one can find a Parallelism Domain – a constituent that reflexively dominates EC and is semantically identical to an antecedent constituent (AC) modulo focus marked material. From this perspective of the parallelism condition, ACD, exemplified in (45)a, raises a question, especially because an antecedent VP contains an elided VP in this construction (Bouton 1970). Therefore, this construction seems to violate the parallelism condition, yet it is grammatical. It is claimed that this antecedent-containment relation is resolved by applying QR to a QP that dominates the elided VP. As a result of QR, an elided
constituent is outside of an antecedent VP (Fox 1995b, Kennedy 1997, Larson and May 1990, May 1985, and Sag 1976; cf., Baltin 1987). Within the copy theory of movement framework, this is the case only if properties of grammar allow us not to represent a relative clause in the original position of the QP, as illustrated in the two possible LF representations in (45)b and in (45)c, the former of which does not satisfy the parallelism condition. 23

(45) a. Polly visited every city Uli did <visit>.
   b. *[[every city [λx. Uli did <visit the city x>]]
      λy. [Polly visited [the city y [λx. Uli did <visit the city x>]]]]
   c. [[every city [λx. Uli did <visit the city x>]]
      λy. [Polly visited [the city y]]]

Pointing out this conflict between the copy theory of movement and the parallelism requirement for licensing ellipsis of EC in ACD, Fox (2002) claims that Lebeaux’s Late Merge operation is at work in deriving the representation in (45)c. Extending Fox and Nissenbaum’s (1999) analysis of adjunct Extraposition, Fox argues that the object QP first undergoes QR, and the relative clause Late Merges with the QRed object in its derived position, as shown in (46). (See Abe and Hoshi 1999, Baltin 1987, and Lasnik

23 Elaborating earlier proposals of the matching analyses of relative clauses (Chomsky 1965 and Lees 1961, 1963), Sauerland (1998) argues that a lexical item corresponding to a head noun of a relative clause undergoes movement relative clause internally, as in (i)b, and that the occurrence of the head noun in the Spec of the relative clause CP is phonologically deleted under semantic identity with the CP-external occurrence of the head noun, as illustrated in (i)c (see also Hulsey and Sauerland to appear).

(i) a. [TP Uli visited [OP city]]

   movement & Trace Conversion

   b. [CP [OP city] λx. [TP Uli visited [the city x]]]

   merger of DP and CP & deletion of the relative clause internal head noun

   c. [DP every city [CP [OP city] λx. [TP Uli visited [the city x]]]]

Note that the lexical content of a head noun is represented relative clause internally in this matching analysis. As I will discuss shortly, this matching analysis is one of the important components of Fox’s analysis, which yields a structure that satisfies the parallelism condition in ACD, as shown in (46). Furthermore, it plays a crucial role to capture Kennedy’s (1994, 2004) and Sauerland’s (1998, 2004) facts, which I will discuss in the next subsection.

91
1999b for other forms of Extrapolation approaches to ACD.) The two underlined λ-predicates are semantically identical, modulo the focus marked subject utorial in (46), and this licenses ellipsis of EC in ACD.

(46)  [Polly visited [every city]]

→ QR

[[λx. [Polly visited [the city x]]] [every city]]

→ Late Merge

[[λx. [Polly visited [the city x]]] [every city [λy. Uli did <visit the city y>]]]

Fox’s approach offers a straightforward solution to the above mentioned apparent conflict with the copy theory of movement by making use of the independently motivated operation Late Merge. It is supported by various facts of ACD, as well. This subsection discusses two of those facts, both of which can be taken as evidence in favor of the claim that Late Merge of relative clauses is an essential component for producing a structure that satisfies the parallelism condition in ACD (see the subsequent subsections for other arguments).

In a sentence with ACD like (45)a, Extrapolation of relative clauses is not reflected in the surface representation of sentences. However, Tiedeman (1995) argues that a visible form of Extrapolation of relative clauses is sometimes necessary to license ellipsis of EC in ACD, as illustrated in (47) and (48) (see also Baltin 1987, Larson and May 1990 and Wilder 1995, 2003 for relevant observation and discussion).

(47)  a. *John believed (that) everyone that you did <believe was a genius> was a genius.
b. John believed (that) everyone was a genius [that you did <believe was a genius>].

(Tiedeman 1995:75)
(48)  a. *I expect (that) everyone that you do <expect will visit Mary> will visit Mary.
b. I expect (that) everyone will visit Mary [that you do <expect will visit Mary>].

(Tiedeman 1995:75)

Let me examine the sentences in (47) to illustrate how Fox’s analysis captures Tiedeman’s fact. In the acceptable ACD in (47)b, the embedded subject QP undergoes QR, which is followed by Late Merge of the relative clause, and the tail of the QR chain is pronounced, just like in Fox and Nissenbaum’s (1999) derivation of adjunct Extraposition:24

(49)  [John believed [everyone] was a genius] 
→ QR  
[[John believed [everyone] was a genius] [everyone]]  
→ Late Merge  
[[John believed [everyone] was a genius] [everyone [that you did <believe was a genius>]]]

As its surface representation indicates, the unacceptable ACD in (47)a does not involve Late Merge of the relative clause. The structure of (47)a before QR applies to the embedded subject would be the one like (50). Given this structure, QR does not help us produce a structure that satisfies the parallelism condition in (47)a.

(50)  [John believed that [everyone that you did <believe was a genius>] was a genius]

24 In most cases, QR obeys the clause boundedness condition, which is circumvented here. See 4.8.3 for relevant discussion. See also Cecchetto (2004), Fox (2000), May (1988), Moltmann and Szabolcsi (1994), and Reinhart (1991) for relevant facts and discussion.
Tiedeman’s observation that Extraposition is relevant for ACD resolution is entirely puzzling in approaches to ACD that postulate QR, but do not assume that Late Merge of relative clauses is involved in the derivation of ACD.

The second piece of evidence in favor of Fox’s approach comes from Fiengo and May’s (1994) fact that QR bleeds Condition C in ACD (see also Fox 1995b, 2000). This is illustrated in (51).

(51)  a. You sent himi the letter that Johni expected you would <send himi>.
    b. You introduced himi to everyone Johni wanted you to <introduce himi to>.
    c. I reported himi to every cop Johni was afraid I would <report himi to>.

    (Fox 2002:84)

The Condition C bleeding fact in ACD is also directly accounted for by Fox’s Late Merge analysis. In this approach, it can be dealt with on a par with the same effect found in wh-movement in (52) and in adjunct Extraposition in (53) because all of these cases involve Late Merge of relative clauses.

(52)  a. Which story that Johni wrote did hei like?
    b. Which argument that Johni made did hei believe?

(53)  a. I gave himi a picture yesterday [from Johni’s collection].
    b. I gave himi an argument yesterday [that supports Johni’s theory].

Owing to Late Merge, the relative clause is not represented in the c-command domain of the pronoun, as illustrated in (54), and thus, it avoids violating Condition C.
In the next subsection, I will discuss further facts that support Fox’s Late Merge approach. However, we will also see that a variant of those facts could be taken as a challenge to the copy theory of movement. I will argue that the Wholesale Late Merge analysis accounts for it within the copy theory of movement framework.

3.4.3 The Lexical Content Effect in Ellipsis

We have seen that Fox’s approach to ACD generates a structure that satisfies the parallelism condition within the copy theory of movement framework. To accomplish this result, the essential component of Fox’s analysis is the Late Merge operation, which allows us not to represent relative clauses in the tail position of a QR chain. It is however crucial to notice that QR does leave a copy of a head noun of a relative clause, and its lexical content is represented within the antecedent VP. This part of Fox’s analysis has a consequence for explaining Kennedy’s (1994, 2004) and Sauerland’s (1998, 2004) facts.

To discuss Kennedy’s and Sauerland’s facts, let us first consider the ungrammatical ACD case in (55), which is first discussed in Kennedy (1994), and further explored in Heim (1997), Jacobson (2004a), Kennedy (2004), Sauerland (1998, 2004), and Williams (1995).

(55) *John visited a town that’s near the lake Mary did <visit>.

(Sauerland 1998:104)
Sauerland (1998, 2004) suggests that what makes the ACD in (55) ungrammatical is that the head noun of the object of the matrix predicate is semantically distinct from that of the elided predicate within the relative clause, and argues that this property has a consequence for licensing ellipsis of EC in (55). Sauerland's suggestion is corroborated by his fact that when head nouns of two objects are semantically identical, ACD of this sort is grammatical, as illustrated in (56). In (56)b, the head noun one is anaphoric to town, and hence, the two head nouns are semantically identical.

(56)  
   a. John visited a town that's near the town Mary did <visit>.  
   b. John visited a town that's near the one Mary did <visit>.  

(Sauerland 1998:104)

These facts lead us to draw the following generalization:25

(57) **Sauerland's Generalization**  
When an elided VP (VP_E) is dominated by an argument DP of an antecedent VP (DP_A), ellipsis of VP_E is licensed only if a head noun of DP_A is semantically identical to that of the corresponding argument DP of VP_E.

Sauerland (1998, 2004) argues that (57) is captured as a consequence of the copy theory of movement. In the ungrammatical ACD in (55), QR leaves a copy of town in the complement of the matrix predicate, and movement of a head noun within the relative

---

25 The difference in a lexical content of argument constituents between an elided VP and an antecedent VP has an effect on licensing ellipsis only in the so-called Argument Contained Ellipsis (ACE) context (Kennedy 1994, 2004). The ACE context is described in (57). The lexical content effect does not emerge in cases of ellipsis other than ACE (Evans 1988, Jacobson 1992a, 2004a, and Sauerland 1998, 2004, among others):

(i)  
   b. I know which cities Mary visited, but I have no idea which lakes she did <visit>.  
   c. The cities Mary visited are near the lakes Bill did <visit>.  

(Sauerland 1998:138)

See Sauerland (1998, 2004) for possible reasons why the lexical content effect is not found in these cases.
clause leaves a copy of lake in the complement of the predicate in the relative clause (see fn. 23 for Sauerland’s matching analysis of relative clauses). Thus, Sauerland suggests that the LF representation postulated for (55) is one close to (58), which does not satisfy the parallelism condition.

(58)  [[a town that’s near the lake [CP [OP lake] [Mary did <visit lake>]]]
[John visited town]]

In contrast to (55), QR and movement of a head noun within the relative clause leave the copies of the semantically identical NP in the acceptable ACD in (56), as shown in (59), and hence, one can find a Parallelism Domain that licenses ellipsis of EC in this case.26, 27

(59)  [[a town that’s near the lake [CP [OP town] [Mary did <visit town>]]]
[John visited town]]

Sauerland’s Generalization is taken as further evidence in favor of the copy theory of movement. However, we have also discussed the conflict between the copy theory of movement and the parallelism requirement in ACD. What component of grammar allows us to generate such structures as (58) and (59)?

As has been made clear by now, Fox’s Late Merge analysis supplies a straightforward way to achieve this result. As discussed in the last subsection, structures of this sort are direct consequences of Fox’s approach, in which only a DP external to a relative clause undergoes QR, and the relative clause Late Merges with the QRed DP, as illustrated in (60) and (61).28 (For the purpose of presentation, QRed material is sometimes represented to the left of its original position.)

---

26 Heim (1997) and Kennedy (1994, 2004) propose different solutions to cases like (55), which do not make use of the consequence of the copy theory of movement. The facts in (56) are not expected in their approaches, as Sauerland (1998, 2004) notes.

27 Jacobson (2004a) discusses a contextual influence on the lexical content effect in the ACE context, which I hope to investigate on another occasion.

28 To produce the LF representations given in (60) and in (61), Sauerland (1998) adopts Fox’s (1995b) earlier analysis of ACD.
Sauerland’s Generalization follows straightforwardly from Fox’s Late Merge approach in a way that is compatible with the copy theory of movement.

However, Sauerland himself also provides an exception to Sauerland’s Generalization. In (62), ellipsis of EC is licensed even though the head noun of the matrix subject (town) is not semantically identical to the head noun of the relative clause (lake), just like in (55).

(62) The town near the lake that was visited by vandals seems to have been <visited by vandals>, as well.

(Sauerland 1998:112)

This is a challenging fact to the above mentioned account of Sauerland’s Generalization. Given the assumptions so far, the LF representation posited for (62) would be something close to (63), and, as in (60), there is no Parallelism Domain that is required for licensing ellipsis of EC:
Notice that (62) is a passive version of (55). Thus, it involves A-movement of the matrix subject DP and the head noun out of the elided VP and the antecedent VP, respectively. Sauerland capitalizes on this property of (62), and argues that this exception to his generalization is an incarnation of the general property of A-movement that it optionally leaves a trace, as has independently been suggested by the fact that A-movement always bleeds Condition C. Due to this property, the lexical contents of the two DPs are not represented inside the antecedent VP and the elided VP, and the parallelism condition is met in (62).

We now have a more principled account of this property of A-movement, and in fact, the Wholesale Late Merge approach can explain the difference between (55) and (62). The Wholesale Late Merge approach allows us to postulate the derivation in (64), which explains why the lexical contents of the relevant DPs do not have any consequence for licensing ellipsis of EC in (62). In (64), only determiners (i.e., the definite determiner the and the null operator) undergo movement as the first step of A-movement, and their restrictor NPs are introduced after they move out of the VPs that are relevant for satisfying the parallelism condition (see Legate 2003 for arguments in favor of the claim that a VP-adjoined position is an intermediate landing site in A-movement in passives). Consequently, the underlined λ-predicates in (64)a and in (64)b are completely identical, and this licenses ellipsis of EC in (62).
Notice that in (55), the head nouns of the two relative clauses are the objects of the predicates, and they receive Case from the functional head immediately above the predicates. This forces the restrictor NPs to be introduced in the complement position of the predicates. Therefore, Wholesale Late Merge is not an option in (55), and the lexical content effect inevitably emerges there. I have demonstrated here that both Sauerland’s Generalization and its exception are accounted for by the Wholesale Late Merge approach within the copy theory of movement framework.

### 3.4.4 The Complement/Adjunct Asymmetry

Given the Wholesale Late Merge account of the exception to Sauerland’s Generalization, repeated here as (65), we predict that whenever the two relevant DPs undergo A-movement out of an elided VP and an antecedent VP, the lexical contents of their restrictor NPs will not have any effect upon licensing ellipsis of VP_E.
When an elided VP (VP_E) is dominated by an argument DP of an antecedent VP (DP_A), ellipsis of VP_E is licensed only if a head noun of DP_A is semantically identical to that of the corresponding argument DP of VP_E.

In this subsection, I discuss another case that supports the Wholesale Late Merge approach to exceptions to Sauerland’s Generalization.

Let us first consider the sentences in (66). In (66)a and (66)b, the elided VP is dominated by the relative clause that modifies the matrix object DP. In contrast, it is included in the sentential complement of the matrix object DP in (66)c.

(66)  a. I made an argument that was very similar to the argument you did <make>.
    b. I made an argument that was very similar to the one you did <make>.
    c. *I made an argument that we should adopt the argument you did <make>.

(Fox 2002:82)

Thus, the two instances of argument in (66)c are different with respect to their complementation property, and, hence, the contrast between (66)a and (66)b, and (66)c is subsumed under Sauerland’s Generalization. Furthermore, it also follows from Fox’s Late Merge analysis. As discussed in detail in the previous sections, argument constituents of NPs cannot be Late Merged because Late Merge of this sort ends up producing uninterpretable LF representations. Thus, the entire object DP in the matrix clause needs to undergo QR, as shown in (67). Consequently, no Parallelism Domain can be found in the LF representation postulated for (66)c, leading to failure in licensing ellipsis of EC.

(67)  *[an argument that we should adopt the argument
      λx. you did <make [the argument x]>]
      λy. [I made [the argument that we should adopt the argument λx. you did
            <make [the argument x]> y]]]
Let us now move onto cases that could be regarded as an additional exception to Sauerland's Generalization. Kennedy (2004) observes the facts like (68), which suggest that a sentential complement of an NP does not have to be represented within a predicate phrase if a constituent undergoes A-movement out of the elided VP (see also Hardt and Asher 1997, Heim 1997, Kennedy 2004, and Sauerland 2004 for discussion of subject ACE cases).29 If the sentential complement of the NP were represented in the original position of the moved subject, as in (69), the parallelism condition would not be satisfied, just like in (67), contrary to fact.

(68)  a. ?Your proof that my proof is valid is <valid>, as well.
     b. ?History suggests that a proof that God exists never will <exist>.

         (Kennedy 2004)

(69) *[[your proof that [my proof
     \lambda x. is [the proof x] valid]]
     \lambda y. is <[AP your proof that [my proof \lambda x. is [the proof x] valid] y] valid]>]

Within the proposed framework, this exception can be treated on a par with the exception discussed in the last subsection. Since the relevant DPs undergo A-movement out of the predicate phrases that are relevant for satisfying the parallelism condition in (68), the Wholesale Late Merge approach can generate a derivation in which the restrictor constituents of the determiners are not represented in the predicate phrases, as illustrated in (70). In a derivation of this sort, there are two semantically identical \( \lambda \)-predicates (the underlined ones in (70)), which correctly licenses ellipsis of EC in (68).30

29 Kennedy (1994) argues that subject ACE cases like (68) are ungrammatical. In Kennedy (2004), they are, however, regarded as grammatical, though marginal. The proposed account treats them as grammatical and does not capture their marginal status. I need to leave this issue to future research.

30 I assume that possessive expressions are covert definite descriptions, and that they involve a covert definite determiner, which is overtly represented in (70) for expository purposes. See Elbourne (2001, 2005) for the relevant claim that all individual denoting expressions are overt/covert definite descriptions. Here, I also assume that an AP-joined position is a possible landing site for A-movement, extending Legate (2003).
As is clear by now, Wholesale Late Merge is not an option in (66)c because of the Case constraint on NPs. In this subsection, I have demonstrated that the Wholesale Late Merge approach allows us to explain another exception to Sauerland’s Generalization within the copy theory of movement framework.

3.5 A Wholesale Late Merge Approach to the ACD Puzzle

3.5.1 An Outline

In 3.5, we examine some puzzling ACD facts discussed in Wold (1995). As we will see shortly, these facts seem to suggest that QR sometimes leaves a contentless trace. Since this is reminiscent of the property of A-movement discussed above, I will extend the Wholesale Late Merge approach to those ACD cases. Although this approach does not
address Wold’s puzzle completely, I would like to suggest that the Wholesale Late Merge approach could be taken as the first step towards solving Wold’s puzzle.

### 3.5.2 The ACD Puzzle

In 3.4.2, we discussed Fox’s (2002) Late Merge approach to ACD, in which it is argued that a DP first undergoes QR, and then a relative clause that dominates an elided VP Late Merges with the QRed DP, as repeated in (71).

\[(71) \quad \text{Polly visited every city that Uli did <visit>}. \]
\[\text{[Polly visited [every city]]} \]
\[\text{→ QR} \]
\[\text{[[λx. [Polly visited [the city x]]] [every city]]} \]
\[\text{→ Late Merge} \]
\[\text{[[λx. [Polly visited [the city x]]] [every city [λy. Uli did <visit the city y>]]]} \]

As shown in (71), Fox’s approach derives an LF representation of ACD that satisfies the parallelism condition within the copy theory of movement framework. We have seen that it is supported by various facts, as well. Among others, the Late Merge approach straightforwardly accounts for Tiedeman’s fact repeated in (72), which indicates the necessity of Extraposition in ACD.

\[(72) \quad \begin{align*}
\text{a. } & \quad * \text{John believed (that) everyone that you did <believe was a genius> was a genius.} \\
\text{b. } & \quad \text{John believed (that) everyone was a genius that you did <believe was a genius>}. 
\end{align*} \]

Given Fox’s approach to ACD, we predict that ACD and Extraposition of relative clauses would go hand in hand. In other words, we expect that ACD is possible only in contexts where Extraposition of relative clauses is allowed. Building on the observations
made by Hulsey (2001) and Hulsey and Sauerland (to appear), I discuss below a case which suggests that this prediction is not borne out. To set the stage for discussing this puzzle, let us first consider the sentence in (73), discussed in Schachter (1973), among others. It demonstrates that a variable within a head noun of a relative clause can be bound by an element within the relative clause:

(73) The portrait of himself, that John, painted is extremely flattering.

(Schachter 1973:32)

Particularly relevant for our discussion is the fact that a variable binding dependency of this sort prevents a relative clause from being extraposed, as argued in Hulsey (2001) and Hulsey and Sauerland (to appear):

(74) a. I saw the picture of himself, that John, liked.
    b. *I saw the picture of himself, yesterday [that John, liked].

(75) a. Mary discovered the book about himself, that Bob, wrote.
    b. *Mary discovered the book about himself, yesterday [that Bob, wrote].

(Hulsey and Sauerland to appear)

We are now ready to discuss an ACD puzzle. Wold (1995) discusses the fact in (76), which suggests that the relevant variable dependency is compatible with ACD (see also Sauerland 1998:73):

(76) Sue likes every picture of himself, that every boy, hoped that she would <like>.

(Wold 1995 attributing to Danny Fox)

31 As Hulsey and Sauerland discuss, the presence of a variable in a head noun itself does not block Extraposition:

(i) I saw the picture of myself yesterday [that John liked].

(Hulsey and Sauerland to appear)
Taking into consideration the ungrammaticality of (overt) Extraposition in (74)b and (75)b, the fact that the ACD in (76) is grammatical is unexpected in Fox’s approach. What makes this ACD puzzle even more puzzling is the fact in (77), which is also discussed in Wold (1995). It indicates that a variable binding dependency of the relevant sort does sometimes block ACD:

(77) *Sue admires every picture of himself, that every boy, does <admire>.

(Wold 1995)

Wold (1995) draws from these observations the generalization that when a head noun of a relative clause contains a variable bound by a relative clause internal element, ACD is licensed only if the variable is bound by a binder that is outside the Parallelism Domain (Wold’s Generalization). To illustrate Wold’s Generalization, let us consider the (somewhat simplified) LF representations in (78) and (79), postulated for (76) and (77), respectively. The λ-predicate introduced with λx is a potential candidate for a PD both in (78) and in (79). In (78), the variable can be bound by the universal QP that is outside this PD, for example, at the most embedded CP domain. In contrast, there is no such position in (79). The variable is bound by the QP in the complement of the relative clause internal predicate, but it is inside the PD.

(78) ([[every picture of himself, that every boy, hoped that [PD λx. she would <like x>]]
    λy. [Sue likes y]]

(79) ([[every picture of himself, that [PD λx. every boy, does <admire x>]]
    λy. [Sue admires y]]

Wold’s Generalization is further corroborated by the fact in (80) in which the most embedded subject Bill cannot be an antecedent of the reflexive anaphor, but the subject John can.
(80) Mary admires every picture of himself\(_i\) that John\(_i\) claimed that Bill\(_j\) does <admire>.

(Wold 1995 attributing to Danny Fox)

If Bill were an antecedent of the reflexive anaphor, it would have to be bound within the PD. Thus, ACD in (80) is not licensed under this interpretation:

\[
\text{(81) } \left[ \left( \text{every picture of himself}_{i/j} \text{ that John}_{i} \text{ claimed that } \left[ \text{PD } \lambda x. \text{ Bill}_{j} \text{ does } <\text{admire } x> \right] \right) \right] \\
\lambda y. \left[ \text{Mary admires } y \right]
\]

Extending Fox’s Late Merge approach, I make an attempt to develop a possible approach to this ACD puzzle. As we will see below, this approach makes a problematic prediction. However, I would like to suggest that it can be taken as the first step towards capturing Wold’s Generalization. I argue that the counter-cyclic merger of relative clauses plays an essential role in Wold’s cases, as well, but it takes place in a different form from that of Late Merge in Fox’s approach. More specifically, Wholesale Late Merge comes into the picture in Wold’s cases. As I have argued above, movement of a determiner leaves a contentless copy when it is followed by Wholesale Late Merge of its restrictor constituent. Capitalizing on this and other properties of grammar, which we will discuss shortly, I suggest that in the structure of the Wold’s acceptable ACD case in (76), a copy of this sort occupies the syntactic positions inside the domain relevant for the parallelism condition, and, hence, its LF representation is very close to (78). As is clear from (78), LF representations of this sort satisfy the parallelism condition. In contrast, an LF representation analogous to (79) cannot be postulated for the unacceptable ACD in (77), even within the framework that I will suggest below. The difference between the two cases has to do with where a constituent involving a bound variable is interpreted, as suggested in Wold’s Generalization.

I have argued that when QR applies to a constituent in object position, it cannot leave a contentless copy because the Case constraint requires a restrictor NP to be introduced before QR. I will suggest below that QR of a constituent in object position can leave a contentless copy when a restrictor constituent of a determiner is a constituent
that does not need Case. This contributes to producing LF representations close to (79). To set the stage for making this suggestion, I first discuss Hulsey’s (2001) and Hulsey and Sauerland’s (to appear) analysis of the fact that Extraposition is unacceptable in (74)b and (75)b.

3.5.3 The First Component: A Raising Structure of Relative Clauses

What is a structural property of the relative clauses in (74) and (75) that might account for the fact that Extraposition is impossible in (74)b and (75)b? A possibility that suggests itself is that in those cases the head noun of the relative clauses is exclusively interpreted in a relative clause internal position because the head noun contains the variable bound by the subject within the relative clause. For this reason, among others, Sauerland (1998) and Hulsey and Sauerland (to appear) claim that a so-called raising structure should be available for restrictive relative clauses, which is exemplified in (82)a.\(^\text{32}\) It is crucial that there is no occurrence of the head noun external to the relative clause CP in a raising structure and that the moved head noun in the Spec of CP is exclusively interpreted inside the relative clause CP as a consequence of the copy theory of movement, which yields the desired interpretation in (82)b.

It is clear that the matching structure of relative clauses discussed in the previous sections cannot be postulated for cases in which a head noun must receive its interpretation only relative clause internally. For example, if a matching structure in (83)a is postulated for (74)a, the variable in the CP-external occurrence of the head noun remains unbound because the binder is within the relative clause and the CP-external occurrence of the head noun does not participate in a movement dependency with the CP-internal one. This results in the uninterpretable LF representation in (83)b.

Investigating various other instances of a raising structure, Hulsey (2001) draws the generalization that a raising structure blocks Extraposition of relative clauses (Hulsey’s Generalization).
Hulsey (2001) and Hulsey and Sauerland (to appear) argue that Hulsey’s Generalization straightforwardly follows from Fox and Nissenbaum’s (1999) approach to adjunct Extraposition. As discussed above, Fox and Nissenbaum argue that Extraposition of relative clauses involves QR of a DP followed by Late Merge of a relative clause. The discontinuity between the DP and the relative clause is derived by pronouncing the tail of the QR chain:

\[(84) \quad [I \text{ saw } [a \text{ picture } \text{ yesterday}]]\]

\[\rightarrow \text{ QR & Late Merge}\]

\[\quad [[I \text{ saw } [a \text{ picture } \text{ yesterday}]] [a \text{ picture } [cP \text{ that John liked}]]]\]

\[\rightarrow \text{ Pronunciation of the tail of the QR chain}\]

\[\quad [[I \text{ saw } [a \text{ picture } \text{ yesterday}]] [a-pi\text{tc}l\text{ure } [cP \text{ that John liked}]]]\]

As is clear from (84), Fox and Nissenbaum’s approach imposes the necessary condition for deriving Extraposition. Namely, Extraposition of a relative clause is possible only when the head noun of the relative clause is external to the relative clause CP. Since there is no such DP involved in a raising structure, as exemplified in (82)a, a raising structure blocks Extraposition of relative clauses (i.e., Hulsey’s Generalization).³³

Wold’s ACD cases still remains puzzling because in Fox’s (2002) approach, a CP-external DP is also prerequisite for the Late Merge process that derives a structure that satisfies the parallelism condition in ACD. However, the next subsection suggests that the proposal in Hulsey (2001) and Hulsey and Sauerland (to appear) help us develop a possible approach to Wold’s cases, which makes crucial use of Wholesale Late Merge.

³³ Movement cannot derive Extraposition, either when a relative clause involves a raising structure. The intermediate projection C’, which is assumed to be immovable, would need to undergo movement in order to produce the observed linear order.
3.5.4 The Second Component: Wholesale Late Merge

As argued in Hulsey (2001) and Hulsey and Sauerland (to appear), in a raising structure for relative clauses, there is no CP-external DP that can undergo QR and merge with a relative clause counter-cyclically. Thus, the version of Late Merge adopted in Fox (2002) would not help us derive a structure that meets the parallelism requirement in Wold’s cases. Notice however that there is a determiner that is external to the CP in a raising structure (see the representation in (82)a). We have already argued that determiners are allowed to move by themselves and that their restrictor constituents can Wholesale Late Merge with the determiners in their derived position. I suggest that the first step towards developing a possible approach to Wold’s puzzle is to extend this Wholesale Late Merge approach to the ACD case above, as illustrated below:

(85)  

a. Sue likes every picture of himselfi that every boyi hoped that she would <like>.

b. [Sue likes [every]]

[move of a determiner]

[[Sue likes [every]] [every]]

[Wholesale Late Merge]

[[Sue likes [every]] [every [CP [picture of himselfi] that every boyi hoped that she would <like>]]]

[Trace Conversion]

[[X. [Sue likes [the x]]] [every [CP [picture of himselfi] that every boyi hoped that she would <like>]]]

Here, the Wholesale Late Merge approach capitalizes on a structural property of a raising structure of relative clauses, namely, that the category of a restrictor constituent of a determiner is a CP. The essential property of CPs relevant for this context is that they do not need Case. Thus, the restrictor constituent of a determiner in a raising structure can be introduced at any stage of the derivation. The relevance of this structural property
can be appreciated by comparing (85) with the cases like (86)a. We have claimed that since the restrictor NP of the determiner *which* needs to receive Case from the functional head directly above the predicate (i.e., *v*), the derivation involving Wholesale Late Merge in (86)b is not viable.

(86)  
a. ??/*Which argument that Johni is a genius did he, believe?  
b. *[v believe [which]] 
   \[\rightarrow \text{movement of a determiner & Wholesale Late Merge}\]  
   *[[which [NP argument that Johni is a genius]] [he, v believe [which]]]*

As illustrated in (85)b, the Wholesale Late Merge approach contributes to resolving the antecedent containment problem in ACD. However, it is not obvious whether the parallelism condition is satisfied in (85)b, especially because we have not made explicit the internal structure of the relative clause. Let us first consider the possibility that the head noun in the Spec of CP is interpreted in its original position as a consequence of the copy theory of movement, as shown in (87).\(^{34}\) If the LF representation in (87) is postulated for the relative clause in (85)b, the parallelism condition would not be met.

(87) \[[cP \lambda y. \text{that every boy}_i \text{hoped that she would } <\text{like } \text{[the picture of himself]}_i y>]\]

As suggested by Wold’s Generalization, the head noun should not be interpreted in such a low position. More specifically, if the $\lambda$-predicate sister to the QRed QP in (85)b is a domain relevant for the parallelism condition, the head noun should be interpreted

\(^{34}\) The LF representation in (87) is derived by the processes illustrated in (i). Since the moved head noun does not contain a determiner in a raising structure of relative clauses, Trace Conversion adds a definite determiner into the lower copy of the head noun:

\[(i) \quad [cP [\text{picture of himself}]_i \lambda y. [TP \text{every boy}_i \text{hoped that she would like } [\text{picture of himself}_i]]] 
\[\rightarrow \text{deletion of the higher copy}\]  
\[[cP [\text{picture of himself}_i] \lambda y. [TP \text{every boy}_i \text{hoped that she would like } [\text{picture of himself}_i]]] 
\[\rightarrow \text{Trace Conversion}\]  
\[[cP \lambda y. [TP \text{every boy}_i \text{hoped that she would like } [\text{the picture of himself}_i y]]]\]
outside the corresponding domain in the relative clause. For example, one such structure is depicted in (88).

(88) \[\text{[\text{CP that every boy} hoped [\text{CP [picture of himself]} [\lambda y. that she would like y]]]}\]

In the next subsection, I suggest another property of grammar, which allows us to postulate a structure of relative clauses analogous to the representation above.

Before discussing this property of grammar, I need to make some comments on the issue of the word orders predicted by the Wholesale Late Merge approach. The Wholesale Late Merge approach captures Hulsey’s Generalization, just as Hulsey (2001) and Hulsey and Sauerland (to appear) approaches do. If a determiner moves across an adverbial, the Wholesale Late Merge approach would produce the two word orders in (89), depending upon which copy of the determiner is pronounced. The heavy NP shift word order in (89)a results if the head of the chain is pronounced. However, if the tail of the chain is phonologically realized, it ends up in the ungrammatical word order in (89)b.

(89)  a. I saw yesterday the picture of himself that John liked.
       b. *I saw the yesterday picture of himself that John liked.

The ungrammatical word order such as (89)b has been noticed as a potential problem to approaches that assume that the interpretability condition on LF representations determines when Late Merge is possible (Bhatt and Pancheva 2004:fn. 33 and Fox 2002:fn. 11). I do not have a principle solution to this problem. However, I speculate that the sentence in (89)b might indicate that movement of determiners is always overt, just like most cases of movement. If this is the right way to approach this issue, we would need to explore what properties of grammar conspire to make movements such as QR of DPs covert.
3.5.5 The Third Component: Null Operator Movement

In the previous subsection, we have seen that Wholesale Late Merge allows us to postulate the structure in (90) for the Wold’s acceptable case. The antecedent containment relation is resolved in this structure.

\[(90) \quad [[\lambda x. \text{[Sue likes [the x]]}] \text{[every [CP [picture of himself]], that every boy, hoped that she would <like t,>]]}]\]

If this suggested approach is the right way to solve the antecedent containment problem in Wold’s ACD case, it implies that there must be a \(\lambda\)-predicate inside the relative clause, which is semantically identical to the \(\lambda\)-predicate in the matrix clause (\(\lambda x\)), modulo focus marked material. Since the complement position of like within the relative clause is inside the elided constituent, a focus marked element would not be present in this position. This suggests that the complement position in the relative clause must be occupied by a contentless copy, as is the corresponding position in the matrix clause. Wold’s Generalization also seems to indicate the same point.

Independently of Wold’s ACD cases, there is a phenomenon which leads us to the impression that relative clause internal movement leaves a contentless trace, rather than a copy of a head noun. As illustrated in (91) and (92), a violation of Condition C is not induced by a pronoun within a relative clause, which is co-referential with an R-expression inside the head noun of the relative clause (Merchant, 2000, Munn 1994, Safir 1999, and Sauerland 1998, 2003b, 2004, among others):

\[(91) \quad \text{In pictures of Ali, which he, lent to us, he is shaking hands with the president.}\]
\[(92) \quad \text{a. I have a report on Bob’s division that he, won’t like.}\]
\[(92) \quad \text{b. I read every report on Bob’s division he, ever submitted.}\]
Whether the relative clauses in (91) and (92) are analyzed as a matching structure or as a raising structure, the absence of a Condition C violation is unexpected if movement of the head noun leaves a copy within the relative clauses:

(93) a. a matching structure

\[ *_{[\text{DP a report on Bob}_i \ [\text{CP [report on Bob}_i]_i\ [\text{TP he}_i \text{ ever submitted [report on Bob}_i]]]} \]

b. a raising structure

\[ *_{[\text{DP a [CP [report on Bob}_i]_i\ [\text{TP he}_i \text{ ever submitted [report on Bob}_i]]]} \]

For these reasons, I suggest that an alternative structure must be available for relative clauses, in which the lexical content of a head noun is not represented relative clause internally. One possible approach is to postulate a structure of relative clauses that involves null operator movement. As illustrated in (94), structures of this sort may be able to explain the fact that Condition C is not violated in cases such as (91) and (92) (see Sauerland 1998, 2003b, 2004 for a different analysis of the Condition C fact above).

(94) the null operator movement structure

\[ [\text{DP a report on Bob}_i \ [\text{CP OP}_1 \lambda x. [\text{TP he}_i \text{ ever submitted OP}_1]]] \]

\[ \rightarrow \text{Trace Conversion} \]

\[ [\text{DP a report on Bob}_i \ [\text{CP \lambda x. [TP he}_i \text{ ever submitted [the x]]}]] \]

This null operator movement also helps us in deriving a structure of Wold’s acceptable ACD case that meets the parallelism requirement. I suggest that a null operator moves from the object position to the most embedded CP domain, and the head noun is introduced into the structure above the moved null operator. (For concreteness,

---

35 In 3.4, we have seen evidence from ellipsis in favor of the effects of the lexical content of head nouns in relative clauses. I will discuss this problem within the current framework in the next subsection.

36 Note that the NP picture of himself is introduced into the structure outside of the c-command domain of the head that assigns Case to object DPs. I will discuss how the NP in question receives Case in 3.6.
the head noun is introduced in the most embedded CP domain in (95)b. Note however that the same result can be obtained if it is introduced in any position between the moved null operator and the surface position of the universal QP every boy.

\[(95)\]
\[a.\] Sue likes every picture of himself that every boy hoped that she would <like>.

\[b.\] \[TP she would <like OP>\]

\[\rightarrow\] null operator movement
\[\text{CP } \text{OP}_1 [\text{TP she would } <\text{like OP}_1>]\]

\[\rightarrow\] insertion of a head noun
\[\text{CP [picture of himself]} [\text{CP } \text{OP}_1 [\text{TP she would } <\text{like OP}_1>]]\]

\[\rightarrow\] movement of the head noun
\[\text{CP [picture of himself]}_2 [\text{TP every boy hoped } [\text{CP [picture of himself]}_2 [\text{CP } \text{OP}_1 [\text{TP she would } <\text{like OP}_1>]]]]\]

\[\rightarrow\] Trace Conversion
\[\lambda x. [\text{TP every boy hoped } [\text{CP [the picture of himself]} [\lambda y. [\text{TP she would } <\text{like [the y]>]]]]]\]

Due to this null operator movement, ellipsis of EC in the Wold’s acceptable ACD case is now correctly licensed because the \(\lambda\)-predicate \(\lambda y\) in the final representation in (95)b can be taken as a Parallelism Domain. It is semantically identical to the relevant \(\lambda\)-predicate in the matrix clause in (96), modulo the focus marked auxiliary \textit{would}. (The structure in (96) illustrates the stage of the derivation before the relative clause in (95)b Wholesale Late Merges with the moved determiner.)

\[(96)\] \[\langle\lambda z. [\text{Sue likes [the z]}] [\text{every}]\]

Let us now move onto the Wold’s unacceptable ACD case repeated here as (97)a. Crucially, null operator movement of the same sort cannot take place. In this case, the head noun needs to be represented in the c-command domain of the subject QP in the
relative clause so that the variable is bound by the QP, as illustrated in (97)b.37 The structure of the matrix clause in (97)a is the same as the one depicted in (96), and there is no constituent in (97)b that is semantically identical to (96), modulo focus marked material. Consequently, ellipsis of EC is not licensed in (97)a.

\[(97) \quad a. \quad *\text{Sue admires every picture of himselfi that every boyi does <admire>.} \\
\quad b. \quad \text{[TP every boyi does <admire [picture of himselfi]>]} \]

\[\rightarrow \text{movement of the head noun} \]

\[\text{[CP [picture of himselfi], [TP every boyi does <admire [picture of himselfi]>]]} \]

\[\rightarrow \text{Trace Conversion} \]

\[\lambda x. \text{[TP every boyi does <admire [picture of himselfi] x]>]} \]

We have seen that the suggested extension of the Wholesale Late Merge approach captures Wold's Generalization, together with the ancillary assumptions. However, the next subsection will reveal that the current analysis of ACD is too weak, and cannot account for the cases that have been explained by making essential use of the lexical content effect as a consequence of movement. The source of the problem is that the current system allows movement to leave a contentless copy in too many cases. As we will see in the next subsection, this problem must remain for future research.

37 For the head noun to be represented in the c-command domain of the subject QP in the relative clause, it does not have to be introduced in the complement position of the predicate, as in (97). An alternative derivation would be that null operator moves from the complement position to the vP-adjoined position and the head noun is inserted at the vP-adjoined position, as illustrated in (i). However, this structure does not satisfy the parallelism condition, either.

\[(i) \quad \text{[CP [picture of himselfi], [TP every boyi [vP [picture of himselfi], [vP OP2 [vP ti admires OP2]]]]]} \]
3.5.6 Further Issues

We have seen that Wold’s Generalization is captured if movement sometimes leaves a contentless copy. In the last subsection, I have suggested that the interplay of Wholesale Late Merge and relative clause internal null operator movement is one possible way to instantiate this idea. However, this analysis of Wold’s ACD cases appears to conflict with the proposed analysis of Sauerland’s Generalization (see 3.4.3), which is based on the contrast between (98)a and (98)b, among others.

\[
\begin{align*}
\text{(98) a. } & \text{ *Polly visited every town that’s near the lake Eric did } \text{ <visit>.} \\
\text{b. } & \text{ Polly visited every town that’s near the town Eric did } \text{ <visit>.}
\end{align*}
\]

(Sauerland 1998:19)

To capture Sauerland’s Generalization, it is crucial that both QR in the matrix clause and relative clause internal movement leave a copy which includes the lexical content of the relevant restrictor constituent (i.e., town and lake in (98)a). Indeed, we will see shortly that the suggested analysis of Wold’s Generalization is too weak in that it cannot account for the ellipsis facts that have been explained by utilizing the lexical content effect as a consequence of movement (including Sauerland’s Generalization). As will be made clear below, the analysis suggested in the last subsection allows movement to leave a contentless copy in too many cases. This problem must remain as a topic for future research. Here, I will suggest postulating one condition on the availability of a raising structure of relative clauses, which is one way to restrict the applicability of movement that leaves a contentless copy. Before discussing this condition, I illustrate the problem that the suggested analysis of Wold’s Generalization is confronted with by examining two cases which are evidence for the lexical content effect in ellipsis.

Let us first consider (98), one of the cases subsumed under Sauerland’s Generalization. We have assumed that the lexical contents of the relevant DPs are represented within EC and AC as a result of movement of the DPs. This ends up
producing a structure that does not satisfy the parallelism condition in (98)a. The derivation in (99) is postulated for (98)a within this framework.

(99)  [John visited [a town]]

→ QR

[[a town] λx. [John visited [the town x]]]

→ Late Merge

*[[a town [that’s near the lake [[OP lake] λy. [Mary did <visit [the lake y>]]]]]

λx. [John visited [the town x]]]

I have pursued this analysis of (98)a, and argued that movement of the determiner followed by Wholesale Late Merge is impossible in (98). Because the restrictor constituent of the determiner is an NP in (98), it must be introduced in the complement of the predicate in order to obey the Case constraint. However, given a raising structure of relative clauses and null operator movement within a relative clause, there are more derivations that can be postulated for (98). Specifically, the derivation in (100), which is posited for (98)a, is particularly important for our purposes because it incorrectly generates a structure that satisfies the parallelism condition. In (100)a, the determiner undergoes QR by itself under the assumption that the relative clause involves a raising structure. In this case, the restrictor constituent of the determiner does not need Case, and can Wholesale Late Merge with the QRed determiner. As a consequence, the complement of the matrix predicate is occupied by a contentless copy. Analogously, null operator movement leaves a copy of the same sort in the corresponding complement position within the relative clause, as shown in (100)b. As one can easily see, the resulting LF representation would meet the parallelism requirement, contrary to fact.
Let us consider another case problematic to my analysis of Wold’s Generalization. We have discussed the fact that QR in ACD bleeds Condition C, as repeated in (101) (Fiengo and May 1994 and Fox 1995b, 2002).

(101)  

(a) You sent him, the letter that John, expected you would <send him,>.  
(b) You introduced him, to everyone John, wanted you to <introduce him, to>.  
(c) I reported him, to every cop John, was afraid I would <report him, to>.

Building on this fact, Merchant (2000) and Sauerland (1998) argue that QR in ACD bleeds Condition C when the relevant R-expression is included in a relative clause, but not in a head of the relative clause:

(102)  

(a) *I gave him, every report on Bob,’s division) you did <give him,>.  
(b) *I reported her, to every cop in Abby,’s neighborhood you did <report to her,>.  
(c) *I showed her, every picture from Abby,’s mantelpiece you did <show her,>.

(Merchant 2000:569)
(103)  a. In the end, I did ask him, to teach the book of Irene’s that David, wanted me to <ask him, to teach>.
   b. *In the end, I did ask him, to teach the book of David,’s that Irene wanted me to <ask him, to teach>.

(Sauerland 1998:38-9)

The contrast between (101), and (102) and (103) is explained by Fox’s Late Merge analysis of ACD in which a lexical content of a relative clause head is represented inside the c-command domain of the relevant pronoun, but not within the relative clause. In contrast, the Wholesale Late Merge analysis cannot account for the Condition C violation in (102) and (103)b. The source of the problem here is the same as we have seen in Sauerland’s Generalization. The Wholesale Late Merge approach allows movement of the determiner in the matrix clause. Together with the possibility of analyzing the relative clause as a raising structure, this in turn allows us not to represent both the head noun and the relative clause in the matrix object position, as illustrated in (104)a. Consequently, the resulting LF representation does not violate Condition C. It also satisfies the parallelism condition because of null operator movement within the relative clause, and thus incorrectly predicts that (103)b is grammatical:

(104)  a. the matrix clause
       [I gave him [every]]
       \[ movement of the determiner
       [[[I gave him [every]] [every]]
       \[ Trace Conversion
       [[[\lambda x. [I gave him [the x]]] [every]]

b. the relative clause
   [you did <give him OP>]
   \[ null operator movement
   [OP, [you did <give him OP,>]]
   \[ Trace Conversion
   [\lambda y. [you did <give him [the y]>]]
If the right characterization of the problem above is that movement is allowed to leave a contentless copy too widely, there could be two possible remedies. One remedy would be to constrain the applicability of null operator movement so that it cannot be applied in (100)b and (104)b, but so it can be applied in the acceptable Wold’s case in (95). I can think of no straightforward way to constrain the applicability of null operator movement this way and, hence, I will not pursue this possibility here. The other possible approach would be to restrict a raising structure of relative clauses. If the relative clauses in (98)a, (102), and (103)b could not be analyzed as having a raising structure, or were forced to involve a matching structure, Wholesale Late Merge would not be feasible, because the restrictor constituent of the determiner would be an NP in this case. Consequently, the lexical content effect would emerge. Therefore, the ACD in (98)a would be ruled out by the failure to satisfy the parallelism condition, as in Fox (2002), regardless of whether null operator movement takes place within the relative clause. The ACD in (102) and (103)b would be excluded by a violation of Condition C and also by the failure to satisfy the parallelism condition, when the relative clause involves null operator movement. I here suggest one possible condition on relative clauses, which makes a raising structure available in cases like Wold’s ACD in (95), but unavailable in cases such as (98)a, (102), and (103)b. The difference between the two cases is that a raising structure and a matching structure produce different meanings only in Wold’s ACD case, due to the presence of the variable in the head noun. Capitalizing on this aspect, I postulate the following economy condition:

(105) **Relative Clauses Economy**

A relative clause can involve a raising structure only if the semantic interpretation derived from a raising structure for the relative clause cannot be derived from a matching structure for it.

Given the condition in (105), a matching structure is forced in (98)a, (102), and (103)b, and the lexical content effect correctly comes into force.

Given the economy condition in (105), the suggested approach to Wold’s ACD puzzle makes a prediction, which appears not to be borne out. That is, we predict that the
lexical content effect should disappear if a relative clause is allowed to involve a raising structure, for example, by adding a bound variable into a head noun of the relevant relative clause. The sentences in (106)b and (107)b are relevant cases, and due to the presence of a bound variable in the head noun, the Wholesale Late Merge approach predicts that they would be acceptable, in contrast to (106)a and (107)a where no such variable is involved. However, there is no contrast between the two cases, as shown below:

(106)  a. *I need to read an article of Mary’s that the professor said summarized the book that I did <read>.
    b. *I need to read the article of his that every professor said summarized the book that I did <read>.

(107)  a. *I want him to read the article of Mary Smith’s about Bush that every reporter said that he didn’t <read>.
    b. *I want him to read the article of hers about Bush that every reporter said that he didn’t <read>.

I cannot provide any way to solve this problem that is consistent with the Wholesale Late Merge approach to Wold’s ACD puzzle. I leave this issue to future research. However, I would like to suggest that the Wholesale Late Merge approach could be taken as the first step towards solving Wold’s puzzle.
3.6 A Note on Case

The discussion about Case in the previous sections deserves further clarifications. Here, I suggest that the Case and agreement system developed in Pesetsky and Torrego (2001, 2004, 2005, to appear) fits fairly well with the proposals in this chapter.38

Let me first introduce two components of Pesetsky and Torrego’s system that are important for our purposes. The first component is their view of Case and the syntactic operation Agree. Following Frampton and Gutmann (2000) and Frampton et al. (2000), among others, Pesetsky and Torrego (to appear) argue that Agree should be viewed as feature sharing, as defined in (108).

(108) Agree (Feature sharing version)

(i) An unvalued feature F (a probe) on a head H at syntactic location α (Fα) scans its c-command domain for another instance of F (a goal) at location β (Fβ) with which to agree.

(ii) Replace Fα with Fβ, so that the same feature is present in both locations.

(Pesetsky and Torrego to appear)

One consequence of this view of Agree is that an unvalued feature as a probe can enter into an Agree relation with another unvalued feature as a goal, and as a result of this Agree relation, a single feature is shared between the probe and the goal, as illustrated in (109). (Following Pesetsky and Torrego’s notations, F[ ] represents an unvalued feature that has not entered into an Agree relation. Indices are employed to represent a single feature shared between two positions.) I suggest below that an Agree relation of this sort occurs between determiners and their restrictor NPs.

38 I am grateful to David Pesetsky for pointing out the relevance of Pesetsky and Torrego’s system for the proposals in this chapter.
The second component of Pesetsky and Torrego’s framework, which I will make essential use of, is their idea that structural Cases (i.e., nominative and accusative) are the incarnations of uninterpretable unvalued tense features on determiners, and uninterpretable unvalued tense features need to enter into an agreement relation with a valued tense feature on a head. This component is crucial for explaining the above adopted assumption that an NP as the restrictor constituent of a determiner must be introduced inside the c-command domain of its Case assigner, but CPs do not have this requirement.

Now, let me discuss the assumptions about Case that I have adopted within Pesetsky and Torrego’s framework. On the basis of the fact that case morphology can be found in nouns as well as in determiners in some languages (e.g., Latin), I suggest that both nouns and determiners bear an uninterpretable unvalued tense feature, and that an uninterpretable unvalued tense feature on a probing determiner enters into an Agree relation with a corresponding feature on an NP in the way depicted in (109). When a head with a valued tense feature is introduced into the derivation, as a result of valuation a single valued tense feature results in being present in three positions in the structure, as illustrated in (110)a. In other word, both the determiner and the NP successfully bear a valued tense feature. (Again, following Pesetsky and Torrego’s notation, Fval[] represents a valued feature.) Suppose that an NP is introduced into the structure outside of the c-command domain of a head that possesses a valued tense feature, as in (110)b. The unvalued tense feature on the NP would not be valued by the valued tense feature on the head X in (110)b because the NP and the head X are not in an appropriate c-command relation. I suggest that the NP cannot receive a value from the determiner, either. There would be two possible reasons for this. One would be that the valued tense feature on the determiner has already been deleted when the NP is introduced. The other would be that, since the tense feature on the determiner has already been valued, it cannot enter into an Agree relation with the unvalued tense feature on the NP. For concreteness, I assume the
former in (110)b. Consequently, the uninterpretable unvalued tense feature on the NP remains as it is in this derivation, which makes the derivation illegitimate.

One question would be whether the unvalued tense feature on the NP in (110)b could be valued by a valued tense feature on a head that enters into the structure above ZP. I conjecture that it could not be valued by the valued feature on that head because the determiner intervenes between the two relevant heads. This could be taken as an instance of defective intervention. However, if there is no intervening determiner, the unvalued tense feature on the NP would be able to be valued by the valued tensed feature on the higher head in this configuration. I suggest that this is what happens within a relative clause in Wold's ACD case in (111)a. In order to derive the structure that satisfies the parallelism condition, I have suggested that null operator movement takes place in the most embedded clause, and the NP picture of himself is introduced in a position above the moved null operator, as depicted in (111)b. Thus, the unvalued tense feature on the NP cannot be valued by a valued feature on a head inside the most embedded clause. Capitalizing on the fact that the NP does not have a determiner as a sister in this construction, I suggest that its unvalued tense feature is valued by a feature on a head in the higher clause within the relative clause, as shown in (111)b.

(111) a. Sue likes every picture of himself, that every boy hoped that she would <like>.

b. [H\textsubscript{Fval} ... [CP NP\textsubscript{f5} [CP OP 1 [TP she would like OP 1]]]]

where the NP is picture of himself.
To sum up the discussion so far, an NP cannot be introduced outside of the c-command domain of the head responsible for its Case assignment, if the NP ends up being sister to a determiner.

Pesetsky and Torrego’s framework also helps in explaining why CPs are different from NPs with respect to the applicability of counter-cyclic merger. Pesetsky and Torrego argue that, although a complementizer also bears an uninterpretable unvalued tense feature, it receives a value from a head within a finite TP in its complement. Pesetsky and Torrego claim that this accounts for the distributional differences between CPs and DPs, which have been taken as evidence in favor of the claim that only DPs need Case. Particularly relevant for our purposes is that, since valuation of a tense feature on a complementizer has nothing to do with the structure outside of a CP, CPs can be merged with determiners counter-cyclically in the context in which NPs cannot. We have seen that this plays an essential role in the account of Wold’s ACD puzzle.

3.7 Conclusion

In this chapter, I have developed a theory of movement and the merger operation, which explains the otherwise puzzling properties of A-movement and A'-movement: A-movement optionally leaves a contentless copy, and A'-movement obligatorily leaves a copy. Some of the facts that illustrate this property of A-movement have been taken as challenges to the copy theory of movement. However, I have argued that the Wholesale Late Merge approach accounts for those facts within the simplest version of the copy theory of movement, in which movement obligatorily leaves a copy. I have also suggested that Wholesale Late Merge is supported by the facts of ellipsis. If the discussion in this chapter is successful, we are able to maintain the copy theory of movement and the copy theory approach to the effects of Scope Narrowing.

---

39 Determiners bear an uninterpretable unvalued tense feature. Therefore, a raising structure of relative clauses cannot be base-generated in a non-Case position, even though the restrictor constituent of determiners does not need Case in a relative clause with this structure.
Chapter 4

More Than Two Quantifiers

4.1 Introduction

4.1.1 The Goal

Chapter 4 discusses the scopal behavior of so-called comparative quantifiers, like *more than three books*. These quantifiers exhibit limited scopal possibilities, a fact which suggests that Scope Shifting Operations (SSOs) should be constrained. In this chapter, I argue for various constraints on SSOs. I also discuss various ramifications for the nature of the scope taking mechanism and of obligatory and optional operations. Specifically, I claim that SSOs are successive cyclic in the sense that each step of an SSO must be independently licensed by grammar.

4.1.2 The Proposal in a Nutshell

According to a widespread view, two quantificational phrases (QPs) can interact freely in a single clause (May 1985). This is based on representative examples like (1), in which an existential and a universal QP are in subject and in object position, respectively.

This chapter is a slightly modified version of Takahashi (to appear).
(1) Some student read every book.

\((\exists > \forall) (\forall > \exists)\)

This generalization, however, does not hold up, to close scrutiny pointed out by Beghelli (1995), Beghelli and Stowell (1997), Liu (1990, 1997), and Szabolcsi (1997), among others. For instance, if the universal QP in (1) is replaced with the quantificational DP more than five books, which I refer to as a comparative QP (or CQP), following Hackl (2000), wide scope of the comparative QP is not available:\(^1\)

(2) Some student read more than five books.

\((\exists > \text{CQP}) * (\text{CQP} > \exists)\)

\[(\text{Beghelli 1995:48})\]

At first sight, this scope fact is unexpected. On the assumption that comparative QPs are, like universal QPs, unanalyzable generalized quantifiers, there is no obvious way to prohibit comparative QPs (and only comparative QPs) from outscoping the subject QP, short of stipulation.

In this chapter, I demonstrate that the scopal contrast between (1) and (2) is a natural consequence of general principles of grammar, once we acknowledge that comparative QPs are structurally more complex than other QPs. Specifically, I pursue a decompositional view of comparative QPs, originally proposed by Bresnan (1973) and defended and elaborated by Hackl (2000), among others. I argue that comparative QPs are actually composed of two generalized quantifiers, namely, the comparative operator and the generalized quantifier many NP (e.g., er than five and many books, respectively, in (2)). Wide scope of an object comparative QP thus requires the subject QP to take scope below both the QPs composing the comparative QP. I argue that a locality condition on SSOs prevents the subject QP from taking scope below both the comparative operator and many NP in one fell swoop, forcing an intermediate step in which the subject QP takes scope between these two generalized quantifiers. However, I

---

\(^1\) As we will observe in subsequent sections, comparative QPs show scopal interaction with other QPs in certain cases. In this respect, they behave similarly to other QPs.
demonstrate that this intermediate step is always ruled out by independently motivated principles, thus preventing wide scope readings of the object comparative QP. The argument here can be taken as an argument for a derivational theory of SSOs. In other words, an SSO proceeds in steps, each of which must be independently licensed by principles of grammar. In the end, it turns out that the apparent peculiarity of comparative QPs provides additional evidence for the core properties of SSOs.

This chapter is organized as follows. In section 4.2, I present the scope facts exhibited by comparative QPs, which constitute the primary empirical concern of the paper. Section 4.3 presents the three components of my analysis. First, I present a decompositional approach to comparative QPs and some arguments in favor of it. Second, I introduce a locality constraint on SSOs that forces an intermediate step in the derivation of wide scope of comparative QPs. Finally, I demonstrate that this intermediate step produces a structure that violates an independent constraint on Degree Phrase (DegP) movement proposed in Heim (2001a), which I call the Heim-Kennedy Constraint. Section 4.4 provides two cases in which a comparative QP actually can take scope over other QPs. I demonstrate that the proposed analysis can distinguish these cases from the cases in which an object comparative QP must take narrow scope relative to a QP in subject position. In section 4.5, I discuss implications of the proposed analysis for the architecture of grammar. It is revealed that obligatory operations are required to apply before optional ones. I argue that this requirement follows from the fundamental nature of obligatory and optional operations. I also explore the derivational character of SSOs in more detail. In section 4.6, I propose a more principled account of the scopal behavior of comparative QPs which is based on an independently motivated constraint on SSOs and a semantic constraint on the comparative operator. In section 4.7, I furthermore present evidence in favor of the constraints proposed in section 4.6 and against the Heim-Kennedy Constraint. All of these facts provide empirical support for the proposed approach, which does not depend on a restriction on DegP movement. Section 4.8 discusses further issues that are pertinent to the nature of the Heim-Kennedy Constraint and SSOs. First, I examine Heim’s (2001a) original facts motivating the Heim-Kennedy Constraint within the proposed framework. While one of Heim’s facts will require further investigation, I suggest that the proposed approach should be taken as
a first step towards understanding the nature of the constraint. Second, I discuss an alternative approach to the Heim-Kennedy Constraint proposed by Stateva (2002, 2004), and suggest some potential problems that arise in this approach. Finally, I take up one peculiar property of QR, which is revealed by looking at the scopal interaction between QPs and quantification bearing elements that are not DPs.

4.2 The Comparative QP Scope Puzzle

The scope puzzle introduced above can be restated as follows. If comparative QPs are in object position, they must take narrow scope relative to the subject QP, in striking contrast to universal QPs. This point is illustrated in (3).²

(3)  a. One student in this class read more than three books.
    (∃>CQP) *(CQP>∃)

   b. Some student in this class read more than three books.
    (∃>CQP) *(CQP>∃)

Thus, the sentences in (3) are judged false if there is no single student who read more than three books even if there are more than three books that were each read by one student.

This scope freezing effect is robust across types of QPs in subject position. Let me first consider the case with a universal QP in (4). The comparative QP is obligatorily assigned narrow scope in this case, too.³

² It has been observed that existential QPs (e.g., some NP) and bare numeral QPs in subject position show the same scopal property, namely, they must take scope over comparative QPs in object position (Beghelli and Stowell 1997, among others).

³ One qualification is in order here. In (4), the inverse scope reading entails the surface scope reading. In other words, no situation makes the inverse scope reading true, while making the surface scope reading false. However, the entailment pattern is reversed in downward entailing contexts (e.g., negative contexts), and there the entailment problem disappears.
(4) Every student in this class read more than three books.

\((\forall > CQP) \ast (CQP > \forall)\)

The sentences in (3) and (4) show that monotone increasing QPs in subject position must outscope monotone increasing comparative QPs in object position.\(^4\) In addition, monotone decreasing comparative QPs such as fewer than three books must also take narrow scope relative to monotone increasing QPs, as shown in (5) and (6).\(^5\)

(5) a. One student in this class read fewer than three books.

\((\exists > CQP) \ast (CQP > \exists)\)

b. Some student in this class read fewer than three books.

\((\exists > CQP) \ast (CQP > \exists)\)

(6) Every student in this class read fewer than three books.

\((\forall > CQP) \ast (CQP > \forall)\)

The scope rigidity above is also observed in cases where the subject is a monotone decreasing QP. This is illustrated in (7).

(7) No student in this class read more than three books.

\((\text{no} > CQP) \ast (CQP > \text{no})\)

(i) It is not the case that every student in this class read more than three books.

\((\forall > CQP) \ast (CQP > \forall)\)

In (i), the surface scope reading entails the inverse scope reading and we can find situations that make the surface scope reading false and the inverse scope reading true. For instance, suppose that for each one of the students in this class, s/he read ten books, but there is just one book that was read by every student. If wide scope of the comparative QP relative to the universal QP were available in (i), it would be judged true in this context. However, this is not the case, which indicates that the comparative QP cannot take scope over the universal QP in (4).

\(^4\) A quantifier Q is monotone increasing if \(X \in Q\) and \(X \subseteq Y \subseteq D\) implies \(Y \subseteq Q\). See Barwise and Cooper (1981).

\(^5\) A quantifier Q is monotone decreasing if \(X \in Q\) and \(Y \subseteq X \subseteq D\) implies \(Y \subseteq Q\). See Barwise and Cooper (1981).
Finally, the scope freezing also holds in relation to non-monotonic QPs in subject position:

(8) Exactly two students in this class read more than three books.

\[(\text{exactly } 2 > \text{CQP}) \ast (\text{CQP} > \text{exactly } 2)\]

The scope facts above lead us to draw the following descriptive generalization:

(9) The Comparative QP Scope Generalization

Comparative QPs in object position must take narrow scope relative to QPs in subject position.

With respect to their scope taking properties, comparative QPs are strikingly different from universal QPs. This might tempt one to pursue the hypothesis that comparative QPs cannot undergo QR at all and are interpretable in situ. However, this hypothesis must immediately be rejected since comparative QPs can host ACD, just like universal QPs, as shown in (10).6

(10) John speaks more than three of the languages that Mary does <speak>.

The fact that ellipsis of EC is licensed in (10) indicates that QR is available even for comparative QPs.

The puzzle is that comparative QPs in object position must take narrowest possible scope even if they can undergo QR. In section 4.3, I present the outline of my account of this puzzle.

---

6 The ACD in (10) is degraded if the comparative QP is not partitive. I leave it to future research to elucidate why this is the case. See Hackl (2000) for the same observation and for relevant discussion.
4.3 The Plot

4.3.1 An Outline

In this section, I provide an explanation for the Comparative QP Scope Generalization. The explanation is composed of three components, all of which are independently motivated. The subsequent three subsections are each devoted to the introduction of one of these components. Finally, I demonstrate how the combination of the three components accounts for the Comparative QP Scope Generalization.

4.3.2 The First Component: A Decompositional Approach to Comparative QPs

As I have already mentioned, there is no straightforward way to explain the scope puzzle as long as we assume that comparative QPs are unanalyzable generalized quantifiers like universal QPs. However, we have good reason to believe that comparative QPs involve a more complex internal structure than universal QPs. More concretely, the comparative QP *more than three books* consists of a comparative operator *er than three* and the DP *many books*, both of which I assume are generalized quantifiers, following Hackl (2000). The original insight into the decompositional approach to comparatives and comparative QPs goes back to Bresnan (1973) and is defended by Bhatt and Pancheva (2004, 2005), Hackl (2000), Heim (2001b), and Wold (1992), among others.

I assume, following Heim (2001a), that the comparative operator *er* is a head of a DegP and, based on the decompositional approach to comparative QPs proposed in Hackl (2000), that it is base-generated as sister of the quantificational determiner *many*, as shown in (11).7

---

7 The structure in (11) does not represent the surface order, namely, *more than three books*. I assume that the structure in (11) is spelled-out in the observed order at the PF level. See Bhatt and Pancheva (2004, 2005) for relevant discussion.
If a comparative QP is the sister of a predicate that is of type \( \langle e, \langle e, t \rangle \rangle \), like in (12)a, neither QP is interpretable in situ, due to type mismatch (see (30) and (50) for the lexical entries of the comparative operator and the determiner \textit{many}, respectively). Both must undergo QR for interpretability, adjoining to a node of type \( t \). Given this, (12)a is parsed as the LF representation in (12)b under the assumption that QR targets \( vP \), which is of type \( t \) (see 4.3.3 for further discussion of this assumption).\(^8\)

(12)  
\begin{enumerate}
  \item John read more than three books.
  \item \([TP \text{John}_1 [vP [DegP er than three]_2 [vP [DP t_2 many books]_3 [vP t_1 \text{read } t_3]]]]\)
\end{enumerate}

One of the most compelling arguments for the decompositional view comes from the scope splitting phenomena discussed by Hackl (2000), Heim (2001a, b) and Wold (1992). Namely, the comparative operator and the DP (e.g., \textit{many books} in (12)b) take scope at distinct positions. Here, I provide Wold’s evidence for the scope splitting phenomenon, which is based on an ACD puzzle originally raised by Diesing (1992) (see 4.6.3 for discussion of another instance of scope splitting, observed by Hackl 2000). To set the stage, let us observe (13).

(13)  
\begin{enumerate}
  \item John read many books.
    \begin{enumerate}
      \item (existential) (presuppositional)
      \item John read many books that Sue did <read>.
      \begin{itemize}
        \item *(existential) (presuppositional)
      \end{itemize}
    \end{enumerate}
\end{enumerate}

The sentence in (13)a is ambiguous in that the object DP can receive either an existential or a presuppositional reading. In the existential reading, the sentence in (13)a asserts the existence of books. The presuppositional reading, on the other hand, presupposes the existence of books. If there were no book, the existential reading would be judged false, \(^8\) In order to avoid a complication that is not relevant for our current concern, I assume here that movement leaves a trace. See Bhatt and Pancheva (2004, 2005) and fn. 13 for relevant discussion.

135
whereas the presuppositional reading would receive no truth value because of presupposition failure. Notice that if the object DP hosts ACD, the only available reading is the presuppositional reading, as illustrated in (13)b (see Diesing 1992 for relevant discussion). As mentioned above, ACD requires quantified expressions to undergo QR out of the VP. It follows from the facts in (13) that QR out of the VP forces inherently ambiguous DPs to be construed as presuppositional.

However, a puzzle arises once we consider the contrast between the following sentences:

(14)  a. *There will be almost no one that there should <be> at the party.
       b. There will be almost no one that there should be at the party.

(Pesetsky 2000:12-13)

(15) John expected there to be more students than Mary did at the party.

(Wold 1992:2)

The ungrammaticality of (14)a can easily be explained by what we have already said. In the existential _there_-construction, post-copular nominals cannot receive a presuppositional reading (Milsark 1974), indicating that they cannot undergo QR. However, ACD requires QR of the post-copular nominal that contains an elided VP in order to generate a structure that satisfies the parallelism condition. In (14)a, the two conflicting requirements cannot be satisfied at the same time, resulting in ungrammaticality. The puzzle is that if we assume that the comparative expression _more students than Mary did_ is an unanalyzable constituent, like the post-copular nominal in (15), we would predict (15) to be ungrammatical on a par with (14)a, contrary to fact. Wold’s solution to this puzzle is that only the comparative operator undergoes QR to the matrix VP, piped-piping the comparative complement introduced by _than_. Thus, _many students_ can receive an existential reading, which is consistent with the requirement of the existential _there_-construction. Furthermore, the LF representation of (15), which is illustrated in (16), satisfies the parallelism condition. Following Chomsky (1977), I

136
assume that there is null operator movement within the comparative complement (cf., Lechner 1999):

(16) \[
[T_P \text{John}_1 [v_P [\text{DegP} \text{ er than OP}_2 \text{Mary}_3 \text{ did}]
\text{[t, expected there to be [DP t many students]]}_4,
\text{[t, expected there to be [DP t many students]]}_4 core]}
\]

In conclusion, (15) is an instance of the scope splitting phenomena since the comparative operator and many students take scope at distinct positions. Therefore, the ACD fact in (15) is an argument in favor of the decompositional approach.

Bearing the decompositional view in mind, let us go back to our main empirical concerns. In the decompositional approach, there are three generalized quantifiers in (17)a, as the LF representation in (17)b shows.

(17) a. One student read more than three books.
   b. \[
   [T_P \text{one student}_1 [v_P [\text{DegP} \text{ er than three}_2 [v_P [\text{DP t many books}_3 [v_P \text{t, read}\text{t}_3]]]]}
   \]

Since there are three generalized quantifiers in the structure, we expect six scopal relations to be potentially available. Notice, however, that the comparative operator undergoes QR from the quantified DP many books and the quantified DP many books contains a trace left behind by the comparative operator. Thus, the quantified DP can never outscope the comparative operator; otherwise, the variable left behind by the comparative operator would be unbound, and therefore, uninterpretable. Consequently, three out of six scopal relations are independently ruled out. The remaining three scopal relations in (17) that we should investigate are illustrated in (18).
In (18)a, the subject QP takes wide scope over the two decomposed QPs, or the comparative QP, which yields a surface scope reading. The representation in (18)c in which the subject QP is outscoped by the two decomposed QPs delivers a so-called inverse scope reading. Let me put aside for the moment the question of what interpretation is derived from the intermediate scope representation in (18)b, which is the main topic of section 4.6. We can now restate the scope puzzle in a new fashion. The puzzle is that the LF representation in (18)c is not available for cases like (17)a.

4.3.3 The Second Component: A Locality Condition on SSOs

One of the strong arguments for a movement approach to scope assignment comes from the fact that scope assignment is governed by island constraints on movement (see Rodman 1976, among others, and see also Chapter 1). However, it has been claimed that SSOs are more severely restricted than other movement operations in terms of locality. Fox (2000) argues this point on the basis of scope facts in VP-ellipsis. Specifically, Fox argues that SSOs must target the closest node of type t.

Let us first consider the following contrast with respect to the availability of wide scope of the object QPs in VP-ellipsis (see Hirschbühler 1982, Sag 1976 and Williams 1977 for earlier observations and discussion, and see also 2.5.1 for relevant discussion).
The only difference between (19)a and (19)b is that the subject in the ellipsis clause is a referential expression in (19)a and it is a QP in (19)b.

(19)  a. Some boy admires every teacher. Mary does <admire every teacher>, too.
     (∃∀) *(∀∃)

     b. Some boy admires every teacher. Some girl does <admire every teacher>, too.
     (∃∀) (∀∃)

(Fox 2000:4)

It is likely that the parallelism condition on ellipsis has the consequence that QPs take parallel scope in the antecedent clause and the ellipsis clause.\(^10\) In (19)b, the non-parallel scopal relations are excluded by the parallelism condition and hence, it is just two-way ambiguous. As discussed above, if a QP is not sister to a predicate that is of type (e, t), it must undergo QR for interpretability. Let us first hypothesize that QR can target any node of type t, and, more specifically, that QR can move the object QP to the TP-adjoined position in which it is interpretable. This hypothesis overgenerates wide scope of the universal QP in (19)a. The entire TP in (20)b can be taken as a Parallelism Domain because it is semantically identical to AC in (20)a, modulo the focus marked subject Mary.

(20)  a. Antecedent Clause: [TP every teacher\(_2\) [TP some boy\(_1\) [TP t\(_1\) admires t\(_2\)]]]

    b. Ellipsis Clause: [TP every teacher\(_4\) [TP Mary\(_3\) <TP t\(_3\) admires t\(_4\)>]]

However, if this type of long distance QR is not an option, we have a way to avoid this problem, as we will see shortly. Fox (2000) formulates the Shortest Move constraint, given in (21) (see Bruening 2001 and Sauerland 2000b for relevant discussion).

\(^10\) Fox (1999b, 2000) proposes a more elaborated version of the parallelism condition. In this system, two QPs sometimes do not take their scope at parallel positions in two clauses. I am assuming the simplified version of the parallelism condition on ellipsis for expository purposes.
Since the closest node of type t from object QPs is the vP node, the first possible landing site for obligatory QR is the vP-adjoined position, according to Shortest Move. The next step is to apply an optional SSO to derive object wide scope. For purposes of concreteness, I assume Johnson and Tomioka’s (1997) approach to object wide scope. Johnson and Tomioka argue that it is derived by movement of the subject QP to the position below the QRed object. I will regard this movement as Quantifier Lowering (QL), which is an optional SSO, in that it is not driven by type mismatch. I would like to suggest that optional SSOs are also constrained by a locality condition slightly stronger than (21):

11 Fox (2000) assumes that optional QR is also available for wide scope of the object.
12 This particular execution of scope taking is not crucial for the proposals that follow. The analysis pursued below can be straightforwardly restated, assuming optional QR approaches to object wide scope.
13 In Chapter 3, I have argued for the copy theory of movement and suggested that this view of movement provides a straightforward way to capture the effects of Scope Narrowing. In this chapter, I instead assume that they are produced by the movement operation, primarily because the constraints that I will argue for below have been formulated as the constraints on movement in the past literature. However, it is certainly possible to reformulate those constraints within the framework of the copy theory approach to Scope Narrowing. Within the copy theory of movement framework, there is one complication caused by a property of the comparative operator. As discussed in Bhatt and Pancheva (2004, 2005) and Heim (2002), the comparative operator is not a conservative determiner. Thus, if Trace Conversion applies to a copy that contains both the comparative operator and its restrictor constituent, the resulting structure produces an incorrect interpretation. Bhatt and Pancheva (2004, 2005) propose a solution to this problem. See fn. 34 for relevant discussion.
14 The formulation in (22) implies that QLed elements move to a position distinct from one of their traces. Some evidence for this claim comes from VP-ellipsis inside the subject QP:

(i) 
[Every student who wanted to <come to the party>] didn’t come to the party.
For example, although both John and Mary wanted to come to the party, John did and Mary didn’t.

As is clarified by the continuation, narrow scope of the universal QP relative to negation is available for the first sentence in (i). However, this scopal relation seems not to be an outcome of Scope Narrowing of the subject QP into its original position inside the vP, as shown in (ii).

(ii) 
[Tp didn’t [vP [every student who wanted to [vP e]] come to the party]]
(22) **Shortest Move (optional SSOs)**

A QP may undergo an optional SSO only if that SSO targets the closest node of type t and crosses a quantificational element.

Given these assumptions, the subject QP must target the vP node sister of the QRed objects, and object wide scope in (19)a is yielded by the LF representations in (23).

(23) Antecedent Clause: \[ \text{TP } [\text{vP every teacher}_1 [\text{vP some boy}_2 [\text{vP t}_2 \text{ admires t}_1]]] \]

Ellipsis Clause: \[ \text{TP } [\text{vP every teacher}_3 [\text{vP Mary}_4 \text{ t}_4 \text{ admires t}_3]]] \]

In this situation, we can now ask the question whether the optional SSO for deriving object wide scope is allowed in (19)a. Fox’s (2000) answer is that QL in the ellipsis clause in (23) is prohibited by **Scope Economy**. The basic idea of Scope Economy is given in (24); its formal definition will be provided in section 4.6.2.

(24) **Scope Economy**

SSOs cannot be semantically vacuous.

(Fox 2000:3)

The subject in (19)a is a referential expression. Since QL of the subject in the ellipsis clause has no semantic effect and is therefore not allowed, the parallelism condition requires the subject QP in the antecedent clause to stay in the Spec of TP. In contrast, the

In this representation, the antecedent vP contains the elided vP and the parallelism condition would not be satisfied, assuming that deletion targets vP, instead of VP. Instead, the LF representation of (i) where the parallelism condition is met must be like the one in (iii).

\[ \text{TP } \text{ didn’t } [\text{vP every student who wanted to } [\text{vP e}]] [\text{vP t}_1 \text{ come to the party}]] \]

It is crucial to notice that in (iii), the subject QP receives its interpretation in a position where it never did stop on the way to the Spec of TP, under the widespread view of subject raising, namely, that the subject directly moves from the Spec of vP to the Spec of TP. This fact can be explained in the QL approach. However, if the subject lands in more places than assumed here, the fact in (i) could be accounted for without making reference to QL (see Johnson and Tomioka 1997 and Sauerland 2003a for relevant observation and discussion).
subject in the ellipsis clause in (19)b is quantificational, and QL has a semantic effect. Thus, QL of the subject is allowed in (19)b, which makes it ambiguous.

Let me now return to the scopal relations in cases like (25).

(25) One student in this class read more than three books.

First, the decompositional approach allows us to take the three scopal relations in (26) into account.

(26) a. widest scope b. intermediate scope c. narrowest scope

Second, Shortest Move forces the subject QP to first target the closest node of type t, which is the sister of the comparative operator. Thus, the combination of the two components ensures that we cannot directly produce (26)c from (26)a without (26)b being established. The next question that I would like to ask is whether the intermediate scope representation in (26)b is legitimate. In the next subsection, I demonstrate that an independently motivated constraint on DegP movement prohibits this representation.

4.3.4 The Third Component: An Intervention Constraint

In this subsection, I introduce an intervention constraint on movement of the DegP, or the comparative operator. This constraint is based on a descriptive generalization proposed by Heim (2001a), which incorporates Kennedy’s (1999) observation on the scopal interaction between the comparative operator and QPs. I will demonstrate that this independently motivated constraint rules out the intermediate scope representation in (26)b. Before going into this, let me briefly introduce the semantics of comparatives developed in Heim (2001a, b) (see also Cresswell 1976, Hellan 1981, Kennedy 1999,

It is assumed that gradable adjectives, such as *tall*, denote functions, which are of type \( \langle d, \langle e, t \rangle \rangle \). In addition, I assume that these functions are monotone. The lexical entry of the gradable adjective *tall* and the definition of monotonicity are given in (27) and (28), respectively.

\[
\text{(27)} \quad [\text{tall}] = \lambda d \in D_d. [\lambda x \in D_e. x \text{ is tall to degree } d]
\]

(Heim 2001a:216)

\[
\text{(28)} \quad \text{Monotonicity}
\]

A function \( f \) of type \( \langle d, \langle e, t \rangle \rangle \) is monotone iff

\[
\forall x \forall d \forall d' [f(d)(x) = 1 \text{ and } d' < d \rightarrow f(d')(x) = 1]
\]

(Heim 2001a:216)

If the first argument of gradable adjectives is occupied by a referential expression of type \( d \) (e.g., *six feet* in (29)a), it remains in situ, as illustrated in the LF representation in (29)b, and the structure is straightforwardly assigned the correct meaning in (29)c.

\[
\text{(29)}
\]

a. John is six feet tall.

b. \([TP \text{ John}_1 \text{ is } [AP t_1 [A^* [\text{DegP six feet} \text{ tall}]]]]

c. John is tall to the degree six feet.

Because of the monotonicity of the gradable adjective, if John is six feet tall, he is also five feet tall, four feet tall, and so on. The LF representation of comparatives is more involved because the comparative operator and the comparative complement (i.e., the DegP) constitute a generalized quantifier over degrees, and the DegP is not interpretable in its original position. Following Heim (2001b), I assume that the comparative operator denotes a proper subset relation between two sets of degrees. Its lexical entry is given in (30).
In order to resolve type mismatch, the DegP must undergo QR to the AP adjoined position, which is the closest node of type t, and leave behind a trace of type d. Thus, the comparative construction in (31)a is analyzed as having the LF representation in (31)b, and it can be compositionally interpreted, which derives the correct interpretation in (31)c.\(^\text{15}\)

\[(31)\]

a. John is taller than Bill is.

b. \([\text{TP John is [AP [DegP \text{er than OP} 2 Bill is [t\text{tall}}] 3 [AP t_1 t_3 \text{tall}]])}\]

c. \{d: Bill is d-tall\} \subset \{d: John is d-tall\}

(i.e., John’s maximal height is higher than Bill’s maximal height.)

The treatment of the comparative operator as a quantificational determiner leads us to expect that it could scopally interact with other QPs (however, see Heim 2001a for an important caveat). Let us consider the less-comparative in (32)a. As shown in (32)b and (32)c, there are potentially two different LF representations, depending upon whether an optional SSO is applied. The resulting interpretations are given below the LF representations:\(^\text{16}\)

\[(32)\]

a. John is less than Bill.

b. \([\text{TP John is [AP [DegP \text{er than OP} 2 Bill is [t\text{tall}}] 3 [AP t_1 t_3 \text{tall}]])}\]

c. \{d: Bill is d-tall\} \subset \{d: John is d-tall\}

(i.e., John’s maximal height is higher than Bill’s maximal height.)

\(^{15}\) In so-called clausal comparatives, such as (31)a, it is assumed that ellipsis of some constituent takes place in the comparative complement.

\(^{16}\) Following Heim’s (2001a) basic idea, I tentatively assume that the lexical entry of the less-comparative operator is the same as (30), except that the direction of the ordering relation is the opposite, as illustrated in (i).

\[(i)\] \[\| \text{less} \| := \lambda D \in D_{(d, \cdot)}\cdot \lambda D' \in D_{(d, \cdot)}\cdot D \supseteq D'\]

However, I will present a decompositional analysis of the less-comparative operator in 4.8.1.
(32)  a. Every girl is less tall than John is.
    b.  (\forall<less)
      [\textit{TP} \text{every girl}_1 \text{ is } [\textit{AP} \text{[DegP less than John is]_2 } [\textit{AP} \text{t}_1 \text{ t}_2 \text{ tall}]]]
      \forall x[x \text{ is a girl } \rightarrow \{d: \text{John is d-tall}\} \supset \{d: x \text{ is d-tall}\}]
    c.  (\text{less<V})
      [\textit{TP} \text{____ is } [\textit{AP} \text{[DegP less than John is]_2 } [\textit{AP} \text{every girl}_1 [\textit{AP} \text{t}_1 \text{ t}_2 \text{ tall}]]]]
      \{d: \text{John is d-tall}\} \supset \{d: \forall x[x \text{ is a girl } \rightarrow x \text{ is d-tall}]\}
      (\text{i.e., The shortest girl is shorter than John.})

The interpretation produced by the LF representation in (32)b represents the intuitive meaning of (32)a. In (32)c, the set of degrees on the right-hand side turns out to be identical to the set of degrees to which the shortest girl is tall, because of the monotonicity of the gradable adjective. Thus, if this LF representation is available for (32)a, it should have a reading on which it is true as long as the shortest girl is shorter than John. In other words, (32)a would be true even if some of the girls are taller than John, as long as not all of them are. From the fact that this reading cannot be obtained, it is clear that the LF representation in (32)c is not available for (32)a. On the basis of this evidence and other facts, Heim (2001a) suggests a descriptive generalization that states that the comparative operator cannot outscope any quantificational DP (see also Kennedy 1999 for discussion about scopal properties of the comparative operator). In order to capture this generalization, for the time being, I stipulate an intervention constraint, which I refer to as the Heim-Kennedy Constraint:

(33) \textit{The Heim-Kennedy Constraint}
    A quantificational DP cannot intervene between a DegP and its trace.
    (based on Heim 2001a:223)

The Heim-Kennedy Constraint correctly prohibits wide scope of the comparative operator in (32)c. At the same time, the intermediate scope representation in (26)b, which is repeated here as (34), is ruled out by this constraint.
While we have not considered what interpretation is produced by (34), it will always be excluded by the Heim-Kennedy Constraint.

With this constraint, we have now introduced all the components necessary for the explanation of the Comparative QP Scope Generalization.

4.3.5 An Account of the Comparative QP Scope Generalization

We are now in a position to capture our initial generalization, which is repeated here as (35).

(35) The Comparative QP Scope Generalization

Comparative QPs in object position must take narrow scope relative to QPs in subject position.

First, the decompositional approach to comparative QPs requires us to take the following three scopal relations into account:

(36) a. widest scope    b. intermediate scope    c. narrowest scope

\[
\begin{align*}
\text{QP}_{\text{subj}} & \quad \text{[er than 3]}_1 \\
\text{[er than 3]}_1 & \quad \text{QP}_{\text{subj}} \\
\text{[t}_1 \text{ many books]} & \quad \text{...} \\
\end{align*}
\]

Notice again that narrowest scope of the subject QP produces the inverse scope reading, which is not available for the cases in question. As mentioned above, the subject QP cannot move across the two decomposed QPs in one fell swoop, but must undergo QL into the position between them, due to Shortest Move. However, the intermediate scope representation in (36)b, which is the initial step towards the inverse scope reading, is
ruled out by the Heim-Kennedy Constraint. I would, therefore, like to propose the following derivational character of optional SSOs. Each step of an optional SSO must be licensed by grammar. Each time an optional SSO is applied, its legitimacy needs to be checked right after its application. Even if subsequent SSOs would ameliorate violations of grammatical principles in an intermediate stage of a derivation, their application would not be justified. This derivational character of optional SSOs will be further discussed in section 4.5. In the cases under consideration, since the first application of QL violates the Heim-Kennedy Constraint, no subsequent QL can be applied in order to move the subject QP to the position below the DP many books even though this subsequent QL would avoid violating the Heim-Kennedy Constraint. Therefore, the SSO that would derive wide scope of the comparative QP is not allowed.  

In conclusion, there is no legitimate way to arrive at the narrowest scope representation in (36)c, and, hence, only the surface scope reading is available. This is the explanation for the Comparative QP Scope Generalization.  

4.4 Wide Scope of Comparative QPs

We have so far only discussed cases where comparative QPs in object position take narrow scope relative to QPs in subject position (i.e., cases that fall under the Comparative QP Scope Generalization). However, it is certainly true that comparative QPs can outscope other QPs in some cases. In this section, I will suggest that the proposed approach can distinguish such cases from the cases presented above. I argue that the availability of wide scope for comparative QPs is dependent upon the location in which comparative QPs split into two generalized quantifiers, relative to other QPs in the structure. In other words, it is contingent on structural considerations.

17 In order to account for the Comparative QP Scope Generalization by the copy theory approach to Scope Narrowing, it would be necessary to assume that the copy deletion procedure takes place in a top-down fashion. More specifically, the highest copy of a chain needs to be interpreted in the first stage of the derivation. It can be deleted and the second highest copy can be interpreted only when the representation in this first stage of the derivation is licensed by grammar.

The explanation for the lack of wide scope of comparative QPs in cases like (37) was that intermediate scope of the subject QP, which is a necessary step for deriving the inverse scope reading, is always illegitimate.

(37) One student in this class read more than three books.

It is crucial here that the closest node of type t to the comparative QP in object position is the vP node, which is structurally lower than the subject QP in the Spec of TP. Hence, the comparative QP is decomposed into two pieces below it. This analysis makes the following prediction. If comparative QPs can move to a position structurally higher than other QPs before they are decomposed into two pieces, they can take scope over those QPs. In such cases, intermediate scope of other QPs is not necessary because both of the decomposed QPs would have already outscoped the other QPs. In this section, I will discuss two such cases.

The first case can be observed in the dative construction. It is well known that scope freezing is observed in the double object construction (Aoun and Li 1989, Bruening 2001, and Larson 1988a). The indirect object must take scope over the direct object in this construction, as illustrated in (38).

(38) The teacher gave a student every book.

\[ (\exists > \forall) \quad *(\forall > \exists) \]

(Bruening 2001:235)

In contrast to this, the DP and the PP can show scope ambiguity in the dative construction, as shown in (39).

(39) The teacher gave a book to every student.

\[ (\exists > \forall) \quad (\forall > \exists) \]

(Bruening 2001:235)
Bruening claims that if there are two QPs which target the same node of type t, the structurally higher one must move first, and then the structurally lower one does "tucking in" beneath the first moved QP, because of a locality condition on movement (see Richards 1997 for the origin of this idea). On the basis of the syntactic and semantic facts, Bruening argues that there is an extra head between the two objects in the double object construction, following Marantz (1993), among others. This extra head (V₁ in the representation in (40)) establishes an asymmetric c-command relation between the two objects in the double object construction. Thus, the original structural hierarchy between the two QPs must remain the same even after they undergo QR, as illustrated in (40).

(40) \[ TP \text{ the teacher}_1 [VP \text{ a student}_2 [VP \text{ every book}_3 [VP \text{ t}_1 \text{ V}_1 [VP \text{ t}_2 \text{ gave t}_3] ]]]\]

This explains the fact that the indirect object must outscope the direct object in this construction. Bruening argues that the extra head does not exist in the dative construction and the direct object DP and the indirect object PP are equidistant from v. In other words, the two arguments are in a sisterhood relation, as shown in (41) (see Bruening 2001 and references cited therein for proposals on the structure of the dative construction).

(41) \[ TP \text{ the teacher}_1 [\text{ v}_1 \text{ t}_1 [\text{ VP gave a book to every student}]]\]

Thus, either one of the two QPs can move first. This is why we observe scope ambiguity in the dative construction.

The decompositional approach to comparative QPs, in tandem with the discussion above, makes the following prediction. Comparative QPs in indirect object position can take scope over QPs in direct object position in the dative construction. This is because a comparative QP can move first, and the other QP then tucks in beneath the decomposed comparative QPs. This is schematically represented in (42).

(42) \[ TP \text{ Subj}_1 [\text{ VP [DegP er than n}_2 [\text{ DP t}_2 \text{ many NP}_3 [\text{ VP QP}_4 [\text{ VP t}_1 \text{ V}_4 \text{ t}_3] ]]]\]
This option is not available in the cases discussed in the previous sections since Shortest Move forces comparative QPs in object position to target the vP node, which is structurally below the subject QP in the Spec of TP, and then forces them to be decomposed there. This prediction is in fact borne out, as illustrated in (43).

(43) a. John submitted some paper to more than five journals this month.
    (∃>CQP) (CQP>∃)

b. John donated two books to more than five churches this month.
    (two>CQP) (CQP>two)

These facts reveal once more that wide scope of comparative QPs is contingent upon structural considerations.

Let me move to the next case: the inverse linking construction discussed in Larson (1987), May (1985), May and Bale (2005), and Sauerland (2005), among many others. This construction is illustrated in (44).

(44) Some representative from every city came to the meeting.

An immediate question that arises for this construction is how to interpret it compositionally. In particular, one issue is how to interpret a QP embedded inside a larger QP (e.g., every city in (44)). Under the simplest assumptions, the closest position where the embedded QP can be interpreted is the TP node at the clausal level (see May 1985 for a different analysis). Here I assume that this QP takes clausal scope by moving out of the embedding QP. I provide a piece of supportive evidence for this assumption. The evidence comes from the scope splitting phenomenon in the inverse linking construction. Let us consider the sentence in (45), bearing the following context in mind.

There are several restaurants from which the host of the party thinks he should order a dish. He asked a group of food critics which of the restaurants he should order a dish from. Each one of these restaurants was recommended to the host by one of the food critics. The food critics do not know the decision made by the host. However, they expect that the restaurant that they recommended to the host has been chosen:
As for the menu at the party, a different food critic expects there to be a dish from every restaurant.

In (45), the DP *a different food critic* can receive a bound reading, which requires the universal QP to take scope over it. On the other hand, *a dish* must take narrow scope relative to the matrix predicate *expect* (i.e., a de dicto construal) because it is in the post-copular position in the existential construction (see 4.3.2 for this requirement). Thus, (45) is an instance of the scope splitting phenomenon, which suggests that the embedded QP in the inverse linking construction takes clausal scope.\(^{19}\)

On this assumption, the interpretable LF representation of (44) is the one in (46).

\[
(46) \quad \left[\text{TP} \left[\text{every city}\right]_1 \left[\text{TP} \left[\text{DP some representative from } t_1\right]_2 \left[\text{vP } t_2 \text{ came to the meeting}\right]\right]\right]
\]

It is crucial to notice that the closest node of type *t* is above the embedding QP. Thus, the embedded QP must QR over the embedding QP for reasons of interpretability. If the embedded QP is a comparative QP, it is decomposed into two pieces at the position above the embedding QP. Thus, we predict that in the inverse linking construction, an embedded comparative QP can take scope over the other QP. This prediction is also borne out, as shown in (47).

\(^{19}\) Contrary to what I suggest here, it has been claimed in the past literature that the scope splitting is impossible, which is based on facts such as (i) (see Barker 2001, 2002, Heim and Kratzer 1998, Larson 1987 and May and Bale 2005 for relevant data and discussion).

\[(i) \quad \text{Two politicians spy on someone from every city.}
\]
\[*(\forall>\text{two}>\exists) \ (\forall>\exists>\text{two})\]

(Larson 1987)

I have no account for the difference in the possibility of scope splitting between (i) and (45). However, I have one speculation. Sauerland (2005) argues that an intensional predicate can take scope between two relevant QPs in the inverse linking construction. In (45), there is a relevant intervening predicate between the two QPs (i.e., *expect*), but there is no such predicate in (i). It might be possible to capitalize on this difference to account for the facts in (45) and (i).
(47) The president invited some politician from more than thirty countries to the party.

This fact, like the evidence from the dative construction, demonstrates that comparative QPs can take scope over other QPs if they can move to a structurally higher position than the other QPs before they are decomposed.

In conclusion, I have argued that the difference between the cases discussed in this section and those in the previous sections is exactly what is predicted by the decompositional approach defended above. In addition, it has been revealed once more that wide scope of comparative QPs is dependent upon structural considerations.

4.5 Implications for the Architecture of Grammar

I have argued that the account of the Comparative QP Scope Generalization is that there is an intermediate stage of the derivation that must be established before a comparative QP in object position can have scope over a subject QP, and this intermediate stage is always ruled out by the Heim-Kennedy Constraint. Two assumptions have been employed to force the intermediate step. One of them is the locality constraint on SSOs, which prohibits long distance SSOs. The other is the assumption that a comparative QP is decomposed into the comparative operator and a quantificational DP before a subject QP undergoes QL. This last assumption, which I have implicitly assumed so far, is very crucial for the proposed analysis, especially because, if we rejected it, there would be an alternative path to yield wide scope of comparative QPs in object position. This path would avoid the illegitimate intermediate stage of the derivation. I will demonstrate that, for principled reasons, this alternative path is not an option in the first place. I will further argue that if the proposed explanation is successful, it can be taken as evidence for a particular derivational order of obligatory and optional operations.
Let me illustrate each step of the alternative derivation for wide scope of a comparative QP. As shown in (48)a, the comparative QP undergoes QR, which is a necessary step for later resolution of type mismatch. The difference between the alternative derivation and the derivation that I have assumed lies in what happens next. I have claimed that QR of the comparative QP occurs at this stage of the derivation. Suppose, alternatively, that before QR of the comparative operator, QL applied to the subject QP and moved it to the position under the comparative QP, as illustrated in (48)c. QR of the comparative operator would then take place at the last stage of the derivation. In this alternative derivation, wide scope of the comparative QP can be derived by SSOs that obey both the Heim-Kennedy Constraint and Shortest Move, as represented in (48)d.

(48) a. \([TP \text{ one student}_1 [vP t_1 \text{ read } [DP \text{ more than three books}]]]\\)
   \rightarrow \text{Obligatory QR of the CQP for interpretability}

b. \([TP \text{ one student}_1 [vP [DP \text{ more than three books}]_2 [vP t_1 \text{ read } t_2]]]\\)
   \rightarrow \text{Optional QL of the subject QP}

c. \([TP \text{ one student}_1 [vP [DP \text{ more than three books}]_2 [vP \text{ one student}_1 [vP t_1 \text{ read } t_2]]]]\\)
   \rightarrow \text{Obligatory QR of the comparative operator for interpretability}

d. \([TP \text{ [DP er than three]}_3 [vP [DP \text{ t many books}]_2 [vP \text{ one student}_1 [vP t_1 \text{ read } t_2]]]]\\)

The main issue that arises here is why QR of the comparative operator and QL of the subject QP must be carried out in the particular derivational order that I have assumed and not the other way around, like in the alternative derivation sketched above. It is noticeable that in (48), the optional SSO (i.e., QL of the subject QP) is sandwiched between the two obligatory operations (i.e., QR of the comparative QP and the comparative operator) that are necessitated by semantic reasons. The alternative derivation would not be a legitimate option if there were a constraint on the derivational order of operations that forced operations driven by syntactic and semantic requirements to be applied before optional operations like QL, henceforth the Derivational Order

20 This alternative possibility was pointed out independently by Sigrid Beck and Uli Sauerland (personal communication).
Given the Derivational Order Requirement, the derivation in which a comparative QP outscopes a subject QP must proceed in the order that I have argued for in previous sections, since QR of the comparative operator, which decomposes a comparative QP into two generalized quantifiers, is an operation required for interpretability.

The Derivational Order Requirement seems to be a natural requirement, and indeed it follows from the nature of optional operations in a principled way. In order to demonstrate this, it is necessary to acknowledge the division of labor between obligatory operations and optional SSOs. It is claimed that obligatory operations driven either by syntactic or semantic requirements are applied so as to derive an interpretable structure from an uninterpretable structure. On the other hand, optional SSOs are not driven by any semantic requirement, since their input structures are already interpretable. Instead, optional SSOs are applied in order to bring forth a designated interpretation. Fox (2000) and Reinhart (1995) argue that due to this very nature of optional SSOs, they cannot be applied unless they have an interpretive motivation for their application. In other words, optional SSOs must have an effect on semantic interpretation; this principle was called Scope Economy in Fox (2000). On this view of the division of labor, optional SSOs can never be applied before applications of obligatory operations are completed. As mentioned above, a structure that needs further applications of obligatory operations is not interpretable. If we applied an optional SSO to such a structure, the resulting structure would still be uninterpretable because optional SSOs are not operations that convert an uninterpretable structure into an interpretable structure. In other words, if optional SSOs applied before obligatory operations were completed, they would always end up yielding no semantic effect on their output. To put this in the context of the alternative derivation for wide scope of the comparative QP in (48), optional QL of the subject QP in (48)c is not permitted because it has no interpretive motivation in the sense

---

21 I assume that we must apply as few operations as possible in order to derive an interpretable representation. Kyle Johnson (personal communication) pointed out that if we did not make this assumption, there would be alternative derivations for wide scope of an object comparative QP relative to a subject QP that obey the relevant constraints. See fn. 26 for discussion about one such derivation.

22 In what follows, I am confined to discussing optional operations on the LF side, or, more specifically, optional SSOs.
that its outcome is still uninterpretable. Therefore, optional SSOs of this sort would be prohibited by Scope Economy. I claim that this is the principled reason for the Derivational Order Requirement.\(^{23}\)

This general view of grammar also provides an answer for the following question that arises from the derivational nature of optional SSOs. In 4.3.5, I have claimed that each step of an optional SSO has to be licensed by grammar. Each time an optional SSO is applied, its legitimacy needs to be checked right after its application. I have argued that the proposed explanation of the Comparative QP Scope Generalization can be regarded as evidence for the derivational character of optional SSOs. However, it is reasonable to ask why optional SSOs have such a derivational property. Another question raised by Sigrid Beck (personal communication) is why only optional SSOs have this derivational character. For instance, suppose that obligatory operations, such as QR for interpretability, also possess the derivational property and that each application of an obligatory operation needs to yield an LF representation licensed by grammar. Within this hypothetical framework, it would be necessary to check whether the obligatory QR of the comparative QP for interpretability is legitimate in (49) immediately after its application. However, this obligatory SSO does not provide an interpretable representation, due to the type mismatch caused by the DegP. Its application is wrongly predicted to be prohibited under the assumption that SSOs that do not yield an interpretable LF representation are not licensed by grammar. This leads us to conclude that obligatory SSOs do not possess the derivational property, in contrast to optional SSOs:

\[(49) \ [TP \text{John}_1 [vP [DP more than three]_2 [P t_1 \text{read} t_2]]] \]

Both of the questions receive a straightforward answer once we take into account the above-mentioned division of labor between obligatory operations and optional SSOs.

\(^{23}\) As Danny Fox (personal communication) pointed out, the suggested account of the Derivational Order Requirement needs to assume Fox’s (1995a) version of Scope Economy, which permits the application of SSOs only when they produce a designated interpretation. In Fox (2000), Scope Economy only refers to a scopal commutativity between two QPs.
Because applications of optional SSOs must have an effect on their output, each step of optional SSOs must be licensed by grammar and produce a designated interpretation. This general requirement of optional SSOs emerges as the derivational property of optional SSOs. On the other hand, obligatory operations need to derive an interpretable representation from an uninterpretable representation, as we have discussed above. Thus, they do not necessarily have an effect on output in the way that optional SSOs do. This difference between the two types of operations means that optional SSOs exclusively possess the derivational property.

To sum up this section, I have argued that the proposed explanation of the Comparative QP Scope Generalization is an argument for the specific derivational order of obligatory and optional operations. This order follows straightforwardly from the nature of the two types of operations, which in turn justifies the derivational character of optional SSOs.

### 4.6 Towards Explaining the Heim-Kennedy Constraint

#### 4.6.1 An Outline

The current analysis of the scope puzzle crucially relies upon the Heim-Kennedy Constraint, whose nature we do not fully understand. In the following subsections, I argue that the intermediate scope representation is ruled out in more principled ways. If this argument is successful, it paves the way for considering the Heim-Kennedy Constraint in a new light. The argument depends on the interpretation of the intermediate scope representation, which I have been conspicuously silent about in previous sections.
4.6.2 Scopal Commutativity: Monotone Increasing Comparative QPs

In order to fully understand the meaning of comparative QPs, let me first introduce the meaning of the determiner *many*. Hackl (2000) argues that it is a scalar determiner whose lexical entry is given in (50) (see Hackl 2000 for evidence in favor of this particular lexical entry). Once its first argument (of type d) is saturated, this determiner is of exactly the same type as other quantificational determiners, such as *every* and *some*, which are non-scalar determiners:

(50) \[ [[\text{many}]] := \lambda d \in D_d. [\lambda f \in D_{(\epsilon,0)} : [\lambda g \in D_{(\epsilon,0)} : d \text{-many } x \text{ are such that } f(x)=1 \text{ and } g(x)=1]] \]

(adapted from Hackl 2000:83)

Combining this lexical entry with the lexical entry of the comparative operator in (30) gives the correct interpretation for the sentence in (51)a, which is represented in (51)b.24

(51) a. John read more than three books.

b. \{d: 0 < d < 3\} \subset \{d: \text{John read } d \text{-many books}\}

(i.e., The number of books read by John is greater than three.)

24 The numeral expression *three* in the comparative complement in (51)a is apparently a referential numeral expression, which is of type d, though the comparative operator needs an element of type \((d, t)\) as its first argument, as illustrated in (30). As a possible way for the comparative complement to denote a set of degrees, I suggest that it involves the clausal structure illustrated in (i) in which there is a covert predicate represented asNUM, and a null operator movement takes place in the same way as in (31)b. The lexical entry of this predicate is presented in (ii).

(i) \[[\text{than OP}_1 \text{ three } t_1 \text{-NUM}]]

(ii) \[[\text{NUM}]] := \lambda d_1 \in D_d. [\lambda d_2 \in D_d. 0 \leq d_1 \leq d_2]

Given these assumptions, the comparative complement denotes the set of degrees, \(\{d: 0 \leq d \leq 3\}\), which is a suitable argument of the comparative operator (see fn. 27 for further discussion of the clausal analysis). See Hackl (2000) for a different clausal analysis of the comparative complement.
I assume that this determiner is monotone in the sense that, if John read \( n \) many books, it is also true that he read \( n-1 \) many books and so on.

Under this semantics for comparative QPs, let me now examine cases where the subject is an existential QP:

(52) One student in this class read more than three books.

I argue that the widest scope reading of the existential QP and its intermediate scope reading in (51) are truth-conditionally equivalent because existential QPs and the comparative operator are scopally commutative, as was observed by Heim (2001a). The scopal commutativity between the two is shown in Lemma I:

(53) Lemma I:

**Scopal commutativity between existential QPs and the comparative operator**

For every model \( M \), every assignment function \( g \), and every monotonic function \( \phi \in \mathcal{D}(a, \infty) \),

\[
\begin{align*}
\llbracket QP_1 \rrbracket^M_g (\lambda x. \llbracket QP_1 \rrbracket^M_g (\lambda d. \phi(d)(x))) &= \llbracket OP_{er} \rrbracket^M_g (\lambda d. \llbracket QP_1 \rrbracket^M_g (\lambda x. \phi(d)(x))) \\
\end{align*}
\]

where \( QP_1 \) is an existential QP and \( OP_{er} \) is a comparative operator.

**Proof:**

\[
\begin{align*}
\llbracket QP_1 \rrbracket^M_g (\lambda x. \llbracket OP_{er} \rrbracket^M_g (\lambda d. \phi(d)(x))) &= \llbracket OP_{er} \rrbracket^M_g (\lambda d. \llbracket QP_1 \rrbracket^M_g (\lambda x. \phi(d)(x))) \\
\rightarrow
\end{align*}
\]

\[\exists x[P(x) \text{ and } D \subseteq \{d: \phi(d)(x)\}] \equiv D \subseteq \{d: \exists x[P(x) \text{ and } \phi(d)(x)]\}\]

where \( P \) is a restrictor of an existential QP and \( D \) is a restrictor of a comparative operator.
a. \( \exists x[P(x) \text{ and } D \supset \{d: \varphi(d)(x)\}] \rightarrow D \supset \{d: \exists x[P(x) \text{ and } \varphi(d)(x)]\} \)

Suppose that \( \exists x[P(x) \text{ and } D \supset \{d: \varphi(d)(x)\}] \).

Choose an arbitrary individual \( \alpha \) that makes the formula above true.

\( P(\alpha) \text{ and } D \supset \{d: \varphi(d)(\alpha)\} \)

\rightarrow

\( D \supset \{d: P(\alpha) \text{ and } \varphi(d)(\alpha)\} \)

\rightarrow

\( D \supset \{d: \exists x[P(x) \text{ and } \varphi(d)(x)]\} \)

b. \( D \supset \{d: \exists x[P(x) \text{ and } \varphi(d)(x)]\} \rightarrow \exists x[P(x) \text{ and } D \supset \{d: \varphi(d)(x)\}] \)

Suppose that \( D \supset \{d: \exists x[P(x) \text{ and } \varphi(d)(x)]\} \).

Then \( D \supset \bigcup_{\alpha \in D} \{d: P(\alpha) \text{ and } \varphi(d)(\alpha)\} \).

By monotonicity, each of these sets is an interval, \( \{d: 0 \leq d \leq d_{\alpha}\} \), where \( d_{\alpha} \) is the maximal degree \( d \) such that \( P(\alpha) \text{ and } \varphi(d)(\alpha) \).

As these all intervals start at 0, their union is itself \( \{d: 0 \leq d \leq d_{\alpha'}\} \), where \( \alpha' \) has the greatest \( d_{\alpha'} \).

Then \( D \supset \{d: 0 \leq d \leq d_{\alpha'}\} \)

\rightarrow

\( \exists x[P(x) \text{ and } D \supset \{d: \varphi(d)(x)\}] \).

QED.

I propose that since an SSO for deriving intermediate scope has no semantic effect, it is prohibited by Scope Economy in (24), the formal definition of which is given in (54).\(^{25}\)

\(^{25}\)Although Scope Economy was originally defined only for generalized quantifiers over individuals, it can be straightforwardly extended in order to accommodate the scopal commutativity between generalized quantifiers over both individuals and degrees.
(54) **Scope Economy**

An SSO can move XP₁ from a position in which it is interpretable only if the movement crosses XP₂ and <XP₁, XP₂> is not scopally commutative.

(Fox 2000:26)

We have now seen that this explanation accounts for the fact that a comparative QP in object position cannot take scope over an existential QP in subject position.

The same explanation can be applied to (55), in which the subject is a universal QP.

(55) Every student in this class read more than three books.

In order to illustrate this point, I present Lemma II, which demonstrates the scopal commutativity between universal QPs and the comparative operator. This was also observed in Heim (2001a):

(56) **Lemma II:**

**Scopal commutativity between universal QPs and the comparative operator**

For every model M, every assignment function g, and every monotonic function ϕ∈D(d, c, t),

\[
\llbracket QP_1 \rrbracket^M, g (\lambda x. \llbracket OP_{er} \rrbracket^M, g (\lambda d. \, \phi(d)(x))) = \llbracket OP_{er} \rrbracket^M, g (\lambda d. \llbracket QP_1 \rrbracket^M, g (\lambda x. \phi(d)(x)))
\]

where QP₁ is a universal QP and OPₚₚ is a comparative operator.
Proof:
\[
[[\text{QP}_1]]_{M}^{M,g} (\lambda x. [[\text{OP}_{er}]]_{M}^{M,g} (\lambda d. \varphi(d)(x))) = [[\text{OP}_{er}]]_{M}^{M,g} (\lambda d. [[\text{QP}_1]]_{M}^{M,g} (\lambda x. \varphi(d)(x)))
\]

\[
\forall x [P(x) \rightarrow D \subset \{d: \varphi(d)(x)\}] \equiv D \subset \{d: \forall x [P(x) \rightarrow \varphi(d)(x)]\}
\]
where \(P\) is a restrictor of a universal QP and \(D\) is a restrictor of a comparative operator.

To prove that this is true, we must show that
\[
\forall \alpha [P(\alpha) \rightarrow D \subset S_{\alpha}] \equiv D \subset \bigcap_{\alpha \in P} S_{\alpha}, \text{ where } S_{\alpha} = \{d: \varphi(d)(\alpha)\}.
\]

If \(D\) is a proper subset of each set \(S_{\alpha}\), it must be the case that \(D\) is a proper subset of the intersection of the sets, assuming that the domain is finite.

Conversely, if \(D\) is a proper subset of the intersection of the sets, it must be the case that \(D\) is a proper subset of each set \(S_{\alpha}\).

QED.

Just as in (52), Scope Economy disallows the application of an optional SSO in order to derive the intermediate scope representation in (55), since this SSO has no semantic effect.\(^{26}\)

\(^{26}\) As mentioned in fn. 21, I assume that we must apply as few operations as possible in order to derive an interpretable representation. Let me illustrate why I need to adopt this assumption by examining the derivation in (i), which is postulated for wide scope of the comparative QP relative to the subject QP in (55). This alternative derivation, which was pointed out by Kyle Johnson (personal communication), seems to obey all the constraints that we discussed so far. First, only the comparative operator -er is base-generated within the DP, following Bhatt and Pancheva’s (2004, 2005) insights (see 3.3.3 for relevant discussion), and it undergoes movement successively cyclically to a position above the subject QP, as illustrated in (i)a and (i)b, respectively. The restrictor constituent of the comparative operator merges with the comparative operator counter-cyclically, as shown in (i)c. The representation in (i)c is an interpretable structure which produces the surface scope reading. Finally, the application of an optional SSO derives narrowest scope of the subject QP in (i)d by lowering it. This application of an optional SSO obeys Scope Economy because the universal QP and the DP many books are not scopally commutative, which incorrectly generates the inverse scope reading:
In this subsection, I have demonstrated that the intermediate scope reading is not derivable in (52) and (55), due to an independently motivated principle: Scope Economy. Thus, we can arrive at the same result as before without relying on the Heim-Kennedy Constraint. The next subsection demonstrates that an analogous analysis can be applied to cases with monotone decreasing comparative QPs.

### 4.6.3 Scopal Commutativity: Monotone Decreasing Comparative QPs

As discussed above, the formula in (57)b captures the truth conditions of the sentence in (57)a.

\[(57)\]
\[
\begin{align*}
\ a. & \text{ John read more than three books.} \\
\ b. & \{d: 0 \leq d \leq 3\} \subset \{d: \text{John read } d\text{-many books}\} \\
& \text{(i.e., The number of books read by John is greater than three.)}
\end{align*}
\]

In the semantics of comparatives employed here, the truth conditions of (58)a, which involves a monotone decreasing comparative QP, can be represented by the formula in (58)b.

\[(58)\]
\[
\begin{align*}
\ a. & \text{ every student read } [\text{DP } [\text{DegP -er} \text{ many books}]] \\
\ b. & [\text{TP } [\text{DegP -er}] [\text{TP every student} [\_p t_1 [\_p [\text{DP } t_1 \text{ many books}] [\_p t_2 \text{ read } t_3]]]]] \\
\ c. & [\text{TP } [\text{DegP -er [than three]}] [\text{TP every student} [\_p t_1 [\_p [\text{DP } t_1 \text{ many books}] [\_p t_2 \text{ read } t_3]]]]] \\
\ d. & [\text{TP } [\text{DegP -er [than three]}] [\_p t_1 [\_p [\text{DP } t_1 \text{ many books}] [\_p every student} [\_p t_2 \text{ read } t_3]]]]]
\end{align*}
\]

If we adopt the assumption mentioned at the beginning of this footnote, the second step of the successive cyclic movement of the comparative operator in (i)b would not be allowed because this step is not necessary to derive an interpretable representation.
(58)  a. John read fewer than three books.
    b. \{d: John read d-many books\} \subseteq \{d: 0 \leq d \leq 3\}
        (i.e., The number of books read by John is smaller than three.)

The question is how to derive this formula compositionally from the LF representation of (58)a. The set-theoretic principle in (59) helps us answer this question.

(59)  For any sets A, B such that A \subseteq U and B \subseteq U: A \subseteq B iff U-B \subseteq U-A
       (Heim 2001b:11)

Given (59), the formula in (58)b is equivalent to the following one:

(60)  \{d: \sim 0 \leq d \leq 3\} \subseteq \{d: John read \sim d-many books\}

In (60), the two sets of numbers are complement sets of the corresponding sets in (58)b. In order to derive (60) or, equivalently, (58)b, following Heim’s (2001b) insights, I argue that monotone decreasing comparative QPs contain a negative component little as well as the comparative operator and the scalar determiner many. The lexical entry of little is given in (61).

(61) \llittle := \lambda d \in D_d. [\lambda P \in D_{(d,0)}, P(d)=0]
       (Heim 2001b:11)

This negative component takes a set of degrees, D, and returns the complement set of D. As for the syntax, I would like to suggest that the comparative operator and the negative component little form a DegP shell structure, as illustrated in (62).

(62) \text{[DP [DegP [DegP er than three] little] many books]}
As is clear from the lexical entry of *little* in (61), once its first argument (of type d) is saturated, the result is a generalized quantifier. Thus, the constituent that consists of *little* and its first argument must undergo QR to a node of type t for resolution of type mismatch. The interpretable LF representation in (63) is assigned to (58)a and the formula in (58)b is compositionally derived from this LF representation.\(^{27}\)

\[
(63) \quad [\text{TP John}_1 \quad [\text{vP} \quad [\text{DegP} \quad \text{er than three}]] \quad [\text{vP} \quad [\text{DegP} \quad \text{t}_2 \quad \text{little}]] \quad [\text{vP} \quad [\text{DP} \quad \text{t}_3 \quad \text{many books}]]] \\
\quad \quad [\text{vP} \quad \text{t}_4 \quad \text{read} \quad \text{t}_4]]
\]

At this point, one might wonder why we need to postulate this particular version of the decompositional approach to monotone decreasing comparative QPs. An alternative way to analyze them would be to decompose them into the comparative operator and a scalar determiner *few*, which is an antonymous counterpart of the scalar determiner *many* and hence, incorporates a negative meaning in itself. The lexical entry of *few* is given in (64).

\[
(64) \quad \llbracket \text{few} \rrbracket := \lambda d \in D_d. \ [\lambda f \in D_{(0,0)} \cdot \lambda g \in D_{(0,0)} \cdot \sim_d \text{-many} \ x \text{ are such that } f(x)=1 \ \text{and } g(x)=1]
\]

I will now present an argument in favor of the more fine-grained decompositional approach proposed above. The evidence comes from the scope splitting phenomenon observed by Hackl (2000). Let us consider the sentence in (65), in which there are an intensional predicate and a monotone decreasing comparative QP.

---

\(^{27}\) Notice that the set of numbers denoted by the comparative complement is \{d: 0\leq d\leq 3\} in (63). However, what we need is a complement set of this set (see (60) for this point). In order to obtain the desired result, I capitalize on the clausal analysis of the comparative complement discussed in fn. 24 and argue that the negative element *little* in the matrix clause is copied into the comparative complement. This yields the complement set of the set mentioned above. See Heim (2001b) for relevant discussion.
(65) a. At MIT one needs to come up with fewer than five brilliant ideas to get tenure. 

(Hackl 2000:156)

b. (de dicto: need>er>little>many)
#Too much productivity excludes one from getting tenure.

c. (de re: er>little>many>need)
#One has to come up with some brilliant ideas that already exist.

d. (scope splitting: er>little>need>many)
Anyone who comes up with more than four brilliant ideas is not excluded from getting tenure.

As illustrated in (65)b-(65)d, there are three truth-conditionally distinct meanings to (65)a.\(^{28}\) The first interpretation of (65)a is a de dicto reading in which all generalized quantifiers take narrow scope relative to the intensional predicate need. In this particular sentence, the de dicto reading is pragmatically less salient, given our knowledge of what is required to get tenure. The de re reading in (65)c is also excluded because the presupposition of the existence of the ideas is incompatible with the creation verb come up with. The pragmatically most salient reading of (65)a is the scope splitting reading in (65)d. We have here an argument in favor of our decompositional approach. In order to derive the scope splitting reading in (65)d, the negative element little needs to take its scope at a position distinct from the determiner many. Notice that the meanings of little and many constitute a single determiner few in (64). Assuming the lexical item few in (64), there would be three scopal relations in (65)a, as illustrated in (66). Under the assumption that need involves universal quantification over worlds, intermediate scope of need in (66)b produces the de dicto interpretation because the comparative operator is scopally commutative with need (see Lemma II in (56) and also 4.8.3 for relevant discussion).

\(^{28}\) There is another potential scopal relation in (65)a, namely, one in which the comparative operator outscopes the intensional predicate need, which takes scope over other QPs. This scopal relation provides the same truth conditions as the de dicto reading in (65)b. See 4.8.1 for relevant discussion.
a. widest scope of need
   need > er > few → de dicto
b. intermediate scope of need
   er > need > few → de dicto
c. narrowest scope of need
   er > few > need → de re

Since the meanings of little and many constitute the single determiner few in (64) and the negative component could not take a different scope than the determiner many, we would have no way to derive the scope splitting reading in (65)d. Therefore, the fact in (65) can be taken as evidence in favor of our decompositional approach.

Let us now return to the sentences in (67), which are our main empirical concerns and examine them in the proposed decompositional approach.

(67) a. One student in this class read fewer than three books.
   b. Every student in this class read fewer than three books.

In these cases, there are four potentially legitimate scopal relations, as illustrated in (68).

(68) a. widest b. intermediate₁ c. intermediate₂ d. narrowest

However, the relevant concern is whether the intermediate₁ scopal relation in (68)b, which is derived from the surface scope relation by the first application of an optional SSO, observes the relevant principles of grammar. As we have already shown in Lemma I in (53) and Lemma II in (56), existential and universal QPs are scopally commutative with the comparative operator. In (67)a and (67)b, we can also prevent the first application of an optional SSO by making reference to Scope Economy, with no reliance
on the Heim-Kennedy Constraint. Consequently, only the surface scope reading is available in these cases.

Before closing this subsection, let me point out one implication of the proposed analysis of the scope freezing in (67) for the decompositional analysis of monotone decreasing comparative QPs. We have argued against the presumed decompositional analysis in which they are composed of the comparative operator and the scalar determiner few in (64). It has been shown that this version of the decompositional approach fails to yield the scope splitting reading in (65)d because it does not allow the negative component little and the determiner many to take scope at different positions. However, there is another alternative decompositional analysis of monotone decreasing comparative QPs, which is compatible with the scope splitting phenomenon in (65). This alternative approach, pursued in Hackl (2000), assumes that the comparative operator and little constitute a single comparative operator. This operator has a meaning identical to that of the comparative operator that Heim (2001a) assumes in less-comparatives (see fn. 16 for the less-comparative operator):

\[
(69) \quad [[\text{less}]] := \lambda D \in D_{(d, 0)}. [\lambda D' \in D_{(d, 0)}. D \supset D']
\]

However, this account cannot rule out the unavailable readings in (67) like the proposed analysis does. If the less-comparative operator in (69) took scope over the subject QP, it would yield the reading that corresponds to the intermediate 2 scopal relation in (68)c, which is truth-conditionally different from the surface scope reading in (68)a. As a consequence, there would be no simple manner to prohibit the inverse scope reading in (67). In tandem with the scope splitting phenomenon in (65), the frozen scope in (67) offers support to the specific decompositional approach that I am arguing for here.
4.6.4 Cross-Polar Anomaly

So far, Scope Economy is the only principle that excludes the intermediate scopal relation. I will now demonstrate that it is sometimes excluded by a semantic constraint which applies specifically to comparatives.

Let us first consider the following anomalous sentence, which Kennedy (1999, 2001) refers to as “cross-polar anomaly”:

(70) ?Alice is shorter than Carmen is tall.

(Kennedy 2001:36)

On the basis of sentences such as (70), among others, Kennedy argues that the polarity of adjectives in antonymous pairs (e.g., *tall* vs. *short*) should be captured in terms of the structural distinction between positive degrees and negative degrees. Positive adjectives, like *tall*, relate an object to positive degrees; they are intervals that extend from a lower end to a certain point of a scale. On the other hand, an object is projected to negative degrees by negative adjectives, such as *short*. Negative degrees are intervals from a certain point to the upper end of a scale. These two types of degrees are formally defined below:

(71) For any scale S,

a. \( \text{POS}(S) = \{d \subseteq S \mid \exists p_1 \in d \forall p_2 \in S[p_2 \leq p_1 \rightarrow p_2 \subseteq d]\} \)

b. \( \text{NEG}(S) = \{d \subseteq S \mid \exists p_1 \in d \forall p_2 \in S[p_1 \leq p_2 \rightarrow p_2 \subseteq d]\} \)

(Kennedy 2001:53)

Given the distinction between the two types of adjectives, (70) can be illustrated by the following diagram (assuming that Alice is 5’ 2” tall and Carmen is 5’ 4” tall):  

168
I would like to suggest that the ordering relation denoted by the comparative operator is constrained by a semantic condition on the polarity of sets of degrees. The basic idea is shown in (73).

(73) The Cross-Polar Anomaly Constraint

The partial ordering relation denoted by the comparative operator is defined only over sets of degrees that contain the same end of a scale.

This is formally implemented by building it into the lexical entry of the comparative operator as a definedness condition, as suggested by Irene Heim (personal communication). The modified lexical entry of the comparative operator is given in (74).

(74) $[\cdot -er \cdot] := \lambda D \in D_{(d, 0)}: \{\lambda D' \in D_{(d, 0)}: D' \subseteq D \text{ or } D \subseteq D', \text{ } D \cap D'\}$

In (70), one of the two sets of degrees is not a subset of the other, as is clear from (72). Thus, the definedness condition on the comparative operator is not satisfied in (70) and the operator receives no semantic value, due to presupposition failure. I suggest that this results in anomaly.²⁹

²⁹ Alternatively, one can keep pursuing the original lexical entry of the comparative operator in (30), and assume that the cross-polar anomaly cases like (70) are degraded because they are contradictions. In the past literature, several attempts have been made to derive ungrammaticality from purely semantic considerations. For instance, Barwise and Cooper (1981) argue that the fact that strong quantifiers such as every NP cannot appear in post-copular position in the existential there-construction is attributed to semantic insignificance of sentences, like tautologies and contradictions. von Fintel (1993) proposes an account of a co-occurrence restriction on the exceptive phrase but with resort to immediate falsity of sentences. It is certainly true that not all sentences of this sort are ungrammatical. But, see Gajewski (2002) for a formal procedure to determine when semantic insignificance of sentences is reflected in their grammatical status. See also Fox and Hackl (2005) for further discussion.
In what follows, I demonstrate that intermediate scope leads to the sentence having no truth value, due to the definedness condition on the comparative operator, in cases where the subject is a monotone decreasing QP and ones where it is a non-monotonic QP. First, I present the surface scope and intermediate scope reading of the monotone decreasing QP in (75).

(75)  
  a. No student in this class read more than three books.  
  b. Surface Scope (no>er>many) 
     \[ \neg \exists x \{ x \text{ is a student in this class and } \{ d : 0 \leq d \leq 3 \} \subseteq \{ d : x \text{ read } d\text{-many books} \} \} \]  
  c. Intermediate Scope (er>no>many) 
     \{ d : 0 \leq d \leq 3 \} \subseteq \{ d : \neg \exists x \{ x \text{ is a student in this class and } x \text{ read } d\text{-many books} \} \} \]

The formula in (75)b correctly represents the surface scope reading of (75)a. In (75)c, the set of numbers on the right-hand side is identical to the set of numbers \( n \) such that no student in this class read \( n \) many books. Suppose that the student who read the most books in this class read ten books. In such a situation, the set of numbers on the right-hand side in (75)c is represented as the interval from eleven to the upper end of the scale, as shown in the diagram in (76). However, the set of numbers on the left-hand side in (75)c is identical to the interval from zero to three:

(76)  

\[ \begin{array}{cccccccc}
0 & 3 & 11 & \ldots & \infty
\end{array} \]

As is clear from the diagram in (76), one of the two sets of numbers is not a subset of the other; hence, the ordering relation between the two sets of numbers is undefined, given the lexical entry of the comparative operator in (74). Thus, the intermediate scope relation in (75)c is without truth value.
A similar problem arises in (77)a, in which the subject is a non-monotonic QP. The two relevant scopal relations are represented in (77)b and (77)c.

(77) a. Exactly two students in this class read more than three books.
    b. Surface Scope (exactly 2>er>many)
       \[ \{x: x \text{ is a student in this class and } \{d: 0 \leq d \leq 3\} \subseteq \{d: x \text{ read } d\text{-many books}\}\} \] = 2
    c. Intermediate Scope (er>exactly 2>many)
       \[ \{d: 0 \leq d \leq 3\} \subseteq \{d: \{x: x \text{ is a student in this class and } x \text{ read } d\text{-many books}\} \] = 2

Like in the previous case, the surface scope reading is indicated by (77)b. The set of numbers on the right-hand side in (77)c is the set of numbers \( n \) such that exactly two students read \( n \) many books. A concrete context will help in understanding what the intermediate scope reading in (77)c implies. Suppose that there are three students A, B, and C. Assume further that A read one book, B, seven books, and C, nine books. In such a situation, the set of numbers on the right-hand side in (77)c contains the numbers from two to seven, as illustrated in the diagram in (78).

(78)

\[
\begin{array}{c}
0 & 1 & 7 & 9 & \infty \\
A & & B & & \\
2-7 & \hline \hline C
\end{array}
\]

It is the case that the set of numbers on the right hand side in (77)c contains neither end of the scale whenever there are more than two students. It is clear that the set of numbers from zero to three is never a subset of the set of numbers that contains neither end of the scale. Thus, the ordering relation denoted by the comparative operator is undefined in the intermediate scopal relation in (77)c and is without truth value.
I would like to claim that grammar disallows representations that are necessarily truth-valueless interpretation and that, hence, cross-polar anomaly cases and the intermediate scope representations in the cases under discussion are not legitimate. Since the first step of an optional SSO violates principles of grammar, we cannot apply any further SSO in order to derive wide scope of the comparative QP. This is the reason why only the surface scope reading is available for the sentences in (75)a and (77)a.

4.6.5 Interim Conclusion

The account of the Comparative QP Scope Generalization provided in section 4.3 depended crucially upon the Heim-Kennedy Constraint. The present section argued that the intermediate scope representation can be ruled out based on independently motivated constraints, namely, Scope Economy in some cases and the semantic constraint on the comparative operator in other cases. Thus, we now obtain a more principled account of the generalization.

4.7 The Heim-Kennedy Constraint or Scope Economy?

4.7.1 An Outline

In this section, I will discuss cases in which adopting either the Heim-Kennedy Constraint or the alternative discussed in section 4.6 leads to different predictions about the available scopal possibilities. Throughout this section, I will limit discussion to cases of comparative QPs in subject position. First, I will present evidence from Szabolcsi (1997) that comparative QP subjects may take narrow scope relative to QPs in object position, contra the predictions of the Heim-Kennedy Constraint-based analysis in section 4.3. On the basis of a fact about ellipsis, I will furthermore demonstrate that QPs
can take scope between the comparative operator and its trace. This is, again, unexpected by the Heim-Kennedy Constraint, but entirely natural under the Scope Economy explanation given in section 4.6. Finally, I point out that a specific prediction is made by the decompositional approach to monotone decreasing comparative QPs. That is, a certain reading that has not been noticed in the past literature is predicted to be available in addition to the surface and inverse scope readings. I argue that this prediction is borne out.

### 4.7.2 Monotone Increasing Comparative QPs in Subject Position

All of the cases that we have observed so far are ones in which comparative QPs occupy an object position. Let me now discuss cases in which monotone increasing comparative QPs are in subject position. As illustrated in (79), a universal QP in object position can take scope over a comparative QP in subject position (see Szabolcsi 1997 for relevant data).

(79) More than three students in this class read every book.

(CQP>V) (V>CQP)

The surface scope reading is straightforwardly produced by the scopal relation in which the two decomposed QPs outscope the universal QP. To this surface scope representation, an optional SSO could in principle apply, deriving the intermediate scope for the universal QP illustrated in (80)a. It is clear that this representation violates the Heim-Kennedy Constraint:
However, it turns out that the intermediate scope representation in (80)a delivers the so-called inverse scope reading of (79), as illustrated in (80)b. Widest scope of the universal QP produces the same interpretation, since universal QPs and the comparative operator are scopally commutative (Lemma II in (56)). Indeed, given the general framework defended above, the SSO that yields widest scope of the universal QP is prohibited by Scope Economy. Note, however, that intermediate scope of the universal QP creates a structure violating the Heim-Kennedy Constraint. Hence, if we were to adopt the Heim-Kennedy Constraint to rule out the intermediate scopal relation, as we did in section 4.3, we would incorrectly predict that (79) should not have an inverse scope reading. Since our current explanation for the Comparative QP Scope Generalization allows an intermediate scopal relation in this context, we correctly predict that (79) does have an inverse scope reading.

The difficulty in arguing against the Heim-Kennedy Constraint on the basis of (79) is that there is no empirical evidence against the claim that the inverse scope reading is derived from widest scope of the universal QP. In other words, there is no way to distinguish intermediate scope of the universal QP from widest scope based on intuitions about interpretations derived from these scopal relations. This is the same situation as that arising from sentences such as (81). In (81), the surface scope reading and the inverse scope reading are truth-conditionally equivalent because the two universal QPs are scopally commutative. Thus, it is impossible to identify the scopal relation between the two universal QPs by utilizing intuitions about interpretations derived from these scopal relations:

(81) Every student admires every teacher.
However, Fox (1995a, 2000) argues that the relative scope of scopally commutative QPs can be identified in ellipsis contexts because it is likely that the parallelism condition has the consequence that QPs take parallel scope in the antecedent clause and the ellipsis clause (but, see fn. 10 for caveat). To illustrate this point, let us consider (82).

(82) A boy admires every teacher. Every girl does <admire every teacher>, too.

\((\exists \forall) \neq (\forall \exists)\)

(Fox 2000:38)

From this perspective of the parallelism condition, the fact that the antecedent clause in (82) is restricted to the surface scopal relation suggests that the inverse scopal relation is impossible in the ellipsis clause in (82), which is explained by Fox’s (2000) Scope Economy, as discussed above.

Following Fox’s methodology, I employ ellipsis in order to identify the scopal relation between the comparative operator and a universal QP in sentences such as (79). Assuming the context given in (83)a, let us consider (83)b, in which the sentential complement of the verb told in the comparative complement is elided (see Heim 2001a for an analogous case). It is judged to be a valid complaint (see fn. 31):

(83) a. If every class last semester was attended by ten various students, the department is eligible for additional funds. The department chair asked the administrator how many people attended every class. The administrator answered that nine students attended every class because nine, ten and eleven students attended class A, B and C, respectively. However, it turned out that twelve students attended class A and the department chair complained to the administrator as follows:

b. More students attended every class than you had told me.

The rational behind the complaint is that ten students actually attended in the class that was attended by the smallest number of the students and that this number is greater than the number that the administrator had told the department chair. To get this interpretation,
the universal QP must be interpreted within the restrictor constituent of the comparative operator. More specifically, the universal QP needs to take scope over the QP many students within this constituent, as illustrated in (84)a, whose meaning is illustrated in (84)b.30

\[(84)\]
\[a. \text{[than Op}_1\text{ you had told me }<[\text{every class}_2 [[\text{DP t}_1 \text{ many students}] \text{ attended } t_2]]>]\]
\[b. \{d: \text{you had told me that } \forall x[x \text{ is a class } \rightarrow \text{d-many students attended } x]\} \]
\[(\text{i.e., the number } n \text{ such that the administrator told the department chair that } n \text{ many students attended the class that was attended by the least students})\]

The fact that ellipsis of EC is licensed in (83)b under the relevant interpretation indicates that the universal QP takes scope at the same position in the antecedent clause (i.e., in the matrix clause). As illustrated in the LF representation in (85)a, this is the intermediate scope of the universal QP. Thus, the sentence in (83)b is an argument against the Heim-Kennedy Constraint. However, the fact in (83)b is expected in our Scope Economy-based approach. Because the universal QP and the DP many students are not scopally commutative, an optional SSO that produces intermediate scope of the universal QP is licensed:

\[(85)\]
\[a. \text{[[DegP er than Op}_1\text{ you had told me }<[\text{every class}_2 [[\text{DP t}_1 \text{ many students}] \text{ attended } t_2]]>]>\]
\[\quad [[\text{every class}_4 [[\text{DP t}_3 \text{ many students}] \text{ attended } t_4]]]\]
\[b. \{d: \text{you had told me that } \forall x[x \text{ is a class } \rightarrow \text{d-many students attended } x]\} \subset \{d: \forall x[x \text{ is a class } \rightarrow \text{d-many students attended } x]\}\]

30 It has been observed that QPs in the comparative complement must take scope over the comparative operator (see Larson 1988b, Rullmann 1995, and Schwarzchild and Wilkinson 2002 for relevant observation and discussion). However, Danny Fox and Elena Guerzoni (personal communication) both pointed out that this requirement is obviated in cases where QPs in the comparative complement are in the scope of intensional predicates.
In contrast to (83)b, the sentence in (86), which is the passive version of (83)b, is not judged to be a valid complaint in the context given in (83)a.\textsuperscript{31}

(86) Every class was attended by more students than you had told me.

This also follows from the proposed approach. As is clear from the simplified LF representation postulated for the surface scope reading of (86) in (87)a, the comparative operator needs to take scope over the universal QP in subject position in order for this universal QP to be interpreted within the restrictor constituent of the comparative operator, as in (85)a.

(87) a. \[
[TP \text{every class was } [VP [\text{DegP } \text{er than you had told me}] \text{]} [VP [\text{DP } t_1 \text{ many students} [VP ... ]]]]
\]

b. \[\forall x \{x \text{ is a class } \rightarrow \{d: \text{ you had told me that } x \text{ was attended by } d\text{-many students} \subseteq \{d: x \text{ was attended by } d\text{-many students}\}\}\]

However, since the universal QP and the comparative operator are scopally commutative (Lemma II in (56)), an optional SSO for producing intermediate scope of the universal QP is prohibited by Scope Economy. Therefore, the sentence in (86) is restricted to the surface scope interpretation given in (87)b. For this scope reading to be true, it must be the case that, for every class, the number of its attendees is greater than the number of attendees told to the department chair by the administrator. However, this is not a valid complaint in the context in (83)a because the information given about class B and C was correct.

Thus, we have a more direct argument for the approach defended in section 4.6 and have good reason to believe that the Heim-Kennedy Constraint should be rejected.

\textsuperscript{31} Six out of the eight native speakers who I consulted got the contrast between (83)b and (86). For the two informants who did not get the contrast, neither (83)b nor (86) is a valid complaint in the context in (83)a.
4.7.3 Monotone Decreasing Comparative QPs in Subject Position

In order to demonstrate that the Heim-Kennedy Constraint and the current analysis based on Scope Economy make different predictions, it is illuminating to consider cases in which a monotone decreasing comparative QP is in subject position. The relevant example is given in (88), in which the inverse scope reading is obtainable.

(88) Fewer than three students in this class read every book.

\[(\text{CQP} \succ \forall) \quad (\forall \succ \text{CQP})\]

As argued in the previous sections, the comparative QP fewer than three students is composed of the comparative operator, little, and many students, and there are four possible scopal relations for (88). They are illustrated in (89).

(89) a. narrowest b. intermediate₁ c. intermediate₂ d. widest

\[
\text{[er than 3]₁} \quad \text{[er than 3]₂} \quad \text{[er than 3]₃} \quad \text{[er than 3]₄}
\]

First, notice that the intermediate₂ scope and the widest scope reading of the universal QP are truth-conditionally equivalent (Lemma II in (56)). Second, the inverse scope reading of (88) is produced from the intermediate₂ scope of the universal QP because an SSO yielding its widest scope is blocked by Scope Economy since the two readings are truth-conditionally equivalent. Thus, the scopal relations in (89)a and (89)c offer the surface scope and the inverse scope readings, respectively. Notice that there is another scopal relation, namely, intermediate₁ scope, which is produced by the first application of an optional SSO, as illustrated in (89)b. The question is what truth conditions are given by this scopal relation and whether it observes the principles of grammar. If (89)b were to violate any principle of grammar, the proposed approach...
could not generate the intermediate scope reading (i.e., the inverse scope reading) either, contrary to fact. First, let me illustrate the truth conditions under this scopal relation, which are provided in (90).

\[(90) \quad \text{Intermediate Scope (er>little>\forall>many)} \]

\[\{d: -0 \leq d \leq 3\} \subseteq \{d: \forall x [x \text{ is a book} \rightarrow d\text{-many students in this class read } x]\} \]

by (59)

\[\{d: \forall x [x \text{ is a book} \rightarrow d\text{-many students in this class read } x]\} \subseteq \{d: 0 \leq d \leq 3\}\]

The reading in (90) is true if fewer than three students read the book that was read by the smallest number of the students. In other words, it turns out to be true even if some popular books were read by more than two students. Since this reading is truth-conditionally distinct from the surface scope reading, the SSO producing this scopal relation does not violate Scope Economy. In addition, the meaning is semantically well-defined because one of the two sets is a subset of the other. Although the intermediate scopal relation observes the principles of grammar discussed in section 4.6, the issue is whether this interpretation is available for the sentence in (88). It is certainly true that it is quite difficult to obtain this meaning in (88).

However, my suggestion here is that the intermediate scope reading can actually be observed in some sentences. Such examples are given in (91) and (92). The crucial sentence is uttered by speaker B in (91) and administrator B in (92):

\[(91) \quad \text{Speaker A: One thousand people died in every modern war.} \]

\[\text{Speaker B: No, you are wrong.} \]

\[\text{A million people died in some of the modern wars.} \]

\[\text{But, I am pretty sure that fewer than one thousand people died in every modern war.} \]

---

\[32\quad \text{Five out of the eight informants that I consulted found the dialogues felicitous. Some speakers need a pitch accent on the determiner } \text{every} \text{ in order to get the relevant reading. I have no account of this fact.}\]
(92) Administrator A: Eight students registered for every class.

Administrator B: No, that’s wrong.

Ten students registered for many of the classes.

But, I am pretty sure that fewer than eight students registered for every class.

In (91), the intended construal of the final statement is that fewer than one thousand people died in the modern war in which the smallest number of people died. If the intermediate₁ scope reading is available for the last sentence in (91), the dialogue is judged felicitous. Similar reasoning applies to (92). In fact, at least some native speakers found the dialogues in (91) and (92) felicitous.

Let me sum up the crucial points of the discussion in this subsection. The first three scopal relations in (89) observe the principles of grammar, and the SSO deriving widest scope of the universal QP is blocked by Scope Economy because it is truth-conditionally equivalent to its intermediate₂ scope. The two intermediate scopal relations in (89) are counterevidence to the Heim-Kennedy Constraint, which no longer needs to be relied upon. More importantly, the availability of the intermediate, scope reading is a specific prediction made by the proposed decompositional approach, which turns out to be supported.

The conclusion of section 4.7 is that the facts involving comparative QPs in subject position can be dealt with by the proposed system. The facts observed above demonstrate that the analysis based on Scope Economy is empirically superior to the one which adopts the Heim-Kennedy Constraint. As a consequence of investigating the decompositional approach, we have found an attested reading that has never been discussed in previous literature.
4.8 Further Issues

4.8.1 More on the Heim-Kennedy Constraint

In the previous sections, it was argued that the Heim-Kennedy Constraint is not an essential component in the explanation of the Comparative QP Scope Generalization, since Scope Economy and the Cross-Polar Anomaly Constraint conspire to yield the same result as the Heim-Kennedy Constraint does. Furthermore, we have discussed several cases that favor the proposed approach and constitute evidence against the Heim-Kennedy Constraint. At this point, it is reasonable to ask whether the proposed analysis can also account for Heim’s (2001a) original facts motivating the Heim-Kennedy Constraint, without resorting to it.

Heim (2001a) argues that there are two cases in which scope ambiguity does not emerge even though confounding factors, such as scopal commutativity and cross-polar anomaly, do not enter into the picture. The two cases are less-comparatives and exactly-differentials, the former of which has already been discussed in section 4.3.4. If the absence of scope ambiguity in these cases could also be dealt with by independently necessary principles, positing the Heim-Kennedy Constraint would no longer be justified. I take this to be a desirable outcome, since we have never had a good grasp on its nature. I will argue that the lack of scope ambiguity in less-comparatives turns out not to be evidence for the Heim-Kennedy Constraint, since the proposed analysis can explain it, in tandem with a fine-grained decompositional approach to less-comparatives. On the other hand, the scope freezing phenomenon in exactly-differentials does not straightforwardly follow from what I have claimed so far and remains a challenge to the attempt to do away with the Heim-Kennedy Constraint, a challenge which I hope to investigate in detail on another occasion.

Let me first discuss the lack of scope ambiguity in less-comparatives such as (32), which is repeated here as (93).
(93) a. Every girl is less tall than John is.
   b. (\forall > \text{less})
      \[
      \{\text{every girl}_1 \text{ is } [\text{AP } \text{DegP less than John is}]_2 [\text{AP } t_1 \text{ t}_2 \text{ tall}]}\]
      \[
      \forall x\{x \text{ is a girl } \rightarrow \{d: \text{ John is d-tall}\} \supset \{d: x \text{ is d-tall}\}\}
      \]
   c. (\text{less} > \forall)
      \[
      \{\text{any} \text{ is } [\text{AP } \text{DegP less than John is}]_2 [\text{AP } \text{every girl}_1 [\text{AP } t_1 \text{ t}_2 \text{ tall}]}\]
      \[
      \{d: \text{ John is d-tall}\} \supset \{d: \forall x\{x \text{ is a girl } \rightarrow x \text{ is d-tall}\}\}
      \]
      (i.e., The shortest girl is shorter than John.)

As we have already seen, if the DegP could take scope over the universal QP in subject position, as shown in (93)c, the less-comparative in (93)a would be true as long as the shortest girl is shorter than John. The unavailability of this reading was one of the original arguments for the Heim-Kennedy Constraint. In Heim (2001a), the less-comparative operator is assumed to be an unanalyzable operator which forms an antonymous pair with the comparative operator, since the directions of the ordering relation between the two sets of degrees are opposite. The lexical entry of the less-comparative operator is repeated here as (94).

(94) \[
[[\text{less}]] := \lambda D \in D_{(d, 0)}. [\lambda D' \in D_{(d, 0)}. D \supset D']
\]

As mentioned in 4.6.3, Hackl (2000) claims that the monotone decreasing comparative QP \textit{fewer than three books} is decomposed into the less-comparative operator in (94) and the quantificational DP \textit{many books}. On the basis of the frozen scope observed in the cases like (95), I have argued for the more fine-grained decompositional approach in which the less-comparative operator is decomposed into the comparative operator and \textit{little}.

(95) Every student in this class read fewer than three books.
    (\forall > \text{CQP}) \ast (\text{CQP} > \forall)
Exactly the same decompositional analysis of the *less*-comparative operator can be assumed in the *less*-comparative in (93), which allows us to offer a unified account of the absence of scope ambiguity in (93) and (95). In the more fine-grained decompositional approach, we can apply an optional SSO to the quantificational elements after we obtain the interpretable LF representation given in (96).

(96) \[ TP \text{ every girl}_1 \text{ is } \text{AP} \text{ [DegP er than John is]}_2 \text{[AP [DegP t_2 little]}_3 \text{[AP t_1 t_3 tall]]} \]

The problematic truth conditions in (93)c are derived from the scopal relation in which both the comparative operator and *little* take scope over the universal QP. This is blocked by the Heim-Kennedy Constraint within Heim’s framework. Notice, however, that the proposed decompositional approach to *less*-comparatives must yield the scopal relation in which the universal QP takes scope between the comparative operator and *little* before the unattested scopal relation is established. We have seen that universal QPs and the comparative operator are scopally commutative (Lemma II in (56)). Thus, Scope Economy prohibits the application of optional SSO and, hence, we cannot generate the unattested scopal relation from it. Consequently, we can deal with the absence of scope ambiguity in the *less*-comparative in (93)a without reliance on the Heim-Kennedy Constraint.

Let us now move to the scope facts in *exactly*-differentials. Heim (2001a) points out that if the comparative operator is modified with an *exactly*-differential expression, it is not scopally commutative with a universal QP. In addition, an LF representation that violates the Heim-Kennedy Constraint yields a well-defined meaning. For instance, the available interpretation of (97)a is expressed by narrow scope of the *exactly*-differential comparative operator, relative to the universal QP, as illustrated in (97)b. However, if the comparative operator were allowed to outscope the universal QP, as shown in (97)c, (97)a could be judged true in the situation where the shortest girl is exactly 4’1” tall and other girls are taller than that. This interpretation is uncontroversially not available for (97)a. Hence, this is an argument for the Heim-Kennedy Constraint.
(97) a. Every girl is exactly 1" taller than 4'.

(b) \((\forall>\text{exactly})\)

\[ [\text{TP every girl}_1 \text{ is } [\text{AP } \text{DeqP exactly 1" er than 4'}]_2 [\text{AP } t_1 \text{ t2 tall}]]] \]

\[ \forall x [\text{x is a girl } \rightarrow \{d: 0 \leq d \leq 4.1\} = \{d: x \text{ is } d\text{-tall}\}] \]

c. *(exactly>\forall)

\[ [\text{TP } \text{is } [\text{AP } \text{DeqP exactly 1" er than 4'}]_2 [\text{AP } \text{every girl}_1 [\text{AP } t_1 \text{ t2 tall}]]]] \]

\[ \{d: 0 \leq d \leq 4.1\} = \{d: \forall x [x \text{ is a girl } \rightarrow x \text{ is } d\text{-tall}]\} \]

(i.e., The shortest girl is exactly 4' 1" tall.)

For the approach that I am pursuing here, the lack of scope ambiguity in exactly-differentials remains crucial since it may interfere with the attempt to reduce the Heim-Kennedy Constraint to other principles.

I must defer addressing this issue until another occasion. However, I here suggest one promising approach to this challenging fact. Capitalizing on the fact that exactly is a focus-sensitive adverbial, I conjecture that the scope freezing effect in exactly-differentials would have something to do with focus, as pointed out by Sigrid Beck (personal communication). The fact that another focus-sensitive adverbial only seems to exhibit limited scopal possibilities is suggestive. In (98), the object only Mary cannot take wide scope relative to the subject universal QP.

(98) Every professor likes only Mary.

\((\forall>\text{only}) * (\text{only}>\forall)\)

As far as I can tell, the Heim-Kennedy Constraint does not enter into the picture in the scope freezing phenomenon in (98). Therefore, exploring scopal properties of focus-sensitive elements might provide a clue to account for the puzzling fact in exactly-differentials.

For now I am optimistic that the frozen scope in exactly-differentials can receive a principled explanation that does not rely on the Heim-Kennedy Constraint. In place of
completely rejecting my proposed approach to the facts accounted for by the Heim-Kennedy Constraint on the basis of the facts in *exactly*-differentials, I suggest that the proposed approach should be taken as a first step towards understanding the nature of the constraint, since we now have several new facts which counter-exemplify the Heim-Kennedy Constraint. These facts offer a good motivation to try to explore an account of the scope fact in *exactly*-differentials without recourse to the Heim-Kennedy Constraint.

To sum up this subsection, within the current framework the scope facts in *less*-comparatives can be straightforwardly dealt with, on a par with monotone decreasing comparative QPs. The scope properties of *exactly*-differentials will need to be subjected to closer scrutiny in the future.

### 4.8.2 More on Intervention Constraints

Similarly to the approach defended in this chapter, Stateva (2002, 2004) proposes that the Heim-Kennedy Constraint is reducible to an independently motivated constraint. In this subsection, I discuss some potential problems for Stateva’s proposal.

By looking at scopal properties of so-called comparative conditionals (Beck 1997), in which Stateva claims that the Heim-Kennedy Constraint is circumvented, Stateva proposes that the Heim-Kennedy Constraint applies only to covert DegP movement, as stated in (99).

(99) Stateva’s version of the Heim-Kennedy Constraint

A quantificational DP cannot intervene between a DegP and its trace *left behind by covert movement*.

(based on Stateva 2002, 2004)

Since a discussion of comparative conditionals would take us too far away from our concern, I explain Stateva’s version of the Heim-Kennedy Constraint by employing scopal properties of comparative QPs (together with some assumptions that I have adopted). By doing this, I do not think that we lose any part of Stateva’s core ideas.
Let us now consider the scopal asymmetry between object and subject comparative QPs, like in (100)a and (101)a. I have argued that this asymmetry is explained by making reference to the difference in semantic effect of the first application of an optional SSO between the two cases. On the other hand, it is not clear to me how the constraint in (99) helps us capture the scopal asymmetry between (100)a and (101)a. However, the following analysis might be one possible way to account for it within Stateva’s framework. The status of the trace \( t_I \) in (100)b and (101)b would play a crucial role in capturing the scopal asymmetry. The trace \( t_I \) in (100)b could count as a trace left by covert movement because it is left behind by covert QR of the object QP. Since the representation in (100)b is ruled out by (99), object wide scope is absent in (100)a. For the present purposes, let us assume that the trace \( t_I \) in (101)b can be regarded as a trace left behind by overt movement because the DegP movement is part of overt movement of the subject. Given this assumption, wide scope of the universal QP would be allowed in (101)a.

(100) a. Every student in this class read more than three books.
   \((∀>\text{CQP}) ∗ (\text{CQP} > ∀)\)
   b. \([\{\text{Degp er than three}\}, [\{\text{DP every student}\}], [\{\text{DP t1 many books}\}], [t_2 \text{ read } t_3]]\]

(101) a. More than three students in this class read every book.
   \((\text{CQP} > ∀) (∀ > \text{CQP})\)
   b. \([\{\text{Degp er than three}\}, [\{\text{DP every book}\}], [\{\text{DP t1 many students}\}], [t_3 \text{ read } t_2]]\]

Stateva claims that the constraint in (99) is one instance of the independently motivated intervention constraint proposed by Beck (1996a, b). (The analogy between the two constraints has already been hinted at in Heim 2001a.) The general form of Beck’s intervention constraint is presented in (102), and it aims at capturing the contrast between (103)a and (103)b, among many others. The in-situ wh-phrase undergoes covert movement across the QP in (103)a, which violates (102), but not in (103)b, which does not.
To consider Stateva’s approach as an alternative to the proposed analysis, I suggest that there are two potential problems that the former needs to address. First, Beck’s intervention constraint in the form of (102) is more general than Stateva’s version of the Heim-Kennedy constraint in (99), but it is still a descriptive generalization that needs to be explained. Beck (to appear) proposes a more principled account of the intervention constraint, in which a theory of focus interpretation derives its effects. One potential question would be whether the effects of Stateva’s intervention constraint in (99) can also be captured by Beck’s recent proposal.

The second potential problem is empirical. Beck (1996a, c) observes that the adverbial *often* cannot take narrow scope relative to the subject QP in (104)a. Based on (104)a and other facts, Beck argues that a subconstituent of an overtly moved *wh*-phrase cannot take scope in the scopal domain of other QPs. Assuming that the subconstituent taking narrower scope contains a trace of the *wh*-phrase, Beck claims that the lack of ambiguity in (104)a can be viewed as one of the effects of the intervention constraint.
(104)  a. Wie oft war niemand verfügbar?  
how often was nobody available
‘How often was nobody available?’
b. For which n: It was n times that nobody was available.
c. #For which n: There is nobody who was available n times.

(Beck 1996c:131)

As far as I can see, the process for deriving the unavailable reading in (104)c is very much analogous to the QL of the subject DP in (101)b. Thus, another potential problem would be how Stateva’s approach differentiates these two cases.

4.8.3 A Puzzle of QR

So far, I have only discussed the scopal interaction of quantificational DPs. Here, I present one puzzling property of QR, which is revealed by looking at the scopal interaction between quantificational DPs and scope bearing elements that are not DPs (e.g., negation, modals, and intensional predicates). We will see below several cases that suggest that when quantificational DPs undergo QR to take scope over some constituent that is not a DP, this QR seems to circumvent certain constraints on SSOs. To demonstrate this, let us first consider the sentence in (105) where the object comparative QP can outscope negation (object wide scope is more easily obtained if the comparative QP is partitive).

(105) John doesn’t speak more than three (of these) languages.
(not>CQP) (CQP>not)

Given the framework defended here, wide scope of the object comparative QP is unexpected. Within this framework, the whole comparative QP in object position first

---

33 Stateva (2002, 2004) assumes the same type of process in order to produce one reading of comparative conditionals.
undergoes QR targeting the vP, and is decomposed into the comparative operator and the quantificational DP in that position. Since these quantificational elements are still structurally below negation, the resulting representation still produces the narrow scope reading of the comparative QP. Its wide scope reading is yielded when both parts of the QP outscope negation. The first step towards deriving it is to move the DegP above negation by QR. This process results in the LF representation in (106)a, from which we obtain the formula in (106)b. As is easily verified, this results in presupposition failure:

\[
\begin{align*}
(106) & \quad \text{a. } [TP [\text{DegP er than three}]_2 [TP John_1 \text{ doesn't } [vP t_2 [vP [DP t_2 \text{ many languages}]]_3 \\
& \quad \quad \quad [vP t_1 \text{ speak } t_3]]]]] \\
& \quad \text{b. } \{d: 0< d < 3\} \subset \{d: \text{John doesn't speak } d\text{-many languages}\}
\end{align*}
\]

Since this stage of the derivation is not licensed by grammar, we are no longer able to apply the optional SSO necessary to get wide scope, which implies that wide scope of the comparative QP would not be allowed, contrary to fact.

An analogous problem arises in (107), which I have presented as one of the arguments for the decompositional approach to comparative QPs.

\[
\begin{align*}
(107) & \quad \text{At MIT one needs to come up with fewer than five brilliant ideas to get tenure.}
\end{align*}
\]

We have seen that the most salient reading of (107) is derived by the scope splitting configuration in which the predicate need is outscoped by the comparative operator and little. Due to Shortest Move, all of the generalized quantifiers that constitute the comparative QP must target the embedded vP position, which is still structurally lower than need, as illustrated in (108).

\[
\begin{align*}
(108) & \quad [TP one_1 [vP t_1 \text{ needs } PRO_1 \text{ to } [vP [\text{DegP er than five}]_2 [vP [\text{DegP t_2 little}]_3 [vP} \\
& \quad \quad \quad \quad \quad \quad \quad [vP t_3 \text{ many brilliant ideas}]]_4 [vP t_1 \text{ read } t_4]]]]]
\end{align*}
\]
However, the comparative QP cannot undergo an optional SSO to take scope over need because these two elements are scopally commutative under the assumption that need involves universal quantification over worlds (see Lemma II in (56)). Consequently, little cannot take wide scope relative to need, either, which fails to generate the scope splitting reading in (107).\(^{34}\)

This peculiar property of QR is quite independent of the issues of comparative QPs and the proposed framework. Indeed, the same fact can be demonstrated by employing ACD with a QP that would need to take scope over a constituent that is not a DP. QR for ACD resolution also seems to elude some constraints that regulate SSOs in the scope interaction between two quantificational DPs (von Fintel and Iatridou 2003, Wilder 1997). For instance, Wilder (1997) demonstrates that QPs can move across a tensed clause boundary in ACD, as in (109)a. To license ellipsis of EC, the object QP in the embedded clause needs to take scope over the matrix predicate, as illustrated in (109)b. However, this fact is unexpected in light of the generally observed clause-boundedness of SSOs:

(109)  

a. John said that you were on every committee that Bill did <say that you were on>.

(Wilder 1997:435)

b. \[\{\{\text{every committee }\lambda x. \text{[Bill did <say that you were on } x>]\} \quad \lambda y. \text{[John said that you were on } y]\}\]

\(^{34}\) A different problem arises in the derivation in (108) within the copy theory of movement framework. As discussed in Bhatt and Pancheva (2004, 2005) and Heim (2002), the comparative operator is not a conservative determiner. Thus, if the comparative operator with its restrictor constituent moves and leaves a copy, and Trace Conversion applies to the copy, the resulting structure produces an unattested interpretation (more precisely, a contradiction). Bhatt and Pancheva argue that this problem can be avoided by merging the comparative operator with its restrictor constituent counter-cyclically (see 3.3.3 for relevant discussion). However, under the assumption that we must apply as few operations as possible in order to derive an interpretable structure (see fn. 21 and 26), the comparative operator must merge with its restrictor constituent at the vP-position, as in (108). If the DegP, which consists of the comparative operator and its restrictor constituent, moved to a position above need, the resulting structure would produce a contradiction and it would not be licensed. Consequently, the comparative operator and, hence, the negative element little would not be able to take scope over need, contrary to fact. I do not have a solution to this problem.
While I need to explore this peculiar property of QR on another occasion, my hope is that a proper analysis of (109) could be carried over to the account of the puzzling comparative QP cases in (105) and (107) and would allow us to purse the proposed theory of SSOs.

4.9 Conclusion and Outlook

I have proposed that the fact that comparative QPs in object position must take narrow scope relative to QPs in subject position is explained by independently motivated properties of grammar, namely, the decompositional analysis of comparative QPs, Shortest Move, and the relevant constraints that rule out the intermediate scopal relation (Scope Economy and a principle that rules out cross-polar anomaly). In the decompositional approach, comparative QPs are composed of the comparative operator and the quantifier many NP. Shortest Move forces an intermediate step where the subject QP takes scope between the two decomposed QPs. However, intermediate scope of the subject QP is ruled out by other principles of grammar and, hence, the subject QP can only take widest scope, or surface scope. We need no stipulation to deal with this apparently peculiar property of comparative QPs. To the extent that my claims are correct, the scopal behavior of comparative QPs provides an additional argument for the core properties of SSOs (i.e., Shortest Move and Scope Economy). Based on the Comparative QP Scope Generalization, I have furthermore argued that an SSO proceeds in steps, which must each be licensed by principles of grammar, and that obligatory operations must be applied before optional SSOs.

The discussion of comparative QPs leads to draw the generalization in (110), which characterizes the derivational nature of SSOs.35

35 In formulating the MSG, I benefited from discussion with Danny Fox (personal communication).
(110) *The Multiple Scope Generalization (MSG)*

If $\alpha$, $\beta$, and $\gamma$ are quantificational elements such that

- a. $\alpha$ asymmetrically c-commands $\beta$ and $\gamma$ at S-structure,
- b. the scopal relation between $\beta$ and $\gamma$ is rigid at LF ($\gamma$ cannot outscope $\beta$),

then narrowest scope of $\alpha$ (i.e., $\beta > \gamma > \alpha$) is allowed only if intermediate scope of $\alpha$ (i.e., $\beta > \alpha > \gamma$) is licensed by grammar.

Although many facts should likely fall under the MSG, its detailed investigation will have to await another occasion.
Chapter 5

Conclusions and Outlook

At the outset of this dissertation, we saw that the scope of QPs is often not represented in surface structures: a constituent that is, on the surface, a sister to a QP in representations of this sort is not necessarily interpreted as its argument in the semantic component. We have discussed the facts that QPs can take either wider scope or narrower scope than their surface position. The dissertation first raised the question of what properties of grammar derive the effects of Scope Widening and Scope Narrowing. As a point of departure, two major approaches to these effects were taken up, namely, syntactic and semantic approaches.

Capitalizing on the property of syntactic approaches that QPs undergo QR to widen their scope, which ends up creating a variable binding dependency, Chapter 2 has provided evidence from ellipsis in favor of syntactic approaches. In order to set the stage for the supportive evidence, it was first established that when a particular type of variable binding dependencies called Re-binding is present in the structure, the biggest deletable constituent must be deleted. Building on Williams’ (2003) facts, it was argued that the maximization effect is observable in cases in which QPs exhibit the effects of Scope Widening, which suggests that the operation that is responsible for producing the effects of Scope Narrowing is accompanied by the formation of a Re-binding configuration. The maximization effect in this context is straightforwardly explained by syntactic approaches, in which the scope of QPs is structurally represented in Logical Forms. However, it is less clear how these facts would be captured in semantic approaches, in which the information concerning the scope of QPs is not included in syntactic representations,
the effects of Scope Widening are produced by applications of type shifting rules in the semantic component. Thus, Chapter 2 concluded that wide scope of the object QP in (1)a is derived from the LF representation of the following sort.

(1)  
   a. Some fifth year student read every book.

       \( \exists x (\forall y (y \in x)) \)  

   b. \([\text{every book} \lambda x \text{[some fifth year student read } x]\])

Chapter 3 focused on the issue of how the tail of a chain (i.e., \( t_1 \) in (1)b) is represented in Logical Forms. We saw various conceptual and empirical arguments in favor of the copy theory of movement, in which the tail of the chain contains the same lexical content as the moved constituent. As we have discussed, the copy theory of movement also has consequences for the treatment of the effects of Scope Narrowing. Despite advantages, some challenging facts to this theory have been pointed out, which has previously taken to be an argument that movement sometimes does leave a trace (contra the copy theory of movement). Chapter 3 developed a theory of movement and counter-cyclic merger which accounts for those facts and, yet, is compatible with the copy theory of movement. Extending the ideas about Late Merge proposed in the literature (Bhatt and Pancheva 2004, 2005, Fox 2002, Fox and Nissenbaum 1999, and Lebeaux 1988), I argued that determiners can undergo movement by themselves, and their restrictor constituents Wholesale Late Merge with the moved determiners. Fox’s (2002) Trace Conversion converts copies left behind by movement of a determiner by itself into syntactic objects that receive the same interpretation as that assigned to traces. I proposed that, together with the Case constraint on NPs, the Wholesale Late Merge analysis explains the challenging facts. Consequently, we can pursue the copy theory approach to the effects of Scope Narrowing as well as the simplest theory of movement. Given this framework, the LF representation in (1)b should be modified as follow:

(2)  
\([\text{every book} \lambda x \text{[some fifth year student read } x]\])
Chapter 4 explored the properties of the scope assignment mechanism by scrutinizing the peculiar scopal behavior of comparative QPs, namely, that comparative QPs in object position cannot take scope over a subject QP. One representative fact that illustrates this point is repeated here as (3).

(3) Some student read more than five books.

(∃>CQP) *(CQP>∃)

Adopting the decompositional approach to comparative QPs in which they are decomposed into the comparative operator -er and the quantifier many NP, I have argued that the locality constraint on SSOs requires the representation in (4)b to be established before the one in (4)c is generated, from which the wide scope reading of the comparative QP would be derived. However, the operation that produces (4)b is always ruled out either by Fox’s (2000) Scope Economy or by the semantic constraint on the comparative operator. Therefore, there is no way to derive the representation in (4)c, and this is the reason for the scope freezing phenomenon in (3).

(4) a. b. / c.

\[
\begin{array}{ccc}
\text{QP}_{\text{subj}} & \text{[er than 5]} & \text{QP}_{\text{subj}} \\
\text{[er than 5]} & \text{[er than 5]} & \text{[t, many books]} \\
\text{[t, many books]} & \text{[t, many books]} & \\
\end{array}
\]

I have argued that even though the representation in (4)c does not violate any principles of grammar, it is not derivable from the representation in (4)a if the intermediate stage in (4)b is illegitimate. Based on this property of the scope assignment mechanism, I have claimed that this mechanism has a derivational character.

I close this chapter with an outlook of two lines of research suggested by this dissertation. In Chapter 3, I explained the difference in Condition C bleeding between A-movement and A’-movement by the Wholesale Late Merge approach. A further topic that I would like to explore on another occasion is whether this approach could be extended to capture other facts that have been taken as evidence for the A/A’-distinction (e.g., crossover phenomena and parasitic gap licensing). Another topic that I also leave
to future research is pertinent to the nature of Scope Narrowing. In Chapter 3, I argued for the copy theory approach to Scope Narrowing, but I have developed the analysis of the scopal behavior of comparative QPs within a framework in which the effects of Scope Narrowing are derived by a movement operation, or QL. I would like to investigate in future research whether QL is necessary to produce the effects of Scope Narrowing within the copy theory of movement framework.
References


Heim, Irene. 2001b. A compositional analysis of less. Lecture notes, MIT.

Heim, Irene. 2002. Semantics of degree. Lecture notes, MIT.


[available at http://home.uchicago.edu/~ck0/docs/ace-revisited.pdf]


Richards, Norvin. 1997. What moves where when in which language? Doctoral
dissertation, MIT.

van Riemsdijk, Henk and Edwin Williams. 1981. NP-structure. The Linguistic Review
1:171-217.

Academic Press.

Romero, Maribel. 1998a. Focus and reconstruction effects in wh-phrases. Doctoral
dissertation, University of Massachusetts, Amherst.

Romero, Maribel. 1998b. The correlation between scope reconstruction and connectivity
effects. In Proceedings of WCCFL 16, ed. by Emily Curtis, James Lyle, and
Gabriel Webster, 351-365. Stanford, Calif.: CSLI Publications.

Rooth, Mats. 1992a. A theory of focus interpretation. Natural Language Semantics 1:75-
116.

Rooth, Mats. 1992b. Ellipsis redundancy and reduction redundancy. In Proceedings of
the Stuttgart Ellipsis Workshop, Report Series, Sprachtheoretische Grundlagen für
die Computerlinguistik, no 29, SFB 340, ed. by Steve Berman and Arild Hestvik, 1-
26. IBM Heidelberg.

Rooth, Mats. 1996. Focus. In The Handbook of Contemporary Semantic Theory, ed. by


Ross, John Robert. 1969. Guess who? In Papers from the Fifth Regional Meeting of the
Chicago Linguistic Society, ed. by Robert I. Binnick, Alice Davison, Georgia M.
Green, and Jerry L. Morgan, 252-286. Chicago: University of Chicago, Chicago
Linguistic Society.

dissertation, University of Massachusetts, Amherst.


Masako Hirotani, and Nancy Hall, 323-337. Amherst, Mass.: GLSA Publications.

SALT X, ed. by Brendan Jackson and Tanya Matthews, 167-184. Cornell, New
York: CLC Publications.

Tübingen.


209


Wold, Dag E. 1995. Focal structure and parallelism. Ms., MIT.
