An inflexible semantics for cross-categorial operators

by

Aron Hirsch

Submitted to the Department of Linguistics and Philosophy on September 8, 2017, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

Abstract

This thesis studies operators such as and only, which occur in a broad range of environments. And, for instance, appears between sentences, intransitive verbs, quantifiers, and so forth. One line of analysis assigns and/only a “cross-categorial” semantics flexible enough to compose with different arguments. This thesis challenges that view, pursuing the “Semantic Inflexibility Hypothesis” (SIH). Regardless of the surface string, and and only uniformly operate on a meaning characteristic of a sentence — a truth-value or proposition.

The thesis presents four case studies testing a central prediction of the SIH: that when and/only appear to compose with an expression having a non-sentential meaning, there must be covert syntax underlying to furnish an appropriate scope site. Most of the cases involve object DPs: (a) apparent object DP conjunction in basic sentences (John saw every student and every professor) and (b) in pseudo-clefts (What Obama approved was this bill and that bill), along with (c) only preceding an object DP (John learned only one language). The additional case study examines coordination of questions. Novel diagnostics reveal covert syntax in each case, reconciling the data with the SIH — and, in some cases, leading to a new perspective on the construction.

In addition to showing that a range of data may be parsed with covert syntax, I present reason to question whether cross-categorial meanings are available at all. Specifically, I point out that cross-categorial analyses over-generate. First: the mechanisms which give rise to cross-categorial meanings are too powerful, and predict more operators to be cross-categorial than actually are. Second, I show that if and itself were cross-categorial, unattested scope readings would derive. If there are no cross-categorial operators, the over-generation problems resolve without new constraints.

Thesis supervisor: Kai von Fintel
Title: Andrew W. Mellon Professor of Linguistics

Thesis supervisor: Danny Fox
Title: Anshen-Chomsky Professor of Language and Thought

Thesis supervisor: Irene Heim
Title: Professor of Linguistics
Acknowledgments

I wrote this thesis now, since I have been a graduate student for 5 years, and I cannot be one any longer. I see this as a status report, and I hope my readers will see it that way also, and forgive its deficiencies in content and form. Whatever ideas in this thesis stand owe a deep debt to those who have generously discussed this work with me. Let me offer a note of thanks, albeit insufficient, to them, and to the many others who have enriched my life professionally and personally.

My committee: Kai von Fintel, Danny Fox, Irene Heim, David Pesetsky, and Roger Schwarzschild. As everyone already knows, they are uncommon as linguists and as people. They have modeled a level of intellectual curiosity and rigor to which I aspire, and have been generous with their ideas and support to a level I don’t deserve. I thank them all for everything.

I also owe a great debt to Martin Hackl and Sylvain Bromberger. Although he was not an official committee member due to his sabbatical, meetings with Martin would often last multiple hours, and it was a pleasure collaborating with him on experimental work. Sylvain has been a debating partner on a wide range of topics. We often approach questions from different perspectives, and the interchange has been very rewarding. For guidance on different projects, my thanks to: Sabine Iatridou, Edward Flemming, Donca Steriade, Michael Kenstowicz, Adam Albright, and Michel DeGraff. For inspiring classes, I thank Gennaro Chierchia, Norvin Richards, Justin Khoo, Ken Wexler, and Noam Chomsky, as well as everyone above. For somehow saving me from my deep administrative ineptitude, I am grateful to Jen Purdy, Matt Sikorski, Mary Grenham, and Chrissy Graham.

I feel incredibly lucky to have spent these years at MIT. For many conversations, linguistic and not, thanks to all the students and visitors in the department, among them: Sam Alxatib, Nico Baier, Isa Bayirli, Itai Bassi, Ruth Brillman, Anthony Brohan, Brian Buccola, Cassandra Chapman, Keny Chatain, Paul Crowley, Ömer Demirok, Mitch Erlewine, Isaac Gould, Verena Hehl, Natasa Ivlieva, Roni Katzir, Hadas Kotek, Loes Koring, Nick Longenbaugh, Lilla Magyar, Paul Marty, Matt Mandelkern, Sophie Moracchini, Mia Nussbaum, Erin Olson, Sasha Podobryaev, Mitya Privoznov, Milena Sisovics, Juliet Stanton, Sam Steddy, Ben Storme, Ayaka Sugawara, Abdul-Razak Sulemana, Amanda Swenson, Chris O’Brien, Despina Oikonomou, Wataru Uegaki, and Danfeng Wu.

For their close friendship, my thanks to: Athulya Aravind, Kenyon Branan, Ted Levin, Daniel Margulis, Ezer Rasin, Coppe van Urk, Michelle Yuan, and Sam Zukoff. From meals at Veggie Galaxy, My Thai, and Lone Star to abridged seders with vegan latkes to donut runs to frequent nights out, the times I had with them were some of the best of graduate school.

Before coming to MIT, I was at McGill University, where I had the good fortune of meeting Michael Wagner. Over the years, Michael transitioned from my undergraduate advisor into my most frequent collaborator and a close friend. I also thank Luis Alonso-Ovalle, Jessica Coon, Meghan Clayards, Brendan Gillon, Yosi Grodzinsky, Máire Noonan, Bernhard Schwarz, Junko Shimoyama, and Lydia White. I am thrilled that I have the opportunity to return to McGill next year.

Many people have discussed this project with me. My thanks to: Luis Alonso-Ovalle, Athulya Aravind, Itai Bassi, Maria Biezma, Kenyon Branan, Seth Cable, Luka Crnić, Amy Rose Deal, Paul Hagstrom, Andreas Haida, Heidi Harley, Jesse Harris, Roni Katzir, Ed Keenan, Angelika Kratzer, Daniel Margulis, Neil Myler, Jessica Rett, Barbara Partee, Yihui Quek (thanks also for joint work
on focus!), Ezer Rasin, Maribel Romero, Bernhard Schwarz, Yael Sharvit, Dominique Sportiche, Anna Szabo, Michal Wagner, Susi Wurmbrand, and Michelle Yuan. I especially thank Lucas Champollion for a conversation in Amsterdam which first got me thinking about pseudo-clefts, and Barry Schein for a long meeting on a visit to Cambridge. I apologize to those I am forgetting.

Segments of this work have been presented at a variety of venues, including: NELS 46 (Chapter 3), NELS 47 (Chapter 6), Sinn und Bedeutung 21 (Chapter 5), and the 2017 LSA Meeting (Chapter 4), as well as at colloquium talks at UCLA and McGill (Chapters 2, 3, and 6), and locally at MIT’s LF Reading Group and Ling-Lunch. I thank those audiences for generous feedback.

It is remarkable how collegial the entire community is. I want to thank, in addition to those above, many people for their friendship, feedback, and/or collaboration over the years: Karlos Arregi, Akiva Bacovcin, Andras Bárán, Jonathan Bobaljik, Thomas Borer, Rajesh Bhatt, Kajsa Djarv, Ilaria Frana, Stefan Keine, Ezra Keshet, Kyle Johnson, Anton Karl Ingason, Jon Ander Mendia, Dan Goodhue, Patrick Grosz, Prity Patel-Grosz, Jérémy Pasquereau, Ethan Poole, Florian Schwarz, Ulli Steindl, Yasu Sudo, Yimei Xiang, Kristine Yu, and Jérémy Zehr. Apologies again for my faulty memory.

And, finally, my deepest and most heartfelt thanks to my family: my brother, Moshe Hirsch; my unofficial brother, Brandon Trask; my grandmother, Mindel Olenick; and my late grandparents, Esther and Louis Hirsch and Tom Olenick. For my parents, Debby and Brian Hirsch, I’m not sure how to express my love and gratitude. I can only hope they know. None of the wonderful things that have happened to me would have been possible without them.
## 4.4 Class 2 operators ........................................... 55
## 4.5 The resultant picture .................................... 62
## 4.6 The perspective of the SIH ............................... 64
### 5 Conclusion .................................................. 64

### Chapter 3: A Case for conjunction reduction ................. 67

#### 1 Introduction ............................................. 67

1.1 Preview of main proposal ................................... 67
  1.1.1 Review of empirical evidence for vP-conjunction ........ 68
  1.1.2 Preview of evidence against DP conjunction with [[and3]] .. 69
1.2 A Note on sum formation .................................... 69
1.3 Roadmap from here ......................................... 71

#### 2 Reviving conjunction reduction ........................... 71

2.1 Step 1: vPs are type t .................................... 73
2.2 Step 2: CR as vP conjunction .............................. 75
  2.2.1 Gapping ........................................... 75
  2.2.2 From gapping to CR ................................ 79
  2.2.3 Antecedent proposals ................................ 80
2.3 Summary and next steps ................................... 81

#### 3 Case 1: adverbs ........................................ 82

#### 4 Case 2: VP ellipsis ....................................... 84

4.1 VP ellipsis supportive of CR ................................ 85
4.2 The DP analysis cannot account for observed VP ellipsis ... 86
  4.2.2 The Adverbial clause adjoins to the DP in second conjunct 75
4.3 Local summary ............................................. 89

#### 5 Case 3: split scope ....................................... 89

5.1 The Split scope signature .................................. 89
5.2 The CR analysis ........................................... 91
5.3 The DP analysis is insufficient ......................... 92
5.4 Ruling out alternatives ................................... 93
  5.4.1 Alternative 1: type-lifting the quantifiers ............ 93
  5.4.2 Alternative 2: sums and homogeneity ................. 95
5.5 Summary and next steps ................................ 98

#### 6 Case 4: a missing scope reading ......................... 98

6.1 A Missing scope reading ................................... 98
6.2 The DP analysis over-generates ......................... 99
6.3 The CR analysis ........................................... 101
  6.3.1 Aside: a comment on the hydra parse ................ 102
Chapter 5: Coordinating questions

1 Introduction
   1.1 Part I: pursuing the SIH .................................................. 179
   1.2 Part II: dissecting question disjunction .................................. 180
   1.3 Roadmap for the chapter .......................................................... 181

2 Pursuing the SIH
   2.1 Question conjunction ............................................................ 181
   2.2 Question disjunction ............................................................. 184
   2.3 Concluding part I ................................................................. 189

3 Two profiles of question disjunction
   3.1 Profile 1: only one question answerable ...................................... 190
   3.2 Profile 2: both questions answerable ............................................ 191
   3.3 Local summary ............................................................................. 192

4 Question disjunctions are mention-some
   4.1 Mention-all questions .............................................................. 192
   4.2 Question disjunctions are not mention-all ...................................... 193
   4.3 Mention-some questions ............................................................ 195

5 The Analysis ................................................................................. 196
# 5.1 Analyzing simple questions

- 5.1.1 Dayal’s ANS ........................................................................ 198
- 5.1.2 Modifying ANS (Fox 2013) .................................................. 201
- 5.1.3 Illustrating mention-all ....................................................... 202
- 5.1.4 Illustrating mention-some .................................................... 204
- 5.1.5 Constraints on mention-some .............................................. 205
- 5.1.6 Local summary .............................................................. 206

# 5.2 Analyzing the disjunctions

- 5.2.1 Profile 2 disjunctions ..................................................... 206
- 5.2.2 Profile 1 disjunctions ..................................................... 207
- 5.2.3 Disjoining MS questions .................................................. 208
- 5.2.4 Local summary .............................................................. 209

# 6 Taking into account the pragmatics

- 6.1 The Pragmatics of mention-some ........................................ 210
- 6.2 The Pragmatics helps explain further data ................................ 211
- 6.3 Why disjoining questions isn’t fully productive ...................... 213
- 6.4 Local summary .............................................................. 214

# 7 Alternative analyses

- 7.1 Alternatives for conjunction ............................................. 214
  - 7.1.1 Partition semantics for questions ........................................ 214
  - 7.1.2 Pointwise Functional Application ..................................... 217
- 7.2 Alternative for disjunction .................................................. 221
- 7.3 Local summary .............................................................. 224

# 8 Conclusion

# 9 A program for further inquiry

- 9.1 On verb conjunction ..................................................... 224
- 9.2 On subject DP conjunction .................................................. 225
- 9.3 On conjunction within DPs .................................................. 226
- 9.4 Collective prediction .............................................................. 230
- 9.5 Local summary .............................................................. 232

---

Chapter 6: Focus operators

# 1 Introduction

- 1.1 The basic meaning of *only* .................................................. 234
- 1.2 Defining [[only$_3$]] ..................................................... 236
- 1.3 Preview of main claims .............................................................. 238
- 1.4 Roadmap for the chapter ..................................................... 239
2 Test 1: split scope ........................................ 239
  2.1 The Split scope signature .............................. 239
  2.2 Insufficiency of [[only,]] .............................. 242
  2.3 Type-lifting the quantifier? ............................ 243
  2.4 Concluding split scope ................................. 245
3 Test 2: VP ellipsis ........................................ 245
  3.1 Insufficiency of type-ambiguity ....................... 246
  3.2 Concluding VP-ellipsis ................................. 247
4 Proposal .................................................. 247
  4.1 Syntax: bipartite ...................................... 249
  4.2 Mapping to phonology: variable ....................... 250
  4.3 Mapping to semantics: constant ....................... 251
  4.4 Explaining Tests 1 and 2 ............................. 251
  4.5 Local summary ........................................ 253
  4.6 Excursus: negative indefinites ........................ 254
  4.7 Next steps ............................................ 256
5 Scope of pre-vP only ..................................... 256
  5.1 Where scope is frozen .................................. 257
  5.2 Where scope is free ................................... 259
  5.3 Comparison with type-ambiguity ....................... 261
  5.4 Local summary ........................................ 263
6 Cross-linguistic evidence ................................ 263
  6.1 Assumptions 1 and 3 ................................... 263
    6.1.1 Dutch ............................................ 263
    6.1.2 Vietnamese ....................................... 265
    6.1.3 Local summary .................................... 265
  6.2 Assumption 2 .......................................... 266
  6.3 A Semantic precedent .................................. 268
  6.4 Concluding cross-linguistic evidence ................. 269
7 Comparison with alternatives ............................. 270
  7.1 Two-place only ....................................... 270
    7.1.1 Illustrating the approach .......................... 270
    7.1.2 Problem 1: pre-DP only ............................. 272
    7.1.3 Problem 2: pre-vP only ............................. 273
    7.1.4 Problem 3: scope of pre-vP only .................. 274
    7.1.5 Local summary ..................................... 275
  7.2 Decomposing only (von Fintel & Iatridou 2007) ........... 275
  7.3 Section summary ...................................... 279
Chapter 1
The Semantic Inflexibility Hypothesis

1 Introduction

For many predicates of natural language, their semantics and syntax are tightly circumscribed, and relate to one another. The verb *devour*, for instance, has as its meaning a relation between two entities, one doing the devouring and another being devoured. In semantic terms, *devour* requires two entities as its arguments. The semantic argument structure of *devour* tracks its syntactic distribution. *Devour* must occur in a syntactic frame with two DPs, which are the kind of syntactic object which can denote entities. Deviating from that frame results in ungrammaticality.

\[(1)\]
Circumscribed distribution of *devour*

\begin{itemize}
\item a. John devoured the apple.
\item b. *John devoured.
\item c. *John devoured the apple the park.
\item d. *That Bill went to the store devoured.
\end{itemize}

This thesis is about operators whose semantics and syntax seem to mismatch. The conjunction *and* is a case in point. A natural hypothesis would take *and* to be interpreted like the \( \wedge \) connective of propositional logic. \( \wedge \) occurs in formulae \((p \wedge q)\), where \( p \) and \( q \) are associated with truth-values, and says that \((p \wedge q)\) is true iff \( p \) is true and \( q \) is true. By hypothesis, *and*, likewise, operates on two truth-values and says that they are both true. Given this meaning, *and* shows an unexpectedly broad distribution. While it conjoins sentences, which denote truth-values, in (2-a), it seems to occur in a wide variety of different syntactic frames — it appears to conjoin intransitive verbs in (2-b), quantificational DPs in (2-c), and transitive verbs in (2-d).

\[(2)\]
Broad distribution of *and*

\begin{itemize}
\item a. John laughed and Mary danced.
\item b. John laughed and danced.
\item c. John saw every student and every professor.
\item d. John hugged and petted the dog.
\end{itemize}

\[^{1}\text{There are, of course, other respects in which natural language *and* appears not to act like logical conjunction. For instance, while \( \wedge \) is commutative \( ([\text{and}](p)(q) \Leftrightarrow [\text{and}](q)(p)) \), natural language *and* shows some asymmetries dependent on order (e.g. *The sniper shot him and he died* \( \neq *He died and the sniper shot him*; Bjorkman 2013). One approach is to attribute such effects to the pragmatics (e.g. Grice 1975, Schmerling 1975, Carston 1993, among others; but, see Txurruka 2003 and Bjorkman 2013 for quite different views). I will assume the pragmatic approach and set this issue aside.}\]
A similar puzzle arises for *only*. Like *and*, *only* deals in truth and falsity. In (3-a), *only* says that *John learned one language* is true, while alternative sentences *John learned n languages* are false, for different *n*. As a result, we might expect *only* to show a syntactic distribution limited to adjoining to constituents associated with truth-values, like sentences. Yet, *only* again exhibits a surprisingly broad distribution, as the data in (3-b) and (3-c) show.

(3) **Broad distribution of only**
   a. It's only the case that John learned one language.
   b. John only learned one language.
   c. John learned only one language.

To adopt a term neutral to analysis, I will call items like *and* and *only* "Free Operators" (hence, 'FOs'), in reference to the distributional freedom they seem to show. The overarching question of the thesis, then, is: how are the syntax and semantics of FOs reconciled?

This question has received considerable attention in the development of formal semantics. FOs have generally been taken as evidence that the semantics must incorporate mechanisms to allow one operator to *flexibly* compose with different kinds of arguments. In this thesis, however, I pursue a different hypothesis: that FOs always compose with sentence meanings, with their broad distribution an illusion, created by a richer underlying syntax not fully expressed in the overt phonology. The semantics, despite appearances, is *inflexible*.

In the remainder of the chapter, I spell out the problem in a more precise way within a particular theory of natural language semantics (Section 1.1), and then sketch the traditional solution based on semantic flexibility (Section 1.2), as well as my own hypothesis and its precedents in the literature (Section 1.3). Finally, I provide an overview for the remainder of the thesis (Section 2).

### 1.1 Developing the problem

To make the problem precise, let us entertain a heavily constrained theory of how natural language assigns meaning to basic expressions. The foundational work of Richard Montague presented a circumscribed system for building lexical entries, based on a notion of *semantic type* (e.g. Montague 1973). I assume three primitive types: type *e* (corresponding to entities), type *t* (for truth-values), and type *s* (for possible worlds). A complex type \(<\sigma, \tau>\) corresponds to a function from elements of type \(\sigma\) to elements of type \(\tau\). The set of meanings of some type \(\sigma\) are indicated as ‘\(D_\sigma\)’.

(4) **Inventory of semantic types**
   a. *e*, *t*, and *s* are semantic types
   b. If \(\sigma\) is a semantic type and \(\tau\) is a semantic type, \(<\sigma, \tau>\) is a semantic type.
   c. Nothing else is a semantic type.

What are the semantic types associated with some common expressions? A referential DP may be assigned a meaning of type *e*, while a sentence *my* be assigned a meaning of type *t*. At this point, I represent possible worlds as a parameter on the denotation.
Sample primitive meanings

a. \([\text{John}]_w = \text{John}\)  
   \(\text{(type e)}\)

b. \([\text{John saw the dog}]_w = 1\) if \(\text{John saw the dog}\) in \(w\), otherwise \(0\)  
   \(\text{(type t)}\)

The expressions in (6) are associated with complex types. Intransitive verbs denote functions of type \(<e,t>\) (i.e. one-place predicates of individuals), quantificational DPs denote functions of type \(<et,t>\) (i.e. a second-order predicate), and transitive verbs denote functions of type \(<e,et>\) (i.e. two-place predicates of individuals).

Sample complex meanings

a. \([\text{laughed}]_w = \lambda x . x \text{ laughed in } w\)  
   \(\text{(type } <e,t>)\)

b. \([\text{every student}]_w = \lambda f_{et} . \forall x \ [x \text{ is a student in } w \rightarrow x \text{ laughed in } w]\)  
   \(\text{(type } <et,t>)\)

c. \([\text{devour}]_w = \lambda x . \lambda y . \lambda w . y \text{ devoured } x \text{ in } w\)  
   \(\text{(type } <e,et>)\)

Depending on its type, a function will require a particular number and kind of arguments. To create a sentence meaning, an \(<e,t>\) predicate requires one type e argument, an \(<et,t>\) predicate requires one type \(<e,t>\) argument, and an \(<e,et>\) predicate requires two type e arguments. As a thought experiment, I want to make the assumption that the expressions of natural language under consideration are \textit{unambiguous}: each has \textit{one} meaning and is thus associated with \textit{one} semantic type imposing \textit{one} particular semantic argument structure.

Assumption for thought experiment

Expressions are unambiguous, associated with a single semantic type.

If expressions are singly typed, the lexical semantics of an expression tightly constrains its syntactic distribution. Why is this? The syntactic structure is output to a system of semantic interpretation, which determines a meaning compositionally from the syntax. Following Heim & Kratzer (1998), one compositional principle is Functional Application:

Functional Application (‘FA’)

For any world \(w\), if \(\alpha\) is a branching node with daughters \(\beta\) and \(\gamma\) such that \(\beta \in D_{\sigma}\) and \(\gamma \in D_{<\sigma,t>}\), then \([\alpha]_w = [\gamma]_w([\beta]_w)\).

Through FA, the semantic arguments of a predicate are saturated step-by-step with the syntactic structure. Recall, for instance, the simple sentence in (1-a) \((\text{John devoured the apple})\). The VP has the syntax in (9), and meets the structural description for FA: the VP is a branching node, with its daughters the transitive verb \(\text{devour}\) (type \(<e,et>)\) and the DP \(\text{the apple}\) (type \(e\)). FA computes the denotation for the VP by saturating the first argument of \([\text{devour}]\) with the apple.
Interpreting the VP

Syntax:

\[\text{VP} = \text{devoured } \text{the apple} \]

FA:

\[\text{[VP]}^w = \text{[devoured]}^w(\text{[the apple]}^w) = \lambda x. \ x \ \text{devoured the apple} \text{ in } w\]

The meaning for the TP in (10) is computed through a second step of FA. The TP's daughters are the VP, and the subject DP John. Because the first argument of \([\text{devour}]\) has already been saturated within the VP, \([\text{VP}]\) is of type \(<e,t>\), while \([\text{John}]\) is of type \(e\). FA saturates the open argument in \([\text{VP}]\), i.e. the second argument of \([\text{devour}]\), with John.

Interpreting the TP

Syntax:

\[\text{TP} = \text{devoured the apple } \text{in } w\]

FA:

\[\text{[TP]}^w = \text{[VP]}^w(\text{[John]}^w) = 1 \text{ iff } \text{John devoured the apple in } w\]

Given FA, an expression must occur in a syntactic frame which furnishes the right number and kind of constituents to provide all and only the semantic arguments required by its semantic type. If every expression is singly typed, every expression should be confined to a very particular syntactic frame. An intransitive verb must occur in a frame with one DP, a quantifier must occur in a frame with a constituent having a type \(<e,t>\) meaning such as the VP, and a transitive verb must occur in a frame with two DPs, and so forth. For predicates such as \(\text{devour}\) (type \(<e,et>\)), this is exactly right:

Circumscribed distribution of \(\text{devour}\)

a. John devoured the apple.

b. *John devoured.

c. *John devoured the apple the park.

d. *That Bill went to the store devoured.

In (11-b), there are not enough arguments to derive a sentence meaning. John could saturate one argument of \([\text{devour}]\) to derive an \(<e,t>\) meaning, but that's it. In (11-c), there are too many arguments. John and the apple saturate the argument slots of \([\text{devour}]\), and there is no way for the park to compose. In (11-d), there is the wrong kind of argument. The CP that Bill went to the store has a type \(t\) meaning, so cannot saturate the second argument of \([\text{devour}]\). (11-b) has the wrong kind of semantic value, and (11-c) and (11-d) have no defined semantic value at all. The semantics of \(\text{devour}\) makes powerful, correct predictions about its syntax.
This picture makes a strong prediction about what sort of expression cannot not exist: there should be no expression \textit{blick} which has a single meaning, but occurs in different syntactic frames. To occur in the frame in (12-a), for instance, \textit{blick} must have a type beginning in \textit{t}, so as to compose with the TP. Conversely, to occur in the frame in (12-b), \textit{blick} must have a type beginning in \textit{<et,t>}, so as to compose with the quantificational DP. With expressions singly typed, \textit{blick} cannot exist.

(12) \textbf{Prediction: no blick}

a. \([TP \textit{blick} [TP \text{John saw every student}]]\)

b. \([TP \text{John saw} [DP \text{blick [every student]}]]\)

FOs such as \textit{and}, however, seem to show exactly the distribution of \textit{blick}. We can make the problem precise now that we have semantic types in hand. If \textit{and} operates on truth-values with a meaning similar to \&, \textit{and} would be of type \textit{<t,<t,t>>}, with this lexical entry:

(13) \textbf{Lexical entry for and}

\([\textit{and}] = \lambda p_t \cdot \lambda q_t \cdot p = q = 1 \quad (\text{type } \textit{<t,<t,t>>})\)

The semantics, then, makes the very strong syntactic prediction that \textit{and} should always occur in construction with two constituents interpreted with a type \textit{t} meaning. This is the case in (14): the TPs are of type \textit{t}, and each saturates one argument of \([\textit{and}].\)

(14) \textbf{Predicted syntactic frame}

\begin{center}
\begin{tikzpicture}
    \node (root) {\&P} child {node (TP1) {TP} child {node (John) {\text{John laughed} \hspace{1cm} (t) \hspace{1cm} (t)}} child {node (and) {\& \hspace{1cm} (\textit{<t,<t,t>>}) \hspace{1cm} \text{and} \hspace{1cm} TP2} child {node (TP2) {TP} child {node (Mary) {\text{Mary danced} \hspace{1cm} (t) \hspace{1cm} (t)}}}}}
    \node (FA) at (root -| root) {$\text{FA: } [[\&P]]'' = [[\textit{and}]]([[TP_1]])''([[TP_2]])''$
    $= 1 \text{ iff } \text{John laughed in } w \hspace{2cm} \& \hspace{1cm} \text{Mary danced in } w$
    $\wedge \text{Mary danced}$
\end{tikzpicture}
\end{center}

The data in (15) (repeated), however, seem to falsify the overall prediction. If these examples involve the most obvious syntax where \textit{and} directly conjoins the expressions that flank it, they should all suffer from a type-mismatch, since intransitive verbs, transitive verbs, and quantifiers are not of type \textit{t}. These data should have the same ungrammatical status as (11-d), where \textit{devoured} occurred with the wrong kind of argument — contrary to fact.

(15) \textbf{Apparent fact: and is blick}

a. \text{John [laughed and danced].}

b. \text{John [hugged and petted] the dog.}

c. \text{John saw [every student and every professor].}
Hence, the problem. Given the theory of semantic types, if FOs have a single meaning, they require a particular kind of argument and their broad distribution is then unexpected. Something needs to change. The problem is based on both semantic and syntactic assumptions. Semantically, it is assumed that lexical entries accord with Montague’s type system and that FOs are associated with a unique type. Syntactically, it is assumed that and directly conjoins the constituents it appears to at the surface. In principle, any of these aspects of the system could be the incorrect one.

1.2 A widespread response

The most widespread response is to accept the syntax and introduce flexibility into the semantics. The nature of the flexibility takes a different form in different theories. I will discuss the different possibilities in depth in Chapter 2, but offer an initial glimpse here. I illustrate for now with and, but other FOs such as only can be captured in similar way.

The first route is to add expressive power to the lexicon. In one way or another, natural language builds meaning from a richer toolkit than the one sketched in Section 1.1 — one able to state single lexical entries with flexible argument structure. Keenan & Faltz (1978, 1985) draw on the tools of Boolean algebra, and model and as the Boolean ‘meet’ operator which can operate on truth-values, as well as other semantic objects.

A second route is to posit that and is ambiguous between a family of meanings of distinct semantic types. Each meaning has a specific type and thus imposes a specific argument structure, but ambiguity creates the overall appearance of flexibility. In addition to type <t,<t,t>>, and can be interpreted as type <et,<et,et>> (allowing for intransitive verb conjunction), type <ett,<ett,ett>> (quantifier conjunction), or type <et,<et,et>> (transitive verb conjunction), and so forth.

Although pervasive accidental homophony lacks plausibility, Partee & Rooth (1983) note that different meanings may be systematically related to one another (cf. von Stechow 1974, Gazdar 1980). In particular, higher-type meanings can be defined based on the <t,<t,t>> meaning:

(16) Higher-type variants of and

a. $\langle and_2 \rangle = \lambda f_{et} . \lambda g_{et} . \lambda x . \langle \text{and} \rangle (f(x))(g(x)) \quad \text{(type <et,<et,et>>)}$
   ('x is both an f and a g')

b. $\langle and_3 \rangle = \lambda f_{et} . \lambda g_{et} . \lambda f_{et} . \langle \text{and} \rangle (f(f))(g(f)) \quad \text{(type <ett,<ett,ett>>)}$
   ('f is both an F and a G')

c. $\langle and_4 \rangle = \lambda f_{et} . \lambda g_{et} . \lambda x . \lambda y . \langle \text{and} \rangle (f(x)(y))(g(x)(y)) \quad \text{(type <et,<et,et,et>>)}$
   ('x f-ed y and x g-ed y')

Other work has proposed more general type-shifting operations to derive the higher-type meanings (notably, the Geach Rule of Jacobson 1999, 2014). In one conception and is stored in the lexicon just with the meaning $\langle \text{and} \rangle$, while other meanings are derived on an as-need basis by rule.

Overall, then, the most widespread treatment of FOs is, one way or another, to systematically imbue them with a flexible argument structure. To facilitate exposition throughout the thesis, I will frame discussion in terms of one conception of how flexibility comes about: the last approach in terms
of a type-shifting operation. The major claims and arguments of the dissertation could, however, be adapted to other conceptions just as well.

1.3 The Semantic Inflexibility Hypothesis

This thesis challenges the semantic flexibility approach by putting forward what I call the Semantic Inflexibility Hypothesis, stated in (17):

\[(17)\] Semantic Inflexibility Hypothesis (‘SIH’)
Natural language does not have a type-shifting mechanism applicable to FOs.

The intent of the SIH is to return to the constrained theory of meaning presented in Section 1.1, without amendment or augmentation. Lexical entries are built within Montague’s type-system, so semantic type narrowly circumscribes syntax. Moreover, there is no mechanism to systematically derive multiple meanings for one operator. Assuming that positing pervasive accidental homophony for FOs is rejected as implausible,\(^2\) a key corollary of the SIH is that FOs have a uniform meaning, with a unique semantic type across their distribution. Despite what appears to be the case at the surface, FOs always compose with arguments of the same type. Specifically, I pursue the idea that \textit{and}, as well as \textit{only}, uniformly composes with sentence meanings. \textit{And} is interpreted as the type \(<t, <t, t>>\) \textit{[and]} operator defined above — and does not have type-lifted variants.

\[(18)\] Corollary for \textit{and} and \textit{only}
These FOs uniformly apply to sentence meanings.

The SIH makes the prediction that every time an FO appears to operate on something other than a type t meaning that there is a richer underlying syntax, obscured from view by the phonology. Covert syntax must create an appropriate type t scope site for the FO. Much of the thesis is dedicated to testing that prediction in a range of cases, and I flag it below. The viability of the SIH thus relies on the strength of independent evidence for covert syntax.

\(^2\)The SIH still does not rule out that a language could lexicalize one or more of the higher-type meanings in (16), nor does it rule out that such a morpheme might occasionally be homophonous with the one encoding \textit{[and]}. English does not realize any of the higher-type meanings as a phonologically distinct morpheme, and I would expect that accidental homophony with \textit{[and]} should be quite rare cross-linguistically. Thus, I will set aside the possibility that \textit{and} is accidentally homophonous at all, and will vindicate that assumption later on by directly arguing that \textit{and} does not encode at least certain of the higher-type meanings. In the concluding chapter, I will entertain possible extensions of the SIH which do prevent higher type meanings from being lexicalized. Thanks to David Pesetsky, Irene Heim, and Bruce Hayes for discussion.

21
Predicting covert syntax

When an FO appears not to operate on a type $t$ meaning, there must be a richer underlying syntax which furnishes a type $t$ scope site.

From this perspective, understanding FOs requires careful consideration of modules of grammar beyond the semantics. We must understand the syntax of FOs, and we must understand how the syntax non-trivially maps to phonology to obscure the presence of sentential nodes in the overt string. This thesis aims to look at FOs in this holistic way, and argues for the SIH with an appropriate mix of novel semantic, syntactic, and phonological evidence.

The idea of positing covert syntax with FOs is certainly not novel, but rather goes back to the earliest literature, pre-dating the semantic flexibility approach. Indeed, it was proposed in the 1970s and before that and should be analyzed with Conjunction Reduction, which was taken to mean that and always underlyingly conjoins full sentences. Partee (1970) provided clear evidence against that idea, and the move towards covert syntax was largely abandoned. In the decades since, however, our understanding of the syntactic module and its interface with semantics has independently undergone rapid development. As we will see, this has opened up new possibilities for covert structure with sentential meaning that were not available when the ambiguity theory was proposed. These new mechanisms will avoid earlier empirical pitfalls and, I believe, will make re-introducing the covert syntax theory both possible and fruitful.

Still, creating a complete argument for an SIH-based treatment of FOs is far from a trivial matter. Such an argument would have two parts. First: we must show that in every single case where an FO seems to require a different meaning that there is covert syntax, including type $t$ node. This would alleviate the need to posit flexible argument structure. At least in some cases, however, a parse with covert syntax could still co-exist with an additional parse in which the covert syntax is absent and the FO is higher-type. The second feature of a complete argument is, therefore, direct evidence that FOs cannot compose with anything other than type $t$ meanings.

Both of these are difficult and the range of relevant data is vast. As a result, this thesis will not attempt a complete argument for the SIH. Rather, I will undertake case studies. I will present novel diagnostics for covert syntax, and use them to demonstrate that covert syntax is available in a range of cases where the SIH requires it. Moreover, I will present certain data which are most easily understood if FOs do not have the additional option of directly composing with non-type $t$ arguments. Hence, this thesis will present a range of positive results that come from pursuing the SIH and thus intends to establish the SIH as an appropriate basis for further inquiry.

1.4 Schein (2017)

In defending the idea that and is interpreted as type $<t, <t,t>>$ in a broader set of cases than generally thought, this thesis draws crucial inspiration from Schein (2017), though my goals in this thesis are admittedly more modest than Schein’s. Schein’s main concern is with a class of data I have not yet discussed: particular cases where and appears to operate not on meanings of a type ending in $t$, but rather on type e meanings. The relevant data involve collective predicates:
Collective predication

John and Mary met.

Because (20) does not mean ‘John met and Mary met’, it cannot be analyzed with \([\text{and}]\), or its type-lifted variants in a cross-categorial theory. Link (1983) proposed that \(\text{and}\) in (20) encodes an unrelated operator, \([\text{and}_{\text{sum}}]\), which manufactures the mereological sum of two entities. \([\text{and}_{\text{sum}}]\) applies to John and Mary to create the sum \(\text{John} \oplus \text{Mary}\) and \(\text{met}\) is predicated of that sum.

Sum forming \(\text{and}\)

\([\text{and}_{\text{sum}}] = \lambda x . \lambda y . x \oplus y\)

In contradistinction to \([\text{and}_{\text{sum}}]\), what I presented as \([\text{and}]\) is often referred to as \(\text{logical and}\), given its transparent relationship to the \(\land\) connective. An important goal of Schein’s is to re-analyze collective predication with logical \(\langle t, \langle t,t \rangle \rangle [\text{and}]\). In achieving that goal, Schein proposes a novel theory of the syntax/semantics interface, where interpretation involves translation into a logical language with plural event pronouns. Omitting details, (20) would have the approximate paraphrase:

Paraphrase of (20)

\(\exists e \exists e' [\text{John participates in } e \land \text{Mary participates in } e' \land \exists E [E = e \oplus e' \land \text{met}(E)]]\)

The overt conjunction in (20) corresponds to the leftmost \(\land\) in (22). In effect, the distribution of event variables results in \(\text{John}\) and \(\text{Mary}\) each being locally associated with a type \(t\) meaning, and \(\text{and}\) conjoins those type \(t\) meanings. Schein’s comprehensive theory not only captures collective predication with \([\text{and}]\), but also creates type \(t\) meanings for \([\text{and}]\) to conjoin in environments like those in (15), where type-lifted variants of logical \(\text{and}\) seem to be invoked.

I join with Schein in pursuing a uniform analysis of logical \(\text{and}\) as type \(\langle t, \langle t,t \rangle \rangle\) — but remain officially ambivalent about whether logical \([\text{and}]\) should replace \([\text{and}_{\text{sum}}]\). Under the SIH, the grammar lacks type-shifting mechanisms which \(\text{systematically}\) derive multiple meanings for \(\text{and}\) from \([\text{and}]\) (i.e. no mechanism to derive \([\text{and}_2]\), \([\text{and}_3]\), \([\text{and}_4]\), ...). The potential between \([\text{and}]\) and \([\text{and}_{\text{sum}}]\), however, has a different character. Because \([\text{and}]\) and \([\text{and}_{\text{sum}}]\) are unrelated, that would be a case of \textit{accidental} homophony — which is consistent with the SIH in principle.

In pursuing the SIH, I will maintain the framework for syntax/semantics mapping outlined in Section 1.1. In the future, I plan to expand my inquiry past the SIH to analyze sum formation \(\text{and}\) as \([\text{and}]\) within the same framework — and I discuss prospects for doing so at the end of Chapter 5 (building on Winter 2001, Champollion 2015). Once that phase of the project is complete, a more detailed comparison between my approach and Schein’s will be possible. For now, I re-iterate the important relation I see between our hypotheses, but leave open whether the framework I adopt allows for re-analysis of sum formation.

2 Outline of the thesis

Before proceeding, I provide a chapter-by-chapter overview of the thesis. After laying out background in Chapter 2, Chapters 3-5 address coordination, and Chapter 6 addresses only. As noted, since it is
impossible to address these operators across their distribution, I will undertake specific case studies. My overarching aim is to show that in a range of cases where semantic flexibility seems to be needed, that covert syntax creating a sentential scope site is actually present. A core component of the thesis will involve cases where *and* and *only* appear to operate on object DPs.

### 2.1 Chapter 2: Semantic flexibility and the over-generation problem

Chapter 2 begins by providing detailed background on the traditional approach to FOs rooted in semantic flexibility, expanding on the sketch in Section 1.2. With that background in place, I lend initial plausibility to the SIH by pointing out that at least the type-shifting mechanisms that have been proposed to allow for flexible argument structure are too powerful. They predict more operators to be cross-categorial and thus show a broad surface distribution than is the case. The ambiguity mechanisms can apply at least to any operator that composes with sentence meanings — but the set of FOs does not extend much past *and* and *only*. While there may be ways to reign in type-shifting to avoid over-generation, I take the over-generation problem seriously as support for the SIH. The fact that *and* and *only* pattern differently from other operators with similar argument structure justifies the close attention paid to them specifically in the remainder of the thesis.

### 2.2 Chapter 3: A case for conjunction reduction

Chapter 3 focuses on cases where *and* appears to conjoin (quantificational) object DPs, requiring the type-lifted meaning \[\text{[and}_3]\] (cf. (16)). I argue that these data actually involve underlying vP conjunction. The verb in the second conjunct is elided, resulting in *and* being flanked by the two object DPs at the surface. Assuming the now widely accepted vP Internal Subject Hypothesis, vPs, like full sentences, are of type t. The structure for (23) is not (23-a), but (23-b).

(23) **Apparent object DP conjunction**

John saw every student and every professor.

a. \[\text{[TP John saw [DP every student [and [DP every professor]]]]}\]

b. \[\text{[TP John} \text{1 [VP T1 saw every student] [and [VP T1 saw every professor]]]}\]

The proposal revives, at least for this case, a Conjunction Reduction analysis, but one improved over the early literature. By adopting a vP conjunction structure, rather than an underlying full clausal conjunction, empirical problems for traditional Conjunction Reduction are avoided. First, I provide conceptual and empirical arguments that the structure in (23-b) is available. Conceptually, the derivation in (23-b) minimally adapts a mechanism which has been independently invoked for another conjunction structure: gapping. Conjunction Reduction “follows for free” from gapping. Empirically, additional vP structure is required to host sentential adverbs, predict certain cases of VP-ellipsis licensing, and to account for otherwise surprising scope readings. Hence, (23-a) must be available — and further data are most naturally predicted if (23-b) is not available in addition. \[\text{[and}_3]\] composes with two quantifiers to output a new quantifier, and a quantificational conjunction of object DPs would over-generate certain unattested scope readings. The diagnostics for vP syntax developed in Chapter 3 provide powerful tools to probe for hidden structure, and these same tools are
deployed in subsequent chapters, particularly in Chapters 4 and 6. In this way, Chapter 3 establishes an important empirical foundation for the overall study.

2.3 Chapter 4: Constructing pseudo-clefts

Chapters 4 and 5 take as their point of departure sets of facts which, at first, appear to provide a special challenge to the SIH. They aim to resolve the challenge, and to develop a full analysis of the relevant constructions, of interest independent of the SIH. The stimulus for Chapter 4 is a constituency puzzle. The SIH predicts that two DPs can never be directly conjoined with a variant of logical and. Yet, classical constituency diagnostics seem to pick out DP and DP constituents. I consider, in particular, the puzzle posed by specificational pseudo-clefts such as (24).³

(24) A constituency puzzle
What Obama approved was this bill and that bill.

Building on the diagnostics from Chapter 3, I probe into the analysis of (24) by introducing adverbs into the post-copular constituent. “Adverb data” such as (25) have not, to my knowledge, been discussed in the literature, and they provide a fresh analytical foothold for the chapter.

(25) Adverb data
a. What Obama approved was this bill and, with difficulty, that bill.
b. What Obama approved was this bill and possibly that bill.

I argue that data along these lines provide new evidence for a controversial idea, due to Ross (1972), den Dikken et al. (2000), and Schlenker (2003), that the post-copular constituent in a pseudo-cleft is the remnant of a full clause. The diagnostics, moreover, reveal a Conjunction Reduction structure within that clause. I propose that this bill and that bill has the underlying structure in (26): a TP with a vP conjunction. The structure for unembedded apparent object DP conjunctions in Chapter 2 is extended to the post-copular site in the pseudo-cleft — resolving the constituency puzzle.

(26) Underlying post-copular syntax
[TP Obama[TP[t1 approved this bill] [and [n.t1 approved that bill]]]]

The remainder of the chapter fits the post-copular TP into a full LF. Pseudo-clefts are typically interpreted exhaustively such that (24), for instance, conveys that Obama approved this bill, that bill, and no other. A widespread idea is that pseudo-clefts express identity, and that exhaustivity follows from the identity statement. Ross (1972) proposed that the pre-copular constituent is a question, and Schlenker (2003) proposed an identity semantics which would paraphrase (24): “the strongest true answer to ‘What did Obama approve?’ is ‘Obama approved this bill and that bill’”. Exhaustivity comes from strongest. Although an identity semantics is adequate for (24), I argue that it cannot

³This example involves referential DPs, but as we will see, there are mechanisms to interpret referential DPs as quantificational, so this example could be parsed with [and3], since it does not involve collective predication. Moreover, we will see reasons to believe there really is a parse with logical and.
capture certain adverb data, notably (25-b). It is not obvious how ‘Obama approved this bill and possibly that bill’ could be the strongest answer to the question.

I propose a new composition for pseudo-clefts, where the main “glue” is a covert operator with a meaning similar to only. The compositional role of the post-copular TP is to provide the prejacent for covert only. Based on Ross (1972), I take the pre-copular constituent to denote a Hamblin question, i.e. a set of propositions, and I take that question to provide only’s restrictor argument. The identity semantics is abandoned, and (24) paraphrases: “Obama only approved this bill and that bill”. Based on new observations about the interaction between only and adverbs, I show that this semantics can capture intuitive meanings for the data in (25), as well as (24).

2.4 Chapter 5: Coordinating questions

If coordinating conjunctions are uniformly interpreted as operating on truth-values, coordination of non-declarative clauses provides a prima facie challenge, since clauses with different illocutionary force are not obviously associated with truth-values. Chapter 5 takes up this challenge, by addressing conjunction and disjunction of questions. Whether or not questions can be disjoined has been a topic of considerable controversy, and I provide novel data showing that they can, building on recent observations in Ciardelli et al. (2015). Conjunction and disjunction data are illustrated in context in (27). (27-a) must be answered with a conjunction such as ‘Mary is the mother and John is the father’ while (27-b) may be answered with either ‘Mary is the mother’ or ‘John is the father’.

(27) Coordinating with and and or

a. (Need information about both biological parents:)
   Who’s the mother and who’s the father?

b. (Need information about one biological parent:)
   Who’s the mother or who’s the father?

I adopt a Performative Hypothesis for questions, and propose that and scopes above the performative prefix, so that (27-a) is interpreted: “I ask who the mother is and I ask who the father is”. Performatives are typically analyzed with a type t meaning. As in Chapter 4, I adopt a Hamblin semantics for questions, according to which they denote a set of propositions (their possible answers). I present a compositional analysis for the Hamblin denotation which includes a type t expression internal to the question CP itself. Whereas and scopes at a higher type t node above a performative prefix, or scopes at that lower type t node, inside the question CP.

The LF reconciles the data in (27) with the SIH, and sets up the second part of the chapter, which exploits this LF to develop a full analysis of question disjunctions like (27-b). Groenendijk & Stokhof (1989) were the first to controversially propose that question disjunctions exist. In their view, a disjunction like (27-b) offers up two questions ‘Who’s the mother?’ and ‘Who’s the father?’ and allows the responder to choose which one to answer. I pursue a different idea. The effect of scoping or inside the question CP is that (27-b) denotes a single question whose Hamblin set contains possible answers of two forms: x is the mother and y is the father. I propose that (27-b) is interpreted as a mention-some question, so the responder has to give one answer, but can give any one they want.
They can choose to answer ‘Mary is the mother?’ or ‘John is the father?’ and the choice is between answers to one question, not between questions. A contribution of Chapter 5 is to fit (27-b) into the broader typology of question types as a mention-some question.

2.5 Chapter 6: Focus operators

Chapter 6 shifts attention to focus operators, and does for only what Chapter 3 did for and. I consider cases where an object DP is focused or properly contains a focused constituent. In those data, only can precede the vP, or precede the object DP itself, as in (28). In (28-b), only appears to operate on the object DP, interpreted as [only3].

(28) Pre-vP vs. pre-DP only
   a. John only learned ONE language.
   b. John learned only ONE language.

By adapting empirical diagnostics for covert syntax from Chapter 3, I argue that appearances in (28-b) are, once again, deceptive. Only does not semantically compose with the object DP, but rather semantically composes with the vP, regardless of the surface string. As seen in Chapter 3, vPs have a sentential meaning, given the vP Internal Subject Hypothesis. Concretely, I propose that only constructions involve two heads. When the object DP contains the focus, one head (ONLY) attaches at the vP, while the other (a functional F) attaches to the DP.

(29) Underlying syntax for (28-a) and (28-b)
    [Johni [ONLY [vP t1 learned [F [DP one language]]]]]

The crux of the proposal then lies in how semantic and phonological labor is divided between heads. I propose that ONLY is always interpreted as [only], while F is semantically inert. Overt only may, however, realize either head, leading to surface variability between pre-vP and pre-DP only. The phonology is variable, while the semantics is constant, and consistent with the SIH. In effect, pre-DP only ends up as functional Agreement morphology, tracking the presence of higher, interpreted ONLY in the structure. The proposal builds on insights from a number of authors about the analysis of ‘only’ in different languages, notably Barbiers (2014) and Bayer (2016) on Colloquial Dutch, Hölle (2013) and Erlewine (2016) on Vietnamese, Horvath (2007) on Hungarian, and Lee (2004) on Korean.

2.6 How should I read the thesis?

The thesis will cover a wide empirical terrain, and the different chapters may be read together, or as discrete units. Readers familiar with the semantic flexibility approach to FOs may wish to skim the first half of Chapter 2, and skip to the second half which sets up the rest of the study. I see Chapter 3 as the core of the thesis. The empirical tests for covert structure in Chapters 4 and 6 are based on tests introduced in Chapter 3, and the theoretical claims in Chapter 4 are based on Chapter 3, as well. Still, the key points are reviewed in each chapter, so there is no harm in reading them separately. For readers who wish to have an overview of one chapter en route to another, I have included a preview
at the beginning of each chapter of the main data and claims that follow. Since Chapters 4, 5, and 6 build analyses of pseudo-clefts, question disjunction, and focus operators, they should be relevant to readers specifically interested in those topics, independent of matters of type-ambiguity.
Chapter 2
Semantic flexibility and the over-generation problem

1 Introduction

This chapter has two aims. The first is to lay background. I present the widespread approach to FOs rooted in semantic flexibility in more detail, illustrating initially with and. As previewed in Chapter 1, that approach has two prongs. One posits a richer toolkit for formulating lexical entries, which makes it possible to provide and with a single meaning with flexible argument structure. The other posits that and is ambiguous between a family of meanings of different semantic types, systematically derived from [and]. Each meaning has a particular semantic type and thus a particular argument structure, but ambiguity between meanings creates the overall impression of flexibility. I elect to frame the thesis in the type-ambiguity approach — which straightforwardly extends from and to only.

The second half of the chapter questions the ambiguity approach. I point out that the type-shifting mechanisms which have been posited to derive systematic ambiguity are too powerful. They are general enough to apply at least to any operator which, like and and only, has a basic meaning which composes with sentences. As such, all things equal, they predict that any sentential operator should be cross-categorial. A survey of sentential operators in English suggests that this is not so: only a relatively small proper subset of sentential operators pattern as cross-categorial. This over-generation problem, in my view, motivates taking seriously the SIH.

I present the different approaches to imbuing and with a flexible argument structure in Section 2, and then extend from and to only in Section 3. In Section 4, I present the over-generation problem and, in doing so, set up the project to pursue the SIH in the rest of the study.

2 Ways of introducing flexibility

This section overviews a range of approaches that have been taken to introduce flexibility into the semantics so as to allow and to compose with different kinds of arguments. In the final subsection, I re-iterate how the SIH situates within the landscape of different approaches.

2.1 Keenan & Faltz (1978, 1985)

An instructive departure point is the approach of Keenan & Faltz (further pursued in Winter 2001). They propose a single lexical entry for and which allows for a flexible argument structure. To do so, they abandon the system of semantic types based on e and t (and s) presented in Chapter 1 (see Section 1.1), and develop the idea that natural language builds lexical entries using the tools of Boolean algebra. I present a rather rudimentary sketch of their idea. To build up, we need to get a better for exactly the flexibility that and would have built into its meaning. To this end, let us consider just two environments where and occurs: with full TPs and with intransitive verbs:
(1) **Selected distribution of and**
   a. John laughed and Mary danced.
   b. John laughed and danced.

We already have an idea about what role *and* plays in (1-a). *And* operates on truth-values and makes a contribution similar to \( \land \). But, what does *and* do in (1-b)? In Montague's type system, intransitive verbs are modeled as functions. In Boolean semantics, they are modeled as sets. *Laughed*, for instance, denotes the set of laughers, and *danced* denotes the set of dancers.

(2) **Intransitive verbs as sets**
   a. \([\text{laughed}] = \{x : x \text{ laughed}\}\)
   b. \([\text{danced}] = \{y : y \text{ danced}\}\)

Note that there is a correspondence between the earlier predicate meanings for intransitive verbs and their set counterparts. Specifically, the predicate meaning for *laughed* in (3) is the "characteristic function" for the set in (2-a), since that predicate maps all and only elements of the set to true.

(3) **Recall: functional meaning**
   \([\text{laughed}] = \lambda x \cdot x \text{ laughed}\)

With intransitive verbs denoting sets, it is clear what role *and* must play in (1-b): the role of the set intersection operator, \( \cap \). *Laughed and danced* denotes the set of entities who both laughed and danced — the intersection in (4) — and the overall sentence asserts that John is in that intersection.

(4) **And as set intersection**
   \([\text{laughed and danced}] = \{x : x \text{ laughed}\} \cap \{y : y \text{ danced}\}\)

Adding together its roles leads to a hypothesis about the precise respect in which *and* is semantically flexible: it must switch between two roles, the role of \( \land \) in one environment and the role of \( \cap \) in another. Boolean semantics furnishes a single operator which is defined to play these two roles: the "meet" operator, indicated \( \cap \). Keenan & Foltz proposal to modal *and* as meet.

(5) **Flexible roles of and**
   a. *And* fulfills the role of \( \land \).
   b. *And* fulfills the role of \( \cap \).

(6) **Lexical entry for and**
   \([\text{and}] = \cap\)

A Boolean algebra consists of a partially ordered set, \( S \), along with certain operations which can apply to elements of \( S \). One is \( \cap \), which applies to two elements of \( S \) and returns their *infimum* or greatest lower bound. Indicating the ordering relation of elements of \( S \) with \( \leq \) ("less than or equal to"), the meet, \( z \), of \( x \) and \( y \) is the element which satisfies these conditions: (a) \( z \) is ordered no higher than \( x \) and \( y \), but (b) \( z \) ordered above every other element ordered no higher than both \( x \) and \( y \).
To compute the meet of two elements, then, it is crucial to define the ordering. In general, one element, \( x \), is ordered below another element, \( y \), iff the condition in (8) holds: that the meet of \( x \) and \( y \) is \( x \) itself. As we will see, axioms are required to specify when that condition obtains.

(8) Defining \( \preceq \)
\[
x \preceq y \iff x \sqcap y = x
\]

Keenan & Faltz's observe that the sets of possible denotations for different expressions of natural language all have the structure of a Boolean algebra. First, the set of possible denotations for a sentence is the set of truth-values, \{1, 0\}. In Boolean algebra, this set is commonly referred to as \( \mathcal{2} \). The axiom in (9) induces an ordering on elements of \( \mathcal{2} \):

(9) Ordering truth-values
\[
a \sqcap 1 = a
\]

Given this axiom, the meet of \( 0 \) and \( 1 \) is \( 0 \) and, as such, \( 0 \) is ordered below \( 1 \), per the definition of the ordering in (8). It follows, in turn, from this ordering that the \( \sqcap \) operator plays the role of \( \wedge \) when applied to two truth-values:

(10) Applying \( \sqcap \) to truth-values
\[
a. \quad 1 \sqcap 1 = 1 \\
b. \quad 1 \sqcap 0 = 0 \\
c. \quad 0 \sqcap 1 = 0 \\
d. \quad 0 \sqcap 0 = 0
\]

If \( \sqcap \) operates on the truth-values, \( 0, 0, 1, \) or \( 1, 0 \), the greatest lower bound is \( 0 \), due to the ordering. On the other hand, if \( \sqcap \) is applied to the truth-values \( 1, 1 \), the greatest lower bound is \( 1 \), since \( 0 \) is not present. Hence, meet outputs \( 1 \) just in case the two input truth-values are \( 1 \) — just like \( \wedge \). Modeling and as \( \sqcap \) derives the right result in the truth-value case.

Now, what happens with intransitive verbs? With intransitive verbs modeled as sets of individuals, the set of possible denotations for an intransitive verb is the set of all sets of individuals. If the set of all individuals is \( U \) (for universe), the possible denotations for an intransitive verb are elements of the power set of \( U \), \( \text{Pow}(U) \). \( \text{Pow}(U) \) again has the structure of a Boolean algebra, assuming the axiom in (11), which induces a partial ordering on its elements:

(11) Ordering sets
\[
\text{If } X \subseteq Y, X \sqcap Y = X
\]
In this case, the ordering tracks the subset relation. A set, \( X \), is ordered lower than set, \( Y \), just in case \( X \) is a subset of \( Y \). With this ordering in hand, \( \cap \) plays the role of \( \cap \) when it operates on sets. To illustrate concretely, suppose \([\text{laughed}]\) and \([\text{danced}]\) are the sets in (12):

(12) **Illustrative values for **\( \text{laughed, danced} \)

   a. \([\text{laughed}] = \{x : x \text{ laughed}\} = \{\text{John, Bill, Sue}\}\)
   b. \([\text{danced}] = \{y : y \text{ danced}\} = \{\text{John, Bill}\}\)

What is the meet of \([\text{laughed}]\) and \([\text{danced}]\)? The elements ranked no higher than \([\text{laughed}]\) are its subsets, ranked relative to each other in (13-a). Likewise, the elements ranked no higher than \([\text{danced}]\) are its subsets in (13-b). The meet of \([\text{laughed}]\) and \([\text{danced}]\) is the greatest element ranked no higher than either of them, which is \{\text{John, Bill}\} — the intersection of \([\text{laughed}]\) and \([\text{danced}]\). The meet of two sets is their intersection.

(13) \text{Sets} \leq [\text{laughed}]

\{\text{John}, \{\text{Bill}\}, \{\text{Sue}\}\} < \{\text{John, Bill}\}, \{\text{John, Sue}\}, \{\text{Bill, Sue}\}\}

(14) \text{Sets} \leq [\text{danced}]

\{\text{John}, \{\text{Bill}\}\} < \{\text{John, Bill}\}\}

If \( \text{and} \) encodes \( \cap \), \( \text{and} \) plays the role of \( \land \) with truth-values and \( \cap \) with intransitive verbs, just as it ought to. The system scales up. A quantifier may be modeled as a set of sets, and a transitive verb may be modeled as a set of ordered pairs. The set of all possible meanings for those expressions would, then, again have the structure of a Boolean algebra, due to the set-based axiom in (11). Hence, in Boolean semantics, flexibility may be built into a single lexical entry for \( \text{and} \) as \( \cap \). In effect, flexibility stems from the fact that the set of truth-values, i.e. 2, and sets of sets all have the structure of a Boolean algebra, due to axioms which induce on these sets a partial order.

2.2 Inspired by Montague (1973)

We return to the system of semantic types from Chapter 1, where the type associated with an expression tightly constrains its syntax. One way to capture \( \text{and} \) is, then, to assign it multiple meanings of distinct types. Each meaning has a specific argument structure, but ambiguity between meanings results in a flexible argument structure overall. This is the idea in Montague (1973). Montague, in effect, took \( \text{and} \) to be accidentally homophonomous between separate, unconnected lexical entries.\(^1\)

First, to compose with truth-values and fulfill the role of \( \land \), one entry is type \( <t,<t,t>> \):

(15) **Interpreting **\( \text{and} \) **like** \( \land \)

\[ [\text{and}] = \lambda p_t . \lambda q_t . p = q = 1 \]  \hspace{1cm} (type \( <t,<t,t>> \))

How does the role of \( \cap \) come about, when verbs and quantifiers denote functions, rather than sets? Montague distributes that role between a range of additional entries. \( \text{And}_2 \) conjoins intransitive verbs.

---

\(^1\)Montague formulated each meaning for \( \text{and} \) syncategorematically. I formulate categoromatic entries in \( \lambda \)-notation. Overall, the details of my presentation differ from Montague's, but I believe the spirit is the same.
To operate on \(<e,t>\) predicates and output a new \(<e,t>\) predicate characterizing their intersection, \(\text{and}_2\) must be interpreted as type \(<e,t,<e,t,et>>\). The required meaning is (17).

(16) **Recall: intransitive verb conjunction**

John \([v \text{ laughed}]\) and \(2 [v \text{ danced}]\).

(17) **Re-creating \(\cap\) with intransitive verbs**

\([\text{and}_2] = \lambda f_{et} . \lambda g_{et} . \lambda x . f(x) = g(x) = 1 \) (type \(<e,t,<e,t,et>>\))

\([\text{and}_2]\) composes with two predicates, \(f\) and \(g\), and returns a predicate which maps an entity \(x\) to \(1\) just in case \(x\) is in the extension of both \(f\) and \(g\). Concretely, the computation for the conjunction in (1-b) is sketched in (18). The predicate output in (18) maps \(x\) to \(1\) just in case \(x\) laughed and \(x\) danced, i.e. just in case \(x\) is an element of the intersection \(\{x : x \text{ laughed}\} \cap \{y : y \text{ danced}\}\).

(18) **Composition with \([\text{and}_2]\)**

a. \([\text{and}_2][\text{laughed}]^{w}([\text{danced}]^{w})\)

b. \(= \lambda x . [\text{laughed}]^{w}(x) = [\text{danced}]^{w}(x) = 1\)

Just as the semantic type associated with \(\text{and}\) limits it to a syntactic frame with type \(t\) sisters, the semantic type associated with \(\text{and}_2\) limits it to a frame with sisters of type \(<e,t>\). In order to capture quantifier conjunction in (19), another entry, \(\text{and}_3\), will be needed.

(19) **Recall: quantifier conjunction**

John saw \([DP \text{ every student}]\) and \(3 [DP \text{ every professor}]\).

\(\text{And}_3\) needs a meaning of type \(<e,t,<e,t,et>>\), which returns the characteristic function for the intersection of the sets characterized by the two input quantifiers. The required lexical entry is (20), applied concretely below. The function output maps a predicate to \(1\) just in case the extension of that predicate contains all the student entities and all the professor entities, i.e. just in case the set it characterizes is an element of \(\{X : \{x : x \text{ is a student}\} \subseteq X\} \cap \{Y : \{y : y \text{ is a professor}\} \subseteq Y\}\).

(20) **Re-creating \(\cap\) with quantifiers**

\([\text{and}_3] = \lambda F_{et} . \lambda G_{et} . \lambda f_{et} . F(f) = G(f) = 1 \) (type \(<e,t,<e,t,et>>\))

(21) **Composition with \([\text{and}_3]\)**

a. \([\text{and}_3][\text{every student}]^{w}([\text{every professor}]^{w})\)

b. \(= \lambda f . [\text{every student}]^{w}(f) = [\text{every professor}]^{w}(f) = 1\)

Conjunction of transitive verbs necessitates another entry still.\(^2\) If modeled as a set, a transitive verb would denote a set of ordered pairs. For instance, \(\text{hugged}\) would denote a set of hugger-hugged pairs, \(\{<x,y> : x \text{ hugged } y\}\). Then, \(\text{and}\) in (22) performs the role of \(\cap\) to return the set of pairs which are both hugger-hugged and petter-petted, \(\{<x,y> : x \text{ hugged } y\} \cap \{<x,y> : x \text{ petted } y\}\).

\(^2\)Montague only formulates entries for conjunction of truth-values, \(<e,t>\) predicates, and \(<e,t,t>\) predicates. The entry I provide for conjunction of \(<e,t>\) predicates is a natural extension.
Recall: transitive verb conjunction
John \(v\) hugged and4 \(v\) petted the dog.

The function version is generally “Curry’ed” so that there are two individual argument slots, rather than a single argument slot for an ordered pair. This fits better with a binary branching syntax, which provides an object DP and a subject DP one at a time. Modeling transitive verbs as \(<e,<e,t>>\) predicates assumes Currying. Recall the denotation for hugged:

Transitive verb
\[
\text{hugged}'' = \lambda x . \lambda y . y \text{hugged} x \text{ in } w
\]

This meaning is not exactly the characteristic functions for a set of ordered pairs. \([\text{hugged}]\) takes an entity \(x\) as its first argument and then maps that entity to the characteristic function for a set of entities \(y\) that hugged \(x\). Still, there is a clear correspondence to the set \(\{<x,y>: x \text{ hugged } y\}\). \([\text{hugged}]\) applied to two entities, John and the dog, returns 1 iff John hugged the dog, i.e. iff the pair \(<\text{John, the dog}>\) is in the set \(\{<x,y>: x \text{ hugged } y\}\). A variant of and is needed to operate on Curry’ed \(<e,e,t>\) functions to re-create the effect of \(\cap\) operating on sets of ordered pairs. This is and4, with a type \(<e,t,<e,t,e,t>>\) meaning:

Conjoining two-place predicates
\[
\text{and4} = \lambda f_{e,t} . \lambda g_{e,t} . \lambda x . \lambda y . f(y)(x) = g(y)(x) = 1 \\
\text{(type } <e,t,<e,t,e,t>>)
\]

With \([\text{and4}]\) applied concretely below, the output is a new \(<e,e,t>\) function, which would apply to John and the dog to return 1 iff John hugged the dog and John petted the dog, i.e. iff \(<\text{John, the dog}>\) is in the intersection \(\{<x,y>: x \text{ hugged } y\} \cap \{<x,y>: x \text{ petted } y\}\).

Composition with [and4]
\[
a. \quad [\text{and4}]([\text{hugged}]''([\text{petted}]'')) \\
b. = \lambda x . \lambda y . [\text{hugged}]''(y)(x) = [\text{petted}]''(y)(x) = 1 \\
c. = \lambda x . \lambda y . x \text{ hugged } y \text{ in } w \text{ and } x \text{ petted } y \text{ in } w
\]

Overall, what we hear as and is ambiguous between multiple lexical entries, each with a meaning of a different type. Sentences are conjoined with and2 (meaning of type \(<t,t,t,t>>\), intransitive verbs with and2 \((<t,<e,t,e,t>>)\), quantifiers with and3 \((<e,t,<e,t,t,e,t>>)\), and transitive verbs with and4 \((<e,t,<e,t,e,t>>)\). For easy reference, I collect the four required meanings together:

Sampling of meanings for and
\[
a. \quad [\text{and}] = \lambda p_t . \lambda q_t . p = q = 1 \quad \text{(type } <t,t,t,t>>) \\
b. \quad [\text{and2}] = \lambda f_{e,t} . \lambda g_{e,t} . \lambda x . f(x) = g(x) = 1 \quad \text{(type } <e,t,<e,t,e,t>>) \\
c. \quad [\text{and3}] = \lambda f_{e,t} . \lambda g_{e,t} . \lambda x . f(x) = G(f) = 1 \quad \text{(type } <e,t,<e,t,t,e,t>>) \\
d. \quad [\text{and4}] = \lambda f_{e,t} . \lambda g_{e,t} . \lambda x . \lambda y . f(y)(x) = g(y)(x) = 1 \quad \text{(type } <e,t,<e,t,e,t>>)
\]
We could consider more environments in which *and* occurs, motivating more entries for *and*, but the point should be clear by now. The semantic type of an expression narrowly restricts its distribution, and the broad distribution of *and* reflect pervasive type-ambiguity.

### 2.3 Partee & Rooth (1983)

Partee & Rooth (1983) make the crucial point that the different meanings for *and* that Montague posited all relate to one another. As they put it: "by linguist’s usual standards, there is clearly a generalization being missed. It is no accident that the same word *and* [...] is introduced in each rule" (p. 3). They observe that all of the higher-type meanings can be defined based on the type <t,<t,t>> meaning. I repeat the higher-type meanings below. For each, I provide the original formulation in (a), along with an equivalent re-formulation in terms of *and* in (b).

#### (27) Re-formulating and2

a. \[ \text{and}_2 = \lambda f_{et} \cdot \lambda g_{et} \cdot \lambda x \cdot f(x) = g(x) = 1 \]

b. \[ = \lambda f_{et} \cdot \lambda g_{et} \cdot \lambda x \cdot \text{[and]}(f(x))(g(x)) \]

#### (28) Re-formulating and3

a. \[ \text{and}_3 = \lambda F_{et} \cdot \lambda G_{et} \cdot \lambda f_{et} \cdot F(f) = G(f) = 1 \]

b. \[ = \lambda F_{et} \cdot \lambda G_{et} \cdot \lambda f_{et} \cdot \text{[and]}(F(f))(G(f)) \]

#### (29) Re-formulating and4

a. \[ \text{and}_4 = \lambda f_{et} \cdot \lambda g_{et} \cdot \lambda x \cdot \lambda y \cdot f(y)(x) = g(y)(x) = 1 \]

b. \[ = \lambda f_{et} \cdot \lambda g_{et} \cdot \lambda x \cdot \lambda y \cdot \text{[and]}(f(y)(x))(g(y)(x)) \]

Let us put one of the higher-type meanings, \[\text{[and}_2\] under a microscope. \[\text{[and}_2\] composes with two <e,t> predicates (f and g), and an entity (x). In the (a) formulation, \[\text{[and}_2\] computes internal to itself f(x) and g(x). Because f and g are one-place predicates, saturating them with x yields truth-values — i.e. f(x) is of type t and g(x) is of type t — and \[\text{[and}_2\] is defined to output 1 iff f(x) = 1 and g(x) = 1. Being type t, f(x) and g(x) are of the right type to be taken as the arguments of \[\text{[and}\]. Moreover, applying \[\text{[and}\] would yield the very same output: 1 iff f(x) = 1 and g(x) = 1. As such, the (a) formulation of \[\text{[and}_2\] is equivalent to the (b) formulation. Another way to appreciate the relationship between \[\text{[and}\] and \[\text{[and}_2\] is to compare the two truth-conditionally equivalent sentences in (30), the first with \[\text{[and}\] and the second with \[\text{[and}_2\].

#### (30) Two equivalent sentences

a. John laughed and John danced.

b. John laughed and danced.

In the first sentence, there are two crucial steps to the composition: first, \[\text{[laughs}\] and \[\text{[danced}\] are applied to John within their respective conjuncts to yield truth-values, and, second, \[\text{[and}\] operates on those truth-values. The composition is sketched in (31). The composition for the second sentence,
sketched in (32), has two parallel steps, but they occur in the opposite order: first, [[laughed]] and [[danced]] are conjoined and, second, John integrates.

(31) **Composition for (30-a)**

\[
\begin{array}{c}
\text{and}((\text{TP}_1)^\omega)(\text{TP}_2)^\omega) \\
\text{TP}_1 \\
\text{John laughed} \\
\text{and} \\
\text{TP}_2 \\
\text{John danced}
\end{array}
\]

(32) **Composition for (30-b)**

\[
\begin{array}{c}
\text{and}_2((\text{laughed})^\omega)((\text{danced})^\omega) \\
\text{John} \\
\text{and}_2 \\
\text{laughed} \\
\text{and} \\
\text{danced}
\end{array}
\]

To re-iterate, the crucial difference between the two derivations localizes in the order of the key steps. Looking bottom up, in (31), John integrates before conjunction, while, in (32), conjunction integrates before John. Since John is needed to create a truth-value, conjunction takes place after truth-values are created in the first sentence, but before any truth-value is created in the second sentence. Given the formulation of [[and]_2] in terms of [[and]], [[and]_2], in effect, allows [[and]] to operate on truth-values before they are actually present. Consider the &P in the second composition:

(33) **Interpreting the &P in (30-b)**

a. \[ \&P = [\text{and}_2]([[\text{laughed}}]^\omega)([[\text{danced}}]^\omega) \]
b. \[ = \lambda x . [\text{and}][[[\text{laughed}}]^\omega(x))([[\text{danced}}]^\omega(x)) \]

At this point, [[and]_2] has composed with the two predicates. Internal to itself, the variable x is introduced to saturate those predicates and create truth-values. [[and]_2] then applies [[and]] to those truth-values and abstracts over x, thus allowing x to get its value later in the composition, once John integrates. x serves as a "placeholder" to create a type t meaning inside the semantics before a constituent with a type t meaning is built by the syntax. All of the variants of and introduce placeholder variables to let [[and]] prematurely compose with truth-values.

With this in hand, we are in a position to state a recipe for how [[and]_2] was derived from [[and]], and to generalize to a broad recipe capable of productively deriving variants of and to compose with predicates of any type ending in t ("t reducible"), i.e. predicates which yield truth-values once all their arguments are saturated. The specific recipe for [[and]_2]] is given first in (34), and the generalization follows in (35). [[and]_3] and [[and]_4] both follow the general recipe.

(34) **Recipe for formulating [[and]_2]]**

(i) Take as arguments two <e, t> predicates (f and g).
(ii) Saturate f and g with a variable (x).
(iii) Apply [[and]] to the truth-values, f(x) and g(x).
(iv) Abstract over x.
General recipe for deriving variants of \textit{and}

(i) Take as arguments two n-place predicates of the same type ending in t (f and g).
(ii) Saturate the predicates with n variables, each of the appropriate type.
(iii) Apply \texttt{[and]} to the resultant truth-values.
(iv) Abstract over all the variables.

If the ambiguity that Montague proposes is \textit{systematic} rather than accidental, there are ways to capture it without proliferating unrelated entries. To do so, Partee & Rooth define a recursive procedure for interpreting \textit{and}. First, they provide a definition of \textit{conjoinable type}, by which any t-reducible type is conjoinable. In tandem, they define the operator in (36), which can compose with any conjoinable type (cf. von Stechow 1974, Gazdar 1980, and Keenan & Faltz 1979). If \textit{and} conjoins predicates, the clause in (35-b) recursively introduces bound variables until all the arguments of those predicates are saturated. Then, the clause in (35-a) applies to compose the resultant type t meanings with \texttt{\top}.

Conjoinable type

a. \texttt{t} is a conjoinable type.

b. If \texttt{b} is a conjoinable type, then for all \texttt{a}, \texttt{<a,b>} is a conjoinable type.

c. No other types are conjoinable.

Defining \texttt{\top}

a. \texttt{X} \texttt{\top} \texttt{Y} = \texttt{X} \texttt{\top} \texttt{Y}, if \texttt{X} and \texttt{Y} are truth-values

b. \texttt{\phi} \texttt{\top} \texttt{\psi} = \lambda z. \texttt{\phi(z)} \texttt{\top} \texttt{\psi(z)}, if \texttt{\phi} and \texttt{\psi} are functions.

Concretely, the illustration below shows how \texttt{\top} captures the effects of \texttt{[and]}. In (38-a), \texttt{\top} applies to \texttt{[laughed]} and \texttt{[danced]}. (38-b) results from clause (b) of the definition of \texttt{\top}: a placeholder variable \texttt{x} saturates those predicates, and that variable is abstracted over; moreover, \texttt{[laughed]}(\texttt{x}) and \texttt{[danced]}(\texttt{x}) are fed into another application of \texttt{\top}. Because \texttt{[laughed]}(\texttt{x}) and \texttt{[danced]}(\texttt{x}) are truth-values, clause (a) applies this time, and \texttt{[laughed]}(\texttt{x}) and \texttt{[danced]}(\texttt{x}) are conjoined with \texttt{\top} in (38-c). Since \texttt{[and]} encodes the meaning of \texttt{\top}, (38-c) is equivalent to (38-d), which is itself just \texttt{[and]} once its \texttt{<e,t>} argument slots are saturated.

Capturing \textit{and}

a. \texttt{[laughed]}(\texttt{w}) \texttt{\top} \texttt{[danced]}(\texttt{w})

b. = \lambda x. \texttt{[laughed]}(\texttt{w})(\texttt{x}) \texttt{\top} \texttt{[danced]}(\texttt{w})(\texttt{x}) \quad \text{(by clause (b))}

c. = \lambda x. \texttt{[laughed]}(\texttt{w})(\texttt{x}) \texttt{\top} \texttt{[danced]}(\texttt{w})(\texttt{x}) \quad \text{(by clause (a))}

d. = \lambda x. \texttt{[and]}(\texttt{[laughed]}(\texttt{w})(\texttt{x}))\texttt{[danced]}(\texttt{w})(\texttt{x})) \quad \text{cf. \texttt{[and]}}

The illustrations in (39) and (40) show that \texttt{\top} also captures \texttt{[and3]} and \texttt{[and4]}. The derivation for \texttt{[and3]} is identical to \texttt{[and2]}, just with different types. With \texttt{[and4]}, note that clause (b) applies twice to saturate and bind the two individual argument slots of the \texttt{<e,t>} predicates.
(39) Capturing and \(3\)

a. \(\llbracket\text{every student}\rrbracket^w \cap \llbracket\text{every professor}\rrbracket^w\)

b. \(= \lambda f_{et} \cdot \llbracket\text{every student}\rrbracket^w(f) \cap \llbracket\text{every professor}\rrbracket^w(f)\) \hspace{1em} (by clause (b))

c. \(= \lambda f_{et} \cdot \llbracket\text{every student}\rrbracket^w(f) \wedge \llbracket\text{every professor}\rrbracket^w(f)\) \hspace{1em} (by clause (a))

d. \(= \lambda f_{et} \cdot \llbracket(\text{every student})(f)\rrbracket \cdot \llbracket(\text{every professor})(f)\rrbracket\) \hspace{1em} cf. \(\llbracket\text{and}_3\rrbracket\)

(40) Capturing and \(4\)

a. \(\llbracket\text{hugged}\rrbracket^w \cap \llbracket\text{petted}\rrbracket^w\)

b. \(= \lambda x \cdot \llbracket\text{hugged}\rrbracket^w(x) \cap \llbracket\text{petted}\rrbracket^w(x)\) \hspace{1em} (by clause (b))

c. \(= \lambda x \cdot \lambda y \cdot \llbracket\text{hugged}\rrbracket^w(x)(y) \cap \llbracket\text{petted}\rrbracket^w(x)(y)\) \hspace{1em} (by clause (b))

d. \(= \lambda x \cdot \lambda y \cdot \llbracket\text{hugged}\rrbracket^w(x)(y) \wedge \llbracket\text{petted}\rrbracket^w(x)(y)\) \hspace{1em} (by clause (a))

e. \(= \lambda x \cdot \lambda y \cdot \llbracket(\text{hugged})(x)(y)\rrbracket \cdot \llbracket(\text{petted})(x)(y)\rrbracket\) \hspace{1em} cf. \(\llbracket\text{and}_4\rrbracket\)

Note that Partee & Rooth indicate their operator as \(\cap\) — and it is reminiscent of Boolean meet. The first clause specifies that \(\cap\) plays the role of \(\wedge\) when it conjoins type \(t\) meanings, and the second clause specifies how it conjoins predicates to approximate \(\cap\). As in Keenan & Faltz’s approach, Partee & Rooth’s \(\cap\) effectively builds flexibility into the lexical semantics of conjunction.

2.4 Jacobson (1999, 2014)

There is another way to capture the algorithm in (35). And could just have the meaning \(\llbracket\text{and}\rrbracket\), while flexibility comes from a separate type-shifting rule. That rule would apply to \(\llbracket\text{and}\rrbracket\) to derive higher-type meanings on an as-need basis, according to the algorithm. Jacobson (1999, 2014) has pursued this tactic, and subsumed (35) under a still more general rule — “the Geach Rule” (based on Geach 1972). The Geach Rule takes any function \(f\) of type \(<a,<b,c>>\) and derives a new function of type \(<d,a>,<d,b>,<d,c>>\) for some type \(d\), such that:

(41) Geach Rule (as adapted in Jacobson 2015:238)

\[ g_d(f) = \lambda X_{<d,a>} \cdot \lambda Y_{<d,b>} \cdot \lambda D_d \cdot f(X(D))(Y(D)) \]

Being of type \(<t,<t,t>>\), \(\llbracket\text{and}\rrbracket\) has the right profile to be subject to the Geach Rule. \(\llbracket\text{and}\rrbracket\) can play the role of ‘\(i\)’, with ‘\(a\)’, ‘\(b\)’, and ‘\(c\)’ in the statement of the types all being \(t\). Different values for ‘\(d\)’ create different variants of and. \(\llbracket\text{and}_2\rrbracket\) (type \(<e,t,et,et>>\) results if ‘\(d\)’ is type \(e\):

(42) \(\llbracket\text{and}_2\rrbracket\) from the Geach Rule

a. \(g_e(\llbracket\text{and}\rrbracket) = \lambda X_{<e,t>} \cdot \lambda Y_{<e,t>} \cdot \lambda D_e \cdot \llbracket(\text{and})(X(D))(Y(D))\rrbracket\)

b. \(= \llbracket\text{and}_2\rrbracket\)

As shown in (42), applying the Geach Rule effectively executes Partee & Rooth’s algorithm. The Geach Rule ‘lifts’ the first two arguments of and from type \(t\) to type \(<e,t>\). In addition, it introduces a type \(e\) argument, saturates the two initial \(<e,t>\) arguments with the corresponding type \(e\) variable, and applies the meaning originally input to the rule to the resultant type \(t\) meanings. In a similar fashion, \(\llbracket\text{and}_3\rrbracket\) (type \(<e,t,et,ett>>\) results if ‘\(d\)’ is type \(<e,t>>\) — as illustrated:
(43) \( [\text{and}_3] \) from the Geach Rule

a. \( g_{\text{et}}([\text{and}]) = \lambda X_{<e,et>} . \lambda Y_{<e,et>} . \lambda D_{e} . [\text{and}](X(D)(Y(D))) \)

b. \( = [\text{and}_3] \)

\([\text{and}_4]\) requires recursive application of the Geach Rule. A first application creates \([\text{and}_2]\), which is then input to a second application. In this second round, \([\text{and}_2]\) plays the role of 'f', so 'a', 'b', and 'c' are all type \(<e,et>\), and 'd' is type e. The result is a new meaning of type \(<eet,<eet,eet>\), corresponding to \([\text{and}_4]\). The second round is shown:

(44) \( [\text{and}_4] \) from the Geach Rule

a. \( g_{\text{et}}([\text{and}_2]) = \lambda X'_{<e,et>} . \lambda Y'_{<e,et>} . \lambda D'_e . [\text{and}_2](X'(D'))(Y'(D')) \)

b. \( = \lambda X'_{<e,et>} . \lambda Y'_{<e,et>} . \lambda D'_e . [\text{and}](X'(D')(D))(Y'(D')(D)) \)

c. \( = [\text{and}_4] \)

Overall, the Geach Rule is able to take \([\text{and}]\) and, through one or more applications, systematically type-shift it into the range of meanings derivable through Partee & Rooth's algorithm. A variant of \( \text{and} \) can be derived to compose with any t-reducible type. Still a question remains: what exactly is the Geach Rule? Where does it localize in the grammar? I entertain three possibilities.

2.4.1 Lexical rules

First, one could posit “lexical rules” which systematically apply to entries in the lexicon to derive new entries. In effect, there is a “core” lexicon, and a “derived” lexicon. The active-passive alternation, for instance, could involve a Lexical Rule, creating from a “core” active entry a “derived” passive entry. The passive rule changes both the phonology of the core entry (e.g. eat becomes eaten), and the semantic argument structure. Type-shifting operations would be lexical rules which have impact on the semantics, but not the phonology. The “core” entry for \( \text{and} \) would have the meaning \([\text{and}]\), while the type-shifting rule would create infinite “derived” entries, corresponding to different type-lifted variants of \([\text{and}]\). In this conception, \( \text{and}_2, \text{and}_3, \) and \( \text{and}_4 \) are lexical entries, but they are not part of the “core” lexicon. Just the “core” lexicon would be memorized.\(^3\)

2.4.2 Semantic rules

Jacobson takes type-shifting rules to be entirely in the semantics. Linguistic expressions of any complexity may be seen as an ordered triple of phonology, syntax, and semantics. Jacobson calls rules which take just one coordinate as input “unary” rules. Type-shifting operations are unary rules operating on the semantic co-ordinate: they take a meaning and map it to a new meaning, without any reflex in the phonology or the syntax.

In this conception, there is only a single entry for \( \text{and} \) anywhere in the lexicon, with the meaning \([\text{and}]\). \( \text{and}_2, \text{and}_3, \) and \( \text{and}_4 \) do not exist in the lexicon in any form. If so, conjunction in all environments must involve the same lexical item, as illustrated with intransitive verbs. In (45), the

\[^3\]This approach to type-shifting is, I believe, implicit in Rooth (1985), which defines a cross-categorial family of lexical meanings for \( \text{and} \) and \( \text{only} \). I will return to Rooth (1985) in more detail with respect to \( \text{only} \).
lexical item *and* is inserted into the structure. For the structure to be interpretable, the unary Geach Rule must apply to change (45) into (46).

\[(45) \text{ From the lexicon} \quad \text{(46) After type-shifting} \]

Assuming that type-shifting rules can apply at any point in the composition, any constituent in the tree is eligible to type-shift, not just the terminals. Nonetheless, we will focus on terminal expressions such as *and* in this thesis.

### 2.4.3 Syntactically represented morpheme

Supposing again that there is just one lexical item *and* with meaning [and] anywhere in the lexicon, the final option is to encode the type-shifting rule in a morpheme present in the syntax. Let us call the relevant morpheme Op and, for illustration, tailor it to intransitive verb conjunction:

\[(47) \text{ Structure with Op} \quad (48) \quad [\text{Op}] = \lambda F_{t,t}. \lambda f_{et}. \lambda g_{et}. \lambda x. F(f(x))(g(x)) \]

*And*, interpreted as *[and]*, is inserted from the lexicon, and is sister to Op. *[and]* is an appropriate input argument for Op, which effectively derives *[and]₂*.

### 2.4.4 Local summary

We have seen that type-shifting rules can systematically derive multiple meanings from *[and]*, and does so either as a lexical rule, a semantic rule, or a morpheme. The semantic conception is closest to the one Jacobson advocated: the Geach rule is a unary rule with no syntactic or phonological reflex.
2.5 Situating the SIH

We have now concluded our tour through the different ways of imbuing *and* with a flexible argument structure. Setting aside the possibility of accidental homophony in *and*, we can distill from the above discussion two general profiles of analysis:

<table>
<thead>
<tr>
<th>Distilling the flexibility approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Option 1:</strong> assign <em>and</em> a meaning with flexibility directly built in. These proposals make use of ( \cap ), defined one way or another.</td>
</tr>
<tr>
<td><strong>Option 2:</strong> assign <em>and</em> just the meaning ([\text{and}]), but systematically derive a family of other meanings of different semantic types through type-shifting (as a lexical rule, a semantic rule, or a morpheme).</td>
</tr>
</tbody>
</table>

The intent of the thesis is to pursue the idea that neither option exists. Option 1 is straightforwardly blocked if lexical entries are strictly formulated in the original type system, based on type e, t, and s. As we saw at the outset in Chapter 1, there is no way to write a single meaning with flexible argument structure just from those building blocks. Blocking Option 2 requires eliminating type-shifting rules (or at least a particular type-shifting rule, the Geach Rule) from the grammar. To facilitate exposition, I will take strict adherence to the type system for granted, and frame the discussion around the type-shifting approach, in particular Jacobson’s conception of type-shifting as a semantic rule. From this perspective, I formulate the SIH as (49), repeated from Chapter 1:

<table>
<thead>
<tr>
<th>(49) The Semantic Inflexibility Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural language does not have a type-shifting mechanism applicable to FOS.</td>
</tr>
</tbody>
</table>

Granting that *and* is stored in the lexicon as \([\text{and}]\), the question is whether the Geach Rule can derive higher-type variants. The key corollary of the negative answer provided by the SIH is that \([\text{and}]\) is the only available meaning for logical conjunction. Although I will frame discussion around type-shifting, most of the thesis affects both versions of the flexibility approach equally.

3 Generalizing to *only*

We have seen how a cross-categorial family of meanings may be derived for *and*. What about the other FO of primary interest in this thesis — *only*? Rooth (1985) showed that Partee & Rooth’s (1983) algorithm for deriving variants of *and* can also derive variants of *only*. I present that insight, and show that it is captured with the Geach Rule just as well.
3.1 Setting the stage

Abstracting away from the discussion of *and*, we have seen that the Geach Rule makes it possible to type-shift a function which applies to arguments of type $\sigma$ to variants which apply to arguments of any type ending in $\sigma$. For *and*, $\sigma$ was type $t$. Rooth makes the key observation that *only*, like *and*, has a basic meaning by which it composes with a sentence. As we will see, however, *only* must apply to the *intension* of a sentence, i.e. a proposition (type $<s,t>$), rather than an *extension*, i.e. a truth-value. This is because the statement of *only* makes reference to entailment, which is a relation between propositions. As such, for *only*, $\sigma$ must be type $<s,t>$.

To analyze *only* with the Geach Rule, then, the denotations of other elements must be modified so that they end in type $<s,t>$ rather than type $t$. The world parameter must give way to world arguments. Drawing inspiration from Creswell (1973) and Kratzer (1991), I replace each instance of $t$ in the earlier denotations with $<s,t>$. Intransitive verbs are not of type $<e,t>$, but of type $<e,st>$; quantifiers are not of type $<et,t>$, but of type $<est,st>$; and transitive verbs are not of type $<e,et>$, but of type $<e,<e,st>>$. The correspondences are shown:

(50) **Sample denotations**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>$[\text{laughed}]^w = \lambda x . x \text{ laughed in } w$ (type $&lt;e,t&gt;$)</td>
</tr>
<tr>
<td>b.</td>
<td>$[\text{laughed}] = \lambda x . \lambda w . x \text{ laughed in } w$ (type $&lt;e,st&gt;$)</td>
</tr>
</tbody>
</table>

(51) **Quantifiers**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>$[\text{every student}]^w = \lambda f_{et} . \forall x [\text{student}(x)(w) \rightarrow f(x)(w)]$ (type $&lt;et,t&gt;$)</td>
</tr>
<tr>
<td>b.</td>
<td>$[\text{every student}] = \lambda f_{est} . \lambda w . \forall x [\text{student}(x)(w) \rightarrow f(x)(w)]$ (type $&lt;est,st&gt;$)</td>
</tr>
</tbody>
</table>

(52) **Transitive verbs**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>$[\text{saw}]^w = \lambda x . \lambda y . x \text{ saw } y \text{ in } w$ (type $&lt;e,et&gt;$)</td>
</tr>
<tr>
<td>b.</td>
<td>$[\text{saw}] = \lambda x . \lambda y . \lambda w . x \text{ saw } y \text{ in } w$ (type $&lt;e,est&gt;$)</td>
</tr>
</tbody>
</table>

3.2 Analyzing *only*

With an intensional system in place, we are ready to analyze *only*. The basic meaning for *only* may be stated as a function of type $<st,st>$, as in (53):

(53) **Basic meaning for *only***

$[\text{only}]^ALT = \lambda p_{st} . \lambda w : p(w) . \forall p' \in ALT [p'(w) \rightarrow p \subseteq p']$ (type $<st,st>$)

$[\text{only}]$ operates on a proposition (called its 'prejacent'). *Only* is a focus-sensitive operator and its interpretation requires a set of focus alternatives, ALT, in addition to the prejacent. ALT could be an argument of $[\text{only}]$, but for now I indicate ALT as a parameter. $[\text{only}]^ALT$ presupposes that the prejacent is true, and asserts that alternatives in ALT are false, unless entailed by the prejacent. To

\[\text{In formulating denotations, I will sometimes use set notation when it is helpful for clarity. } p \subseteq p' \text{ for instance should be read: "the set characterized by } p \text{ is a subset of the set characterized by } p'\".\]
get a feel for how *only* would compose, let us consider the sentence in (54), and assume, as an idealization, that *only* directly applies to the TP.

(54) **Possible illustration of basic meaning**
It’s only the case that John learned ONE language.

(55) **LF for (54)**

```
TP2
  <s,t>
    only
      TP1
        (st)

John learned ONE language
```

["only"] presupposes that the proposition expressed by its prejacent, (56), is true. The assertion depends on the value of ALT, which is constrained by the placement of focus. In (55), prosody signals that *one* is focused in that *one* is pronounced with the greatest stress. With *one* focused, ALT contains propositions just like *John learned one language*, but with *one* replaced by other numerals, as in (57). *Only* asserts that all of these propositions are false, except (56-a), which is entailed by (in fact, equivalent to) the prejacent. Taking presupposition and assertion together, *only* conveys that John approved one bill, but not more — he approved exactly one bill.

(56) **The prejacent**

\[\llbracket TP_1 \rrbracket = \lambda w . \text{John learned } \geq 1 \text{ language in } w\]

(57) **The ALT parameter**

a. ALT = \{\lambda w . \text{John learned } \geq 1 \text{ language in } w, \n  \lambda w . \text{John learned } \geq 2 \text{ languages in } w, \n  \lambda w . \text{John learned } \geq 3 \text{ languages in } w, \ldots\} \]

Note that the example in (54) had *only* operating semantically on the TP, under the implicit assumption that the additional material *it is the case that* is semantically vacuous. Nonetheless, that LF may be idealized and *only* cannot in general attach to TPs, as the ungrammaticality of (58) shows.

(58) **Distributional restriction on only**

Only John learned ONE language.

For this reason, although Rooth proposed that *only* is lexicalized with the basic meaning of \[\llbracket\text{only}\rrbracket\], he infers its existence indirectly from examples like (59-a) and (59-b), where *only* apparently adjoins to the verb or VP, and DP. He proposed that *only*, like *and*, must be ambiguous in its semantic type, and
derived the different variants of *only* from [[only]]. [[only]] was a “phantom” operator: never attested, but an extant basic meaning, based on which higher-type meanings could be derived.

(59) **Distribution of *only***

a. John *only* learned ONE language.

b. John learned *only* ONE language.

From [[only]], a family of related meanings can be derived to operate on an argument of any type ending in <s,t>. [[only]$_2$] applies to a one-place property (type <e,st>), [[only]$_3$] to a quantifier intension (type <est,st>), and [[only]$_4$] to a two-place property (type <e,<e,st>>). If adjoined to the VP, *only* in (59-a) would be interpreted as [[only]$_3$]; (59-b) would invoke [[only]$_4$].

(60) **Derived meanings for *only***

a. $[[\text{only}]_2]^{\text{ALT}} = \lambda f_{\text{est}} \cdot \lambda x . [[\text{only}]^{\text{ALT}}(f(x))$

b. $[[\text{only}]_3]^{\text{ALT}} = \lambda f_{\text{est},st} \cdot \lambda f_{\text{est}} . [[\text{only}]^{\text{ALT}}(f(f))$

c. $[[\text{only}]_4]^{\text{ALT}} = \lambda f_{e,est} \cdot \lambda x . \lambda y . [[\text{only}]^{\text{ALT}}(f(x)(y))$

It is important to observe that the different variants of *only* are all derived by a parallel recipe to the different variants of *and*. The basic meaning for *only*, like the basic meaning for *and* operates on a sentence meaning (now a proposition). Each higher-type meaning for *only* takes a function which, after one or more arguments are saturated, yields a sentence meaning. The required placeholder variables are introduced to saturate the predicate, and [[only]] is applied to the resultant sentence meaning. Finally, the variables are abstracted back over, so their value can be filled in later in the derivation. The recipe is stated in a familiar way:

(61) **General recipe for deriving variants of *only***

(i) Take as argument an n-place predicate with a type ending in <s,t>.

(ii) Saturate the predicate with n variables of the appropriate type.

(iii) Apply [[only]] to the resultant <s,t> meaning.

(iv) Abstract over all the variables.

Given the close tie between the *and* variants and *only* variants, essentially the same type-shifting rules can derive them. The main complication has to do with their adicity: whereas *and* composes with two sentence meanings, *only* composes with one. The type-shifting rule provided earlier applied specifically to operators with a type of profile <a,<b,c>>. While [[and]] had the right kind of type (<t,<t,t>>, [[only]] does not (its type being <st,st>). The type-shifting rule must apply to functions with different numbers of argument slots, <t,<t,t>> and <st,st> alike.

While I called the earlier rule the Geach Rule, it was actually a generalization of a basic Geach Rule, which applies to lower types, of the form <a,b>. With ‘a’ and ‘b’ both <s,t>, the basic Geach Rule does apply to [[only]]. Much like how the earlier, higher-type version of the rule took a function of some type <a,<b,c>> and derived a new function of type <<d,a>,<<d,b>,<d,c>>> (for some type d), the basic, lower-type rule takes a function of some type <a,b> and derives a function of type.
I write the two rules side-by-side. Rule (a) applies if \( f \) has a type with profile \(<a,b>\), and rule (b) applies if \( f \) has a type with profile \(<a,<b,c>>\).

(62) **The Geach Rule**

\[
\begin{align*}
\text{a. } \text{g}_{e}(f) &= \lambda X_{<c,a>}. \lambda C_e . f(X(C)) \quad \text{(for f of type \(<a,b>\))} \\
\text{b. } \text{g}_{d}(f) &= \lambda X_{<d,a>}. \lambda Y_{<d,b>}. \lambda D_d . f(X(D))(Y(D)) \quad \text{(for f of type \(<a,<b,c>>\))}
\end{align*}
\]

It is clear that the two rules are doing the same thing, just for input functions with different numbers of argument slots. Indeed, the full range of meanings for *only* can be derived from rule (a). Taking ‘f’ to be [[only]] yields [[only2]] if ‘c’ is type e, as in (63), and [[only2]] if ‘c’ is type \(<e,st>\), as in (64). Deriving [[only4]] requires feeding [[only2]] into a second application of rule (a) as ‘f’ with ‘a’ and ‘b’ type \(<e,st>\), and ‘c’ type e, as in (65). This is all perfectly parallel to the earlier derivation of the different variants of *and* from application of rule (b).

(63) **Only2 from the Geach Rule**

\[
\begin{align*}
\text{a. } \text{g}_{e}(\text{[[only]]}) &= \lambda X_{<e,st>}. \lambda C_e . \text{[[only]]}(X(C)) \\
\text{b. } &= \text{[[only2]]}
\end{align*}
\]

(64) **Only3 from the Geach Rule**

\[
\begin{align*}
\text{a. } \text{g}_{estr}(\text{[[only]]}) &= \lambda X_{<est,st>}. \lambda C_{<e,st>}. \text{[[only]]}(X(C)) \\
\text{b. } &= \text{[[only3]]}
\end{align*}
\]

(65) **Only4 from the Geach Rule**

\[
\begin{align*}
\text{a. } \text{g}_{e}(\text{[[only2]]}) &= \lambda X'_{<e,est>}. \lambda C'_{e}. \text{[[only2]]}(X'(C')) \\
\text{b. } &= \lambda X'_{<e,est>}. \lambda C'_{e}. \lambda C_e . \text{[[only]]}(X'(C')(C)) \\
\text{c. } &= \text{[[only4]]}
\end{align*}
\]

For our purposes, it will be enough to note the parallel between rules (a) and (b), and I will refer to them collectively as the Geach Rule. I refer to Jacobson (1999) for a recursive generalization of the Geach Rule able to subsume both of the two rules (a) and (b). For us, the key point is this: variants of the Geach Rule are sufficiently general to apply to both *and* and *only*.

### 3.3 A note on notional convention

Before proceeding, I need to introduce a notational convention which will hold steadfast throughout the rest of the thesis. As noted, if *and* has a basic meaning of [[and]] of type \(<t,<t,t>>\) its type-lifted variants can compose with any arguments ending in \(t\). For this reason, in considering *and* it was sufficient to represent the world as a parameter and keep the types themselves extensional:

(66) **Types ending in \(t\)**

\[
\begin{align*}
\text{a. } \text{[[laughed]]}^w &= \lambda x . x \text{ laughed in } w \quad \text{\(\langle e,t\rangle\)} \\
\text{b. } \text{[[every student]]}^w &= \lambda f_{st} . \forall x \ [\text{student}(x)(w) \rightarrow f(x)(w)] \quad \text{\(\langle et,t\rangle\)} \\
\text{c. } \text{[[saw]]}^w &= \lambda x . \lambda y . y \text{ saw } x \text{ in } w \quad \text{\(\langle e,et\rangle\)}
\end{align*}
\]
Yet, if only is of type <st, st>, its type-lifted variants take arguments ending in <s, t>, and thus verbs and quantifiers must be analyzed with the types themselves intensional:

(67)  **Types ending in <s, t>**

a. \[
\text{[laughed]} = \lambda x . \lambda w . x \text{ laughed in } w
\]  
\(<e, st>\)

b. \[
\text{[every student]} = \lambda f_{st} . \lambda w . \forall x [\text{student}(x)(w) \rightarrow f(x)(w)]
\]  
\(<est, st>\)

c. \[
\text{[saw]} = \lambda x . \lambda y . \lambda w . y \text{ saw x in } w
\]  
\(<e, est>\)

We now have three options: we could attempt to maintain the world parameter system, and try to re-define only so that it operates on extensions; we could move entirely to the system with world arguments; or we could juggle between the two systems, keeping the simpler extensional types when possible, and presenting the intensional types when forced.

The first option is more difficult, since the definition of only must make reference to intensions, as noted. Nonetheless, Rooth (1985) did pursue a version of that option. He adopted a system of interpretation which translated expressions of the object language into a logical language. Operators such as only expressed “instructions” for how to translate. As a result, only could combine with an expression having a type t meaning, and compute the intension of that expression itself. Sentential only was defined along the lines of (68) (Rooth 1985:120). Where a is an expression with a type t extension, \(^\uparrow\) goes ‘up’ to its intension. C is a set of propositions (what I called ALT) and p is a proposition; \(^\downarrow\) goes ‘down’ to the extension of p. Rooth took the truth of the prejacent, a, to be an implicature, rather than presupposition, and conjoined it with the assertion.5

(68)  **Rooth: translation of sentential only**

\[
F_{\text{only}}(a) = \forall p \ [C(p) \land ^\downarrow p \rightarrow ^\uparrow a \subseteq p] \land a
\]

In the system I have adopted, however, only cannot itself take a type t argument and compute an intension. An attempt at re-stating only might be (69), but this is not well-formed. Since p is not an expression with a truth-value as its extension, but rather is a truth-value itself, \(^\uparrow p\) is a constant function, which would map any world to 1, if p = 1.

(69)  **Only cannot operate on truth-value**

\[
[\text{only}]_{\text{ALT}, w} = \lambda p_t : p = 1 . \forall p' \in \text{ALT} \ [p'(w) \rightarrow ^\uparrow p \subseteq p']
\]  
\(<t, t>\)

The only available route to unification, then, would be the second option where we move entirely to a system with world arguments, where quantifiers and verbs have a type ending in <s, t>. That would be coherent. Because <s, t> ends in t, the type-lifted variants of and introduced earlier could compose with arguments ending in <s, t>. Alternatively, type t could be eliminated altogether, and propositions taken to be a basic type, with the basic meaning for and re-defined to operate on propositions. Recall that, in formulating the revised entries for verbs and quantifiers every occurrence of type t was replaced with <s, t>.

5 Rooth defines the assertion of only such that it negates any alternative non-identical to the prejacent. I have adapted the entry so that only negates non-weaker alternatives, as before.
My approach in the thesis, however, will be to juggle so as to keep the types as simple as possible. I will use as a default the initial system where the world is a parameter, and will introduce world arguments only when necessary. In effect, when I discuss *and*, I will deploy the extensional types, and when I discuss *only*, I will deploy the intensional types. Still, it is worth noting that *only* with its basic meaning can be analyzed with extensional types. If we maintain the meanings in (66) for verbs and quantifiers, sentences are type t. While *and* composes straightforwardly through Functional Application, *only* must still take a proposition. In a system with world parameters, it would be defined as (70), of type <st,t>:

(70) **Defining only with world parameter**

\[
\llbracket \text{only} \rrbracket_{w,ALT}^{\Lambda T} = \lambda p : p(w) . \forall p' \in \text{ALT} \ p'(w) \rightarrow p \subseteq p'
\]

Even if sentence meanings are of type t, \llbracket \text{only} \rrbracket can compose with a proposition, given the composition rule of Intensional Functional Application (IFA) in von Fintel & Heim (2011). IFA allows a function requiring an argument of type <s,σ> to compose with an argument of type σ. IFA computes the intension of the argument, and applies the function to it:

(71) **Intensional Functional Application**

For any world w and assignment g, if \( \alpha \) is a branching node with daughters \( \beta \) and \( \gamma \) such that \( \llbracket \beta \rrbracket_{w,g} \in D_{\sigma} \) and \( \llbracket \gamma \rrbracket_{w,g} \in D_{<s,\sigma,\sigma>} \), then \( \llbracket \alpha \rrbracket_{w,g} = \llbracket \gamma \rrbracket_{w,g}(\lambda w' . \llbracket \alpha \rrbracket_{w',g}) \)

Whereas \llbracket \text{and} \rrbracket composes with two type t meanings through Functional Application, \llbracket \text{only} \rrbracket would compose with one sentence meaning through Intensional Functional Application, as shown in (72). Hence, the basic meaning of *only* can compose in a system where sentences are of type t.

(72) **Composing only**

\[
\begin{array}{c}
\text{TP}_2 \\
(t) \\
\text{only} \\
(<st,t>) \\
\text{TP}_1 \\
(t) \\
\text{John learned ONE language}
\end{array}
\]

\[
\text{IFA: } \llbracket \text{TP} \rrbracket^w = \llbracket \text{only} \rrbracket^w(\lambda w' . \llbracket \text{TP} \rrbracket^{w'})
\]

The intensional types are needed for type-shifting *only*. While the basic meaning of *only* can compose through IFA, IFA would not allow its type-lifted variants to compose with the verbs and quantifiers in (66). \llbracket \text{only}_2 \rrbracket, for instance, requires an argument of type <e,st>, and the intransitive verb in (66-a) is of type <e,t>. IFA does not allow those types to compose. If the order of the arguments in \llbracket \text{only}_2 \rrbracket were switched to <s,et>, then it could compose with an <e,t> sister through IFA — but the Geach Rule cannot generate the arguments of \llbracket \text{only}_2 \rrbracket in that order. For the type-lifted variants to compose, they truly must apply to meanings which themselves end in type <s,t>.
Overall, I will allow myself considerable fluidity in switching between the two systems of types, following, as mentioned, the heuristic of using the extensional types wherever possible. In Section 4, I will discuss both operators such as *and*, which most obviously operate on truth-values, and operators such as *only*, which most obviously operate on propositions. I discuss type-shifting and address them together, but I will still present their more obvious meaning, with type $t$ or type $<s,t>$ arguments as the case may be. The inventory of composition rules, as they must be formulated for each system, are in the appendix to the chapter.

3.4 Local summary

In sum, the same type-shifting mechanism that can systematically derive multiple meanings for *and* can derive multiple meanings for *only*. Both compose with a sentence meaning (type $t$ or $<s,t>$), and their type.lifted variants compose with arguments ending in a sentence meaning. The next section introduces the over-generation problem, providing impetus to pursue the SIH.

4 Over-generation problem

To re-iterate, the Geach Rule can apply to any function which operates on an argument of type $\sigma$ and generalize it to apply to arguments of any type ending in $\sigma$. To illustrate the over-generation problem, I will continue to consider operators which, like *and* and *only*, are naturally analyzed as applying to a sentence meaning, i.e. $\sigma$ is type $t$ or $<s,t>$. For any operator, $\text{Op}$, lexically listed as $[\text{Op}]$ of type $<st,st>$, for instance, the Geach Rule can derive the variants in (73), among others. Hence, all things being equal, the Geach Rule leads to the strong prediction in (74).

(73) Family of meanings for $\text{Op}$

a. $[\text{Op}] = ??$ \hspace{1cm} \text{(type $<st,st>$)}

b. $[\text{Op}_2] = \lambda f_{est} . \lambda x . [\text{Op}](f(x))$ \hspace{1cm} \text{(type $<est,est>$)}

c. $[\text{Op}_3] = \lambda F_{est,st} . \lambda f_{est} . [\text{Op}](F(f))$ \hspace{1cm} \text{(type $<est,st>,<est,est>$)}

d. $[\text{Op}_4] = \lambda f_{e,est} . \lambda x . \lambda y . [\text{Op}](f(y)(x))$ \hspace{1cm} \text{(type $<e,est>,<e,est>$)}

(74) Strong prediction of Geach Rule
All sentential operators should be cross-categorial.

This section shows that this prediction is not obviously borne out. I conduct a short survey of sentential operators, and evaluate whether each of them shows evidence of being cross-categorial. This investigation will identify two classes of operators: a small class which clearly do pattern as cross-categorial, and a larger class which, based on the available evidence, do not. The first class will not extend very far beyond *and* and *only*. While the generality of the Geach Rule is what makes it such a compelling hypothesis, its generality is, I believe, also a liability. If only a proper subset of operators eligible for the Geach Rule are FOs, it might be misguided to analyze FOs with a general
ambiguity mechanism widely applicable across operators. I will take the over-generation problem seriously as motivation to pursue the SIH.

4.1 An inventory of sentential operators

One way to first approximate the class of sentential operators is to employ a distributional criterion: an operator is sentential if it can adjoin to either a TP or a vP. So far, the discussion has implicitly assumed that the only node in the structure with a sentential meaning is the TP. As we will see in the next chapter, however, this is not so. Under current assumptions about the syntax and semantics of the verbal domain, it contains a projection, vP, which has a sentential meaning, too. Here is a partial list of operators show the expected distribution of a sentential operator:

- Coordinators (e.g. and, or)
- Negation (not)
- Focus operators (e.g. only, even)
- Modals (e.g. might, must, can)
- Modal adverbs (e.g. possibly, necessarily)
- Evidential adverbs (e.g. allegedly)
- Evaluated adverbs (e.g. luckily)
- Aspectual adverbs (e.g. again, usually)
- Temporal adverbs (e.g. yesterday)

Consistent with their distribution, these operators all seem eligible for very plausible semantic analyses in which they apply to sentence meanings. The coordinators and negation are interpreted as intensionalized versions of the logical connectives, and thus operate on type t arguments:

\[(75) \text{Sentential analyses of coordinators, negation} \]
\[
\begin{align*}
& a. \quad [\text{and}] = \lambda p_t . \lambda q_t . p = q = 1 \\
& \quad \text{("both p and q are true")}
\end{align*}
\[
\begin{align*}
& b. \quad [\text{or}] = \lambda p_t . \lambda q_t . p = 1 \lor q = 1 \\
& \quad \text{("at least one of p or q is true")}
\end{align*}
\[
\begin{align*}
& c. \quad [\text{not}] = \lambda p_t . p = 0 \\
& \quad \text{("p is false")}
\end{align*}
\]

The focus operators, the modals, and the adverbs all operate on sentence intensions. Possible denotations for the listed focus operators — only and even — are:

\[(76) \text{Sentential analyses of focus operators} \]
\[
\begin{align*}
& a. \quad [\text{only}]^{\text{ALT}} = \lambda p_w . \lambda w : p(w) \cdot \forall p' \in \text{ALT} [p'(w) \rightarrow p \subseteq p'] \\
& \quad \text{("p is true and no alternative to p is")}
\end{align*}
\]
b. \[\text{[even]}^{\text{ALT}} = \lambda p_{st} . \lambda w : \forall p' \in \text{ALT} [p \text{ is less likely to be true in } w \text{ than } p'] . p\]

("p is less likely than its alternatives, but true")

Since the seminal work of Kratzer (1977, 1981), modals have received a widespread analysis as quantifiers over possible worlds. As defined in (77), might and its adverbial counterpart possibly are existential modal operators. In all cases we will see, they apply to a proposition and assert that there is some epistemically accessible world at which that proposition is true. Throughout the thesis, \(\mathcal{E}(\text{SP(C)})(w)\) picks out the set of possible worlds epistemically accessible to the speaker (where SP(C) picks out the speaker in context C) from the evaluation world (w).

(77) Sentential analysis of modals

\[\text{[might]}^C = [\text{possibly}]^C = \lambda p_{st} . \lambda w . \exists w' \in \mathcal{E}(\text{SP(C)})(w) [p'(w)]\]

("p is true in some epistemically accessible world.")

Similar entries may be formulated for modals with a different accessibility relation (e.g. deontic or circumstantial) and different quantificational force (universal). The evidential adverb allegedly may receive a similar quantificational treatment, along the lines of (78):

(78) Sentential analysis of allegedly

\[\text{[allegedly]}^C = \lambda p_{st} . \lambda w . \forall w' [w' \text{ is compatible with what is alleged in } w \rightarrow p(w')]\]

("p is true in all worlds compatible with what is alleged.")

The evaluative adverb, luckily, applies to a proposition p, re-asserts p, and adds the entailment that the speaker finds p lucky.\(^6\) We can state a denotation of type \(<\text{st},\text{st}>\):

(79) Sentential analysis of luckily

\[\text{[luckily]}^C = \lambda p_{st} . \lambda w . p(w) \land \text{SP(C) finds p lucky}\]

("p is true and lucky, according to the speaker.")

A sentential analysis of again and the temporal adverbs is somewhat less obvious. These show the distribution of sentential operators, in that they can adjoin to vPs and TPs. Yet, they do not seem to operate on propositions, i.e. a set of worlds, since they need access to information about time. Their semantics can be modeled as that of a sentential operator, however, if we elaborate the meaning of a sentence to not just be a set of worlds, but rather a set of world-time pairs. Denotations for again and yesterday might be stated as (80). Here, s is the type of world-time pairs and, for a world-time pair c, t, is the time co-ordinate and w, is the world co-ordinate, denotations for again and yesterday may be stated as (80). Note that I revert to the earlier conception of proposition as sets of worlds for expediency in subsequent text.

\(^6\)Note that the latter entailment seems to have a different status from the former, such that the entailment that p is lucky is secondary or non-at-issue. Potts (2003, 2005) captures that intuition in a multi-dimensional semantic framework (discussed again in Chapter 4). I provide a more naïve entry for presentation here.
Sentential analysis of temporal operators

a. \([\text{again}] = \lambda p_{st} . \lambda c_s : \exists t' < t_c \ [p(w_c, t')] \cdot p(w_c, t_c)\)
   
   ("p is true in w at time t, and p was true at in w at a time before t."")

b. \([\text{yesterday}] = \lambda p_{st} . \lambda c_s : \exists t' \text{ is in the day before } t_c \wedge p(w_c, t')]\)
   
   ("p is true in w at some time the day before t.")

4.2 Preliminaries to the classification

The Geach Rule should apply to any of the operators listed above. The goal now is to assess which of these operators actually pattern as cross-categorial and which, if any, do not. To this end, I will home in on just one derived meaning for each operator: the meaning that would compose with quantifiers. I refer to any meaning of that form as \([\text{Op}_3]\), referring to the schema in (73). \([\text{Op}_3]\) will, in fact, serve as my primary testing ground throughout most of the thesis from this point on. I take an operator to be cross-categorial if there is evidence it can be interpreted as \([\text{Op}_3]\), i.e. if (a) it can adjoin to a quantificational DP, and (b) the sentence has the overall meaning expected with the operator interpreted as \([\text{Op}_3]\). I will call operators which meet these criteria Class 1 operators, and operators which fail to meet them Class 2 operators.

Two clarifications are in order about the criteria. Regarding Criterion (a), note that I will often construct examples with DPs which appear referential. Partee (1987) proposed a mechanism to interpret referential DPs as quantificational, so if \([\text{Op}_3]\) is an available meaning, it should still compose with these DPs. The mechanism is the Montague Lift operation which, applied to an entity, transforms that entity into a quantifier:

Montague Lifting a referential DP

a. \([\text{John}] = \text{John} \) \hspace{1cm} (type e)

b. \([\text{John}] = \lambda f_{est} . \lambda w . f(\text{John})(w) \) \hspace{1cm} (type <est,st>)

Regarding Criterion (b), note that in all data we will see in this section, the overall predicted interpretation with \text{Op} adjoined to a DP, interpreted as \([\text{Op}_3]\), is equivalent to baseline sentence with \text{Op} adjoined to either the TP or vP, interpreted as \([\text{Op}]. I illustrate in detail with the conjunction:

Recall: conjunction data

a. John saw every student and Mary saw every professor.

b. John saw \([\&P] \) every student and every professor.

How does composition proceed in (83-b)? The first step, repeated from earlier, has \([\text{and}_3]\) compose with every student and every professor to output a quantificational meaning for the conjunction:

Interpreting the &P

a. \([\&P] = [\text{and}_3][([\text{every student}]^w)([\text{every professor}]^w)]\)

b. \( = \lambda f_{est} \cdot [\text{and}][([\text{every student}]^w)([\text{every professor}]^w)](f)\)

c. \( = \lambda f_{est} \cdot [\text{every student}]^w(f) = [\text{every professor}]^w(f) = 1\)
Here, and throughout the thesis, I assume that quantifiers take scope through a covert displacement operation, Quantifier Raising (May 1977, hence ‘QR’), so that the sentence has the full LF:

(84) **Full LF for (82-b)**

\[ TP [\&P every student and every professor] 'l [- TP John saw t1] \]

When the DP moves, there are two syntactic consequences. First, I assume that a trace is inserted into the base position from which the DP has moved. And, second, a binder index is inserted into the structure just below the landing site of the DP.\(^7\) That binder index is indicated as ‘\(\lambda 1\)’ and the trace has the same index, 1. Composition principles are needed for interpreting the trace, and the node just above the binder index. Traces are interpreted according to the rule in (85):

<table>
<thead>
<tr>
<th>(85) <strong>Traces/pronouns rule</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>For a variable assignment (g) and natural number (n) such that (n \in \text{dom}(g)),</td>
</tr>
<tr>
<td>([t_n]^g = g(n))</td>
</tr>
</tbody>
</table>

The new parameter \(g\) is a variable assignment, which is a partial function from natural numbers to meanings. \(g\) is determined by the context, and I will assume that the co-domain of \(g\) includes meanings of different semantic types. To determine the semantic value for a trace with index \(n\), the variable assignment is consulted. \([t_n]^g\) is whatever meaning \(g\) maps \(n\) to. For a structure containing \(t_n\) to be interpretable, therefore, the domain of \(g\) must include \(n\). In (84), the trace must be of type \(e\) to compose with \(\text{saw}\). In other words, \(g\) must map 1 to an entity. If so, the Traces/Pronouns rule and FA deliver the type \(t\) meaning in (86) for the TP. This meaning depends on the variable assignment (it is ‘assignment-dependent’), since the value for the TP depends on \(g(1)\).

(86) **Interpreting the TP**

\([TP]^w = 1 \text{ iff John saw } g(1) \text{ in } w\)

The composition rule for interpreting the next node up is Predicate Abstraction. PA interprets a branching node one of whose daughters is a binder index ‘\(\lambda n\)’ and the other is some expression \(\alpha\). PA modifies the variable assignment so that \(n\) is mapped to a variable \(x\), interprets \(\alpha\) relative to the modified variable assignment, and introduces a lambda term binding (or “abstracting over”) \(x\).

<table>
<thead>
<tr>
<th>(87) <strong>Predicate Abstraction (PA)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>For any world (w) and variable assignment (g) such that (n \in \text{dom}(g)),</td>
</tr>
<tr>
<td>([\lambda n \quad \alpha ]^w g = \lambda x . [\alpha]^{w,g[x/n]})</td>
</tr>
</tbody>
</table>

In (82-b), the effect of PA is to take the assignment-dependent type \(t\) meaning for the TP, and return an assignment-independent meaning of type \(<e,t>\). PA modifies the variable assignment so that 1 is

---

\(^7\)The Copy Theory of Movement (Chomsky 1993) holds that a copy of the moved element is left in the base position rather than a trace, but I will not adopt that approach here, since the Copy Theory of Movement raises certain extra complications when combined with conjunction (see Hirsch 2015 for discussion).
mapped to $x$. Since $g[x/l](1) = x$, $\left[ TP \right]^{w,g}[x/l] = 1$ just in case John saw $x$. Abstracting over $x$ thus yields a predicate true of an entity iff John saw that entity at the evaluation world.

(88) **Output of PA in (84)**

a. $\left[ \lambda 1 \ TP \right]^{w,g} = \lambda x . \left[ TP \right]^{w,g}[x/l]

b. $= \lambda x . \text{John saw x in w}$

In the final step of the composition, the predicate derived in (88) is taken as the argument of the quantificational $\&P$ to derive the overall meaning for (82-b) in (89), which is equivalent to the meaning transparently derived with TP conjunction in (82-a).

(89) **Final predicted meaning**

a. $\left[ TP \right]^w$

b. $= \left[ \&P \right]^w(\lambda x . \text{John saw x in w})$

c. $= \left[ \text{and} \right](\left[ \text{every student} \right]^w(\lambda x . \text{John saw x in w}))(\left[ \text{every prof} \right]^w(\lambda x . \text{John saw x in w}))$

d. $= 1$ iff John saw every student in $w$ and John saw every professor in $w$

Generalizing from the composition just shown, we have now confirmed the overall meaning that $[Op_3]$ should yield. $[Op_3]$ should yield the same meaning as $[Op]$ when it adjoins to the $vP$ or the TP. As we will see in later chapters, the prediction becomes more involved when the sentences contain other scope-bearing operators. Such sentences will not, however, be germane at this point. We are now ready to classify the sentential operators: for each operator, is there evidence for a meaning as $[Op_3]$ (meriting Class 1 status), or is there not (meriting Class 2 status)?

4.3 **Class 1 operators**

We already see in detail that *and* and *only* are Class 1. *Or* patterns in kind with *and*, and *even* patterns in kind with *only*. In each example below, the (a) data point is a baseline example showing TP- or $vP$-attachment, and the (b) data point apparently shows the operator DP-adjoined, with no change in the truth-conditions.

(90) **Class 1: or**

a. Either John saw every student or he saw every professor.

b. John saw every student or every professor.

(91) **Class 1: even**

a. John even won a cooking contest.

b. John won even a cooking contest.

There is also some evidence that negation is cross-categorial. The examples in (93) seem to show that *not* can attach both at the edge of the $vP$, and also as an adjunct to a DP. In (92-b), the DP fronts to sentence-initial position via negative inversion, seemingly transporting *not* along. The overall meaning in (92-b) is identical to the overall meaning for the baseline in (92-a).
(92) **Class 1: not(?)**

a. John did not do a single thing right.
b. Not a single thing did John do right.

It is important to note, however, that pre-DP not is not freely available. Whereas not can precede initial quantifiers, inserting not prior to an object DP is often sharply degraded:

(93) **Restriction on pre-DP not**

a. Not many people saw John.
b. ?*John saw not many people.

Not is able to precede an object DP when there is contrast, as in (94-b), which is natural and interpreted as equivalent to (94-a). This usage is known as *contrastive negation*. When the contrast is removed in (95), pre-DP not sharply degrades once again.

(94) **Contrastive negation**

a. John did not eat radishes, but rutabagas.
b. John ate not radishes, but rutabagas.

(95) **Restricted once again**

a. John did not eat radishes.
b. *John ate not radishes.

Since the Geach Rule should allow not to be interpreted as \[\text{not}_3\] independent of position or contrast, I indicate the classification of not as Class 1 with a question mark in (92). There have been proposals that not in (92-b) is actually adjoined to the TP, rather than the DP (see Lasnik 1972, 1976). In a similar vein, certain proposals assign contrastive negation an entirely separate lexical item from regular negation. Wagner & Bale (2015), for instance, define contrastive not such that it directly composes with the contrastive piece. Still, these analyses are controversial (see e.g. Collins 2016 for a scope argument that initial not is DP-adjoined as \[\text{not}_3\], and e.g. Toosarvandani 2013 for an analysis of certain contrastive negation data with \[\text{not}_3\]).

4.4 **Class 2 operators**

The coordinators, *only* and *even*, and negation are the sole operators which qualify for Class 1 status, as far as I can tell. Categorizing an operator as Class 2 is more difficult than categorizing it as Class 1. An operator qualifies as Class 1 if there are any data apparently evincing \[\text{Ops}_3\]. To qualify as Class 2, there must be no data with \[\text{Ops}_3\]. Since it is never possible to prove the absence of a kind of data, we will have to settle for showing that the remaining operators cannot DP adjoin and be interpreted as \[\text{Ops}_3\] in particular test environments. The first environment is schematically (96), with the different operators to be placed in the position of Op, apparently as a DP adjunct.
(96) **Carrier sentence 1**
    John flew off to Op Paris.

(97) **Apparent structure**
    \[ TP \text{ John flew off } [pp \text{ Op } [DP \text{ Paris}]] ]

This environment controls for an alternative derivation which could place Op linearly prior to the DP without it actually adjoining to the DP. The derivation would involve extraposition: Op would attach on the clausal spine, while the DP extraposes above it (cf. Ross 1967):

(98) **Blocked: extraposition**
    \[ TP [TP [TP \text{ John flew off } [pp t]] \text{ Op} [DP \text{ Paris}]_1] \]

Extraposition is blocked in two ways. First, this kind of extraposition — often called “heavy NP-shift” — is only available when the DP is sufficiently “heavy” in terms of prosodic weight or syntactic complexity. In one view, heavy constituents are easier to process the later they come in the sentence, and heavy NP shift usually occurs only when necessary to aid in processing (see e.g. Gibson 1998, Hawkins 1990, 1994, Frazier 1985, Wasow 2002). Since the DP *Paris* is very light, it should at least resist extraposition. In addition, extraposition of *Paris* would strand the preposition *to, and extraposition categorically cannot P-strand (Ross 1967, Bresnan 1976, Stowell 1981).

The data in (99) cycle through all of the non-Class 1 operators in the test environment. In each case, there is a sharp contrast in acceptability between the baseline example in (a), where Op appears before the TP or vP, and the critical test example in (b). At least in this environment, these operators seem unable to adjoin to a DP, let alone adjoin to a DP and be interpreted as [OP3].

(99) **Class 2: can (modal)**
    a. John can fly off to Paris.
    b. *John flies off to can Paris.

(100) **Class 2: possibly (modal)**
    a. John possibly flew off to Paris.
    b. *John flew off to possibly Paris.

(101) **Class 2: allegedly (evidential)**

(102) **Class 2: luckily (evaluative)**
    a. This is great! Luckily, I’m flying off to Paris.
    b. This is great! *I’m flying off to luckily Paris.
(103) **Class 2: *again* (≈ temporal)**
   a. He again flew off to Paris.
   b. *He flew off to again Paris.

(104) **Class 2: *yesterday* (temporal)**
   a. Yesterday, he flew off to Paris.
   b. *He off to yesterday Paris.

Nonetheless, the first carrier sentence does have a certain limitation in that not all Class 1 operators can freely occur in that environment. While *and* is perfectly acceptable in (105-a), *only* is degraded for many speakers, if substituted for Op, as (105-b) shows.

(105) **Class 1 operators variably acceptable**
   b. ??John flew off to only Paris.

To make clear that the results are not linked to the preposition, I test the non-Class 1 operators in a second environment in (106). Both *and* and *only* are perfectly acceptable in this case.

(106) **Carrier sentence 2**
   John accused Op him of a crime.

(107) **Class 1 operators acceptable**
   a. John accused him and her of a crime.
   b. John assumed only him of a crime.

The target structure for (106) is (108), where Op is adjoined to a DP. Because there is no preposition, the extraposition parse in (109) is not categorically blocked due to P-stranding. I have attempted to strengthen the bias against extraposition in two other ways. First, instead of a proper name such as *Paris*, the DP following Op is a pronoun, which is even lighter. Second, there is a second argument of *accused (of a crime)* following the DP. To right extrapose *him, of a crime* would also have to extrapose. If extraposition is dispreferred when not required to aid in processing a complex DP, extraposing two constituents may be even more marked.

(108) **Target structure**
   \[ TP \text{John} \left[ \text{accused} \left[ DP \text{Op} \left[ DP \text{him} \right] \right] \right] \right] \text{[of a crime]} \]

(109) **Extraposition structure**
   \[ TP \left[ TP \left[ TP \text{John accused} \left[ t1t2 \right] \text{Op} \left[ DP \text{him} \right] \right] \right] \right] \text{[of a crime]} \]

The contrasts in (110) suggest that extraposition is successfully blocked, and corroborate the results from the first carrier sentence. Even in an environment where both hallmark Class 1 operators are acceptable, the non-Class 2 operators are degraded or unacceptable.
(110) **Class 2: can (modal)**
   a. John can accuse him of a crime.
   b. *John accuses can him of a crime.

(111) **Class 2: possibly (modal)**
   a. John possibly accused him of a crime.
   b. ?*John accused possibly him of a crime.

(112) **Class 2: allegedly (evidential)**
   a. John allegedly accused him of a crime.
   b. ?*John accused allegedly him of a crime.

(113) **Class 2: luckily (evaluative)**
   a. This is great! Luckily, John accused him of a crime.
   b. This is great! ?*John accused luckily him of a crime.

(114) **Class 2: again (aspectual)**
   a. John again accused him of a crime.
   b. *John accused again him of a crime.

(115) **Class 2: yesterday (temporal)**
   a. Yesterday, John accused him of a crime.
   b. *John accused yesterday him of a crime.

Note that, with both carrier sentences, the examples with the modal adverbs and the evidential seem marginally improved over the other examples. In fact, there are data in which these adverbs are fully grammatical. Bogal-Allbritten (2014) observed (116), with possibly. In (116-a), possibly adjoins to the vP. In (116-b), possibly seems to adjoin to the DP and the string is still grammatical.

(116) **A third test environment**
   a. John possibly climbed over the tallest mountain in Ireland.
   b. John climbed over possibly the tallest mountain in Ireland.

A natural impulse might be to re-categorize possibly as Class 1 on this basis. On closer inspection, however, Class 2 status is still supported: although (116-b) is grammatical, it does not have the meaning expected with [[possibly3]]. The baseline example in (116-a) is interpreted as (117), where TMI abbreviates the referent of the tallest mountain in Ireland. The meaning paraphrases: “in some epistemically accessible world, John climbed over the TMI”.

(117) **Interpreting (116-a)**
   a. [[possibly3](\lambda \overline{w} . \, \text{John climbed TMI in } \overline{w})
   b. = \lambda \overline{w} . \exists \overline{w'} \in \mathcal{E}(\text{sp(C)})(\overline{w}) \, [\text{John climbed TMI in } \overline{w'}]
If $[\text{possibly}_3]$ were an available meaning, it should be possible to interpret (116-b) as equivalent to (116-a). $[\text{possibly}_3]$ is defined in (118), the LF would be (119), and the computation of the equivalent reading is sketched below the LF.

(118)  
**Type-lifting possibly**

$[\text{possibly}_3]^C = \lambda \text{est}, st . \lambda f_{\text{est}} . [\text{possibly}]^C(F(f))$

(119)  
**LF for (116-b)**

$[TP \{DP_2 \text{ possibly } [DP_1 \text{ the tallest mountain in Ireland}]\} \lambda 1 [TP \text{ John climbed } t_1]]$

(120)  
**Sketching the composition**

a. $[\text{possibly}_3]^C([DP])(\lambda x . \lambda w . \text{ John climbed } x \text{ in } w)$
b. $= [\text{possibly}]^C([DP](\lambda x . \lambda w . \text{ John climbed } x \text{ in } w))$
c. $= [\text{possibly}]^C((\lambda f_{\text{est}}, st . f(\text{TMI})(w))(\lambda x . \lambda w . \text{ John climbed } x \text{ in } w))$
d. $= [\text{possibly}]^C(\lambda w . \text{ John climbed TMI in } w)$
e. $= \lambda w . \exists w' \in C(\text{John climbed TMI in } w') = (117)$

In fact, Bogal-Allbritten observes that (116-b) cannot be interpreted like (116-a). Observe that the predicted meaning does not entail that John actually climbed anything. It is epistemically possible that he climbed the tallest mountain in Ireland, but the evaluation world may not be one at which he climbed that mountain, or anything else for that matter. For (116-a), this is the right prediction: (116-a) does not intuitively entail that John climbed anything. For (116-b), however, the intuition is different: (116-b) does necessarily license an existential inference. Its only reading paraphrases: “John climbed something which was possibly the tallest mountain in Ireland”. Since the reading expected from $[\text{possibly}_3]$ is not attested, this example fails to show that possibly can be interpreted as $[\text{possibly}_3]$. A similar pattern arises with the other quantificational adverb, allegedly:

(121)  
**Replicating with allegedly**

a. John allegedly climbed over the tallest mountain in Ireland.

'It is alleged that John climbed over the TMI.’

b. John climbed over allegedly the tallest mountain in Ireland.

'John climbed over something which was allegedly the TMI.’

Testing evaluating adverbs in this frame requires a somewhat different strategy. Because (122) entails that John climbed the tallest mountain in Ireland, it is impossible to test for bleeding of an existential entailment when the adverb is moved pre-DP.

(122)  
**Inserting luckily in the frame**

John luckily climbed the tallest mountain in Ireland.

Still, it is possible to construct cases which make clear that evaluative adverbs do not yield equivalent meanings pre-TP and pre-DP in this frame. Consider the example in (123), with the evaluative adverb, sadly. On the one hand, (123-a) is acceptable and conveys sadness about the proposition expressed
by the entire clause. On the other hand, (123-b) is sharply odd. In so far as a reading is available, it would be the one paraphrased, where sadness is expressed just about the height of the mountain. This contrast is again not expected if \([\text{luckily}_3]\) can operate on the DP.

(123)  **Bringing out a contrast**

a. John sadly died on top of the tallest mountain in Ireland.
   ‘It is sad that John died on top of the TMI.”

b. #John died on top of sadly the tallest mountain in Ireland.
   ‘John died on top of something which was sadly the TMI.”

The temporal operators, *again* and *yesterday*, do not seem to be grammatical in the frame. The sentence with pre-DP *again* in (124-b) would be sensible with a reading parallel to the earlier (b) examples, but any reading is marginal at best. The counterpart with *yesterday* is worse still.

(124)  **Testing with *again***

a. John again talked to the happiest student in the class.
   ‘John talked to the student before and now again.”

b. ?*John talked to again the happiest student in the class.
   ‘John talked to the student who was happiest before and is again happiest.”

(125)  **Testing with *again* and *yesterday***

a. Yesterday, John talked to the happiest student in the class.
   ‘Yesterday, John talked to the student.”

b. ?*John talked to again the happiest student in the class.
   ‘He talked to the student who, yesterday, was the happiest student in the class.”

Overall, Bogal-Allbritten’s test environment supports the earlier conclusion that all of the operators under consideration are Class 2: even when certain of them appear to grammatically adjoin to a DP, they do not deliver the reading expected with \([\text{Op}_3]\). A question does arise as to how the right meaning for the grammatical (b) examples in Bogal-Allbritten’s frame may be derived, given that DP-adjunction of the operator, interpreted as \([\text{Op}_3]\) does not fit the bill. Bogal-Allbritten entertains a few compositional ideas, one of which is that (116-b) contains hidden syntax. *The tallest mountain in Ireland* obscures a reduced relative clause structure. She focuses on possibly:
(126) **Hidden relative clause structure**

```
    DP_2
     /\  
    /   \ RC
  PRO  AdvP
  \   /  possibly
   \ t_1 IDENT

the tallest mountain in Ireland
```

Bogal-Allbritten invokes a covert IDENT shifter. The DP *the tallest mountain in Ireland* has the type <s,e> meaning in (127-a). IDENT applies to that <s,e> meaning to returns the type <e,st> meaning in (127-b), the singleton property of being the tallest mountain in Ireland:

(127) **Output of IDENT**

a. \[ DP_1^{s,c} = \lambda w . \text{the TMI in } w \]

b. \[ IDENT DP_1^{s,c} = \lambda x . \lambda w . x \text{ is the TMI in } w \]

Like in a relative clause, a covert pronoun, PRO, originates just above IDENT. PRO is semantically vacuous, but moves to a higher position in the clause. I label that higher position RC (for Relative Clause). When PRO moves, it leaves in its base position a trace, and that trace saturates the first argument of \[ IDENT DP \] to derive a type <s,t> meaning, as in (128). *Possibly* then operates on the proposition with its basic meaning as \[ \text{possibly} \].

(128) **Creating a sentential meaning**

a. \[ t_1 IDENT DP_1^{s,c} = t_1 IDENT DP_1^{s,c}(t_1^{s,c}) \]

b. \[ = \lambda w . g(1) \text{ is the TMI in } w \]

(129) **Applying sentential **possibly**

a. \[ [AdvP]^{s,C} = \text{possibly}^{s,C}(t_1 IDENT DP_1^{s,c}) \]

b. \[ = \lambda w . \exists w' \in \varepsilon(SP(C))(w) [g(1) \text{ is the TMI in } w'] \]

The final compositional steps abstract back over the trace, and take the resultant property as the argument of the covert existential quantifier, \( \exists \), to derive the final quantificational meaning for the structure in (130). That quantifier, "something which is possibly the TMI", is then the object of *climbed* to derive the overall meaning: "John climbed something which was possibly the TMI."
Final quantifier meaning

\[ [DP_2] = \lambda f_{est} \cdot \lambda w \cdot \exists x \left[ \exists w' \in \delta(SP(C))(w) \ [g(1) \text{ is the TMI in } w'] \land f(x)(w) \right] \]

The same analysis extends to the other operators. If this analysis is on the right track, the data not only do not provide any evidence for \([Op_3]\), but they can be interpreted with just the basic meaning \([Op]\). There is a low scope position for \([Op]\) inside a hidden relative clause structure.\(^8\)

There are two residual questions, which I leave as open puzzles. First, why are again and yesterday unable to attach at the low scope site within the reduced relative? One possibility is that operators which make reference to times are not really sentential operators after all, and that information about times is not available at that scope site. Rather than re-envisioning propositions as sets of world-time pairs (recall discussion with ex. (80)), propositions may just be sets of worlds, with times introduced separately. And, the second question: why are those operators acceptable in Bogal-Allbritten’s frame not also fully acceptable in the earlier frames? Either the relative clause structure must not always be available, or there must be pragmatic factors biasing against the operators scoping there in the earlier data. This is an issue for further study.

Overall, this section has presented convergent evidence across three environments that not all of the operators in the original inventory pattern as cross-categorial. When the operators discussed in this section appear to DP-adjoin, they are either degraded, or yield a meaning unexpected with \([Op_3]\). These operators are all categorized as Class 2.

4.5 The resultant picture

It is time to take a step back, and return to the inventory of sentential operators. Discussion in the preceding subsections has motivated bifurcating that inventory into the two classes below. There is a small set of Class 1 operators which show the distribution and meaning expected of cross-categorial operators, and a larger set of Class 2 operators which do not pattern in kind.

Classification of sentential operators

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinators (and, or)</td>
<td>Modals (e.g. can)</td>
</tr>
<tr>
<td>Focus operators (only, even)</td>
<td>Modal adverbs (e.g. possibly)</td>
</tr>
<tr>
<td>not(?)</td>
<td>Evidential adverbs (e.g. allegedly)</td>
</tr>
<tr>
<td></td>
<td>Evaluate adverbs (e.g. luckily, sadly)</td>
</tr>
<tr>
<td></td>
<td>Aspeclual adverbs (e.g. again)</td>
</tr>
<tr>
<td></td>
<td>Temporal adverbs (e.g. yesterday)</td>
</tr>
</tbody>
</table>

The results of the survey are not as expected if the Geach Rule is active in the grammar. The Geach Rules makes the strong prediction that all sentential operators should be cross-categorial when, in

\(^8\)It may not even be necessary to invoke type-shifting in the form of \text{IDENT}. Coppock & Beaver (2015) challenge the traditional idea that definite descriptions are of type \(e\) (or \(<s,e>\)), in favor of an analysis where they are of type \(<e,st>\) — equivalent to what Bogal-Allbritten derives with \text{IDENT}. For Coppock & Beaver, definite descriptions must routinely occur with a covert existential quantifier. Fara (2015) advances related ideas for proper names.
fact, only a very restricted subset of sentential operators pattern that way. If the Geach Rule is active, extra constraints must reign in the distribution of the Class 2 operators. 9

One hypothesis might be that the constraints should come from the syntax, rather than the semantics. Perhaps Class 2 operators are subject to the Geach Rule and could be interpreted as [Op3], but we are unable to detect that meaning due to some syntactic constraint preventing them from adjoining to a DP. In other words, perhaps a structure like (132) for a Class 2 operator would be interpretable by the semantics, but is not generated by the syntax in the first place.

(132) **Unavailable structure with Class 2 operators**

\[ TP \text{ John accused} [DP \text{ Op} [DP \text{ him}]] \text{ of a crime} \]

For one sort of Class 2 operator, I find this possibility quite compelling, in particular for the modals *can* and *must*. These have been analyzed syntactically as verbal heads, which, in English, raise to T. I am not aware of any evidence that verbs can adjoin to DPs, so there could be a principled theory which simply fails to generate a structure like (133):

(133) **Instantiation with modal**

\[ TP \text{ John accused} [DP \text{ can} [DP \text{ him}]] \text{ of a crime} \]

However, appealing to a general syntactic ban on the necessary adjunction seems considerably less motivated for the other Class 2 operators (*again*, *yesterday*, *possibly*, and so forth). They are all adverbials and there is a commitment in the approach under consideration that certain adverbials can adjoin to DPs, namely *only* and *even* from Class 1. If the syntax generates structures with *only* and *even* adjoined to a DP, it should generate parallel structures for Class 2 adverbials. I do not see a general way to rule in DP-adjunction with *only* and *even*, but rule out (133). The only way to make sense of Class 2 operators in syntactic terms, then, is to impose specific constraints blocking these particular operators from adjoining to DPs on a case-by-case basis.

The alternative is for the syntax to allow Class 2 operators to DP-adjoin and for us to come at the problem from a semantic perspective. For each Class 2 operator, one would have to explain why that operator cannot type-shift. The viability of the type-shifting theory would depend on there being a principle explanation for each Class 2 operator. For certain operators, I do see plausible directions to try. As noted earlier, if the intension of a sentence is a set of worlds, rather than a set of world-time pairs, temporal operators would not apply to propositions. If quantifiers have a type ending in <s,t>, then *yesterday* could not type-lift to operate on a quantifier. The evaluative adverbs, too, may have a more involved semantics, where they operate on a non-at-issue dimension. Perhaps that could interact with the viability of type-shifting. I cannot, however, envision a principled response for why modal adverbs would be unable to type-shift, since these seem to operate on propositions and contribute at-issue content. Overall, there may be a viable project to reconcile the Class 2 operators

---

9I believe the over-generation problem is less severe for approaches which localize flexibility in the lexical semantics for a given FO, rather than in a widely applicable type-shifting rule. Still, if a recursive meaning for Class 1 operators is available (like Partee & Rooth’s \( \sqcap \) for *and*), the question does still remain why a recursive meaning would not also be deployed with Class 2 operators, at least in some languages.
with type-ambiguity — but I will pursue the opposite tactic. I take Class 2 operators seriously as reason to question the ambiguity approach.

4.6 The perspective of the SIH

The SIH turns the problem on its head. If the Geach Rule is not active in the grammar, then sentential operators are inflexibly interpreted as such, and Class 2 operators are the ones which show the expected distribution. The operators requiring further study are those in Class 1: if they are uniformly interpreted as sentential, how do they show a broad distribution?

As discussed in Chapter 1, the expectation is that these operators involve a richer underlying syntax, which furnishes an appropriate type (s)t scope site. Their broad distribution is not reflective of flexible semantics, but rather is an illusion created by the interfacing modules. The remainder of the thesis zeroes in on the Class 1 operators — specifically coordination and focus operators — and shows that the prediction is borne out in a range of cases. In this way, the bifurcation of operators made here sets in motion the rest of the project.

In order to have a principled account of the typology of operators, we will need some principled understanding of why the syntactic and phonological mechanisms which create a broad surface distribution with Class 1 operators do not do the same with Class 2 operators. I return to this question at the very end of the thesis, in the concluding chapter, after I have presented detailed proposals for covert syntax in the environments of concern in the thesis.

5 Conclusion

This chapter has reviewed the different ideas in the literature about how to imbue FOs with a flexible argument structure. Focusing on the idea of type-shifting, I demonstrated that the Geach Rule faces an over-generation problem: it predicts any sentential operator to pattern as cross-categorial, but only a limited subset do. I took the problem seriously as impetus to pursue the SIH. The next chapter begins that investigation by undertaking a close study of apparent object DP conjunction.

A Compositional rules

The compositional rules for the first system, with the world a parameter, have been introduced over the course of the last two chapters, and are collected here in one place. These rules include both Functional Application and Intensional Functional Application, in addition to Predicate Abstraction and the one role not discussed so far, Predicate Modification.

(134) Functional Application

For any world \( w \) and assignment \( g \), if \( \alpha \) is a branching node with daughters \( \beta \) and \( \gamma \) such that \( [\beta]^{w,g} \in D_{\sigma} \) and \( [\gamma]^{w,g} \in D_{<\sigma,\tau>} \), then \( [\alpha]^{w,g} = [\gamma]^{w,g}([\beta]^{w,g}) \).
(135) **Intensional Functional Application**
For any world $w$ and assignment $g$, if $\alpha$ is a branching node with daughters $\beta$ and $\gamma$ such that $[\beta]^{w,g} \in D_\sigma$ and $[\gamma]^{w,g} \in D_{<s,\sigma,t>}$, then $[\alpha]^{w,g} = [\gamma]^{w,g}(\lambda w'. [\alpha]^{w',g})$

(136) **Predicate Abstraction**
For any world $w$ and assignment $g$ such that $n \in \text{dom}(g)$,
\[ [\lambda n \alpha]^{w,g} = \lambda x . \ [\alpha]^{w,g[x/n]} \]

(137) **Predicate Modification**
For any world $w$ and assignment $g$, if $\alpha$ is a branching node with daughters $\beta$ and $\gamma$ such that $[\beta]^{w,g} \in D_{et}$ and $[\gamma]^{w,g} \in D_{et}$, then $[\alpha]^{w,g} = \lambda x . \ [\beta]^{w,g}(x) = [\gamma]^{w,g}(x) = 1$

In the second system, with world arguments such that every type $t$ in the first system is switched for $<s,t>$, the composition rules update as below. The world parameter is eliminated from the worlds, and Intensional Functional Application will no longer be necessary.

(138) **Functional Application**
For an assignment $g$, if $\alpha$ is a branching node with daughters $\beta$ and $\gamma$ such that $[\beta]^{g} \in D_\sigma$ and $[\gamma]^{g} \in D_{<\sigma,t>}$, then $[\alpha]^{g} = [\gamma]^{g}([\beta]^{g})$.

(139) **Predicate Abstraction**
For any assignment $g$ such that $n \in \text{dom}(g)$,
\[ [\lambda n \alpha]^{g} = \lambda x . \ [\alpha]^{g[x/n]} \]

(140) **Predicate Modification**
For any assignment $g$, if $\alpha$ is a branching node with daughters $\beta$ and $\gamma$ such that $[\beta]^{g} \in D_{et}$ and $[\gamma]^{g} \in D_{et}$, then $[\alpha]^{g} = \lambda x . \ [\beta]^{g}(x) = [\gamma]^{g}(x) = 1$

Given the expositional heuristic to prefer the first system over the second, I will present Chapters 3-5, on coordination, in the first system, and then move to the second with only in Chapter 6.
Chapter 3
A case for conjunction reduction

1 Introduction

In this chapter, I undertake a close study of and in one environment where it appears to require a type-
lifted meaning: data such as (1), where and appears between object DPs.

(1) Testing ground
John saw every student and every professor.

The most obvious syntax for this string is (2), where and directly conjoins the DPs that flank it. For
the conjunction every student and every professor to be interpretable, and must be able to compose with
quantifiers. This requires lifting and from its basic meaning of type <t,<t,t> to a meaning of type
<ett,<ett,ett>>, which I called [and3] in Chapter 1. The required meaning for and is given below the
structure, along with a composition for the &P. I refer to this parse as the “DP analysis”.

(2) Most obvious syntax
[TP John saw [&P [DP every student] [and [DP every professor]]]]

(3) Corresponding meaning for and
a. [and3] = \( \lambda F_{ett} \cdot \lambda G_{ett} \cdot \lambda f_{et} \cdot \lambda (F(f))(G(f)) \)
b. = \( \lambda F_{ett} \cdot \lambda G_{ett} \cdot \lambda f_{et} \cdot F(f) = G(f) = 1 \)

(4) Interpreting the &P
a. [[&P]] = [[and3]][[every student]] [[every prof]]
b. = \( \lambda f_{et} \cdot \lambda \left( \text{every student} \right)^w(f) \land \left( \text{every prof} \right)^w(f) \)
c. = \( \lambda f_{et} \cdot \forall x \left( \text{student}(x) \rightarrow f(x) \right) \land \forall y \left( \text{prof}(y) \rightarrow f(y) \right) \)

Is and interpreted as [[and3]], or is it really interpreted as [[and]]? Does apparent DP conjunction involve
DP conjunction, or is there really a richer underlying syntax? I take up these questions, and develop a
series of novel arguments for the perspective of the SIH: a richer underlying syntax is available, allowing
these data to be parsed with and operating on sentence meanings. Moreover, additional data are most
straightforwardly understood if direct DP conjunction with [and3] is unavailable.

1.1 Preview of main proposal

I pursue a simple idea: that apparent object DP conjunction involves a particular sort of Conjunction
Reduction (‘CR’) structure, one involving underlyingly vP conjunction. vP structure is obscured from

1I write ‘student(x)’ as shorthand for ‘x is a student in w’.

(5) **Main proposal**

Apparent object DP conjunction = vP conjunction.

As mentioned briefly in the last chapter, current syntactic assumptions predict that vPs, as well as TPs, are type t. This is due to the vP-Internal Subject Hypothesis, which holds that subjects enter the structure not in the specifier of the TP, but lower in spec-vP. All of the arguments of the verb are then saturated within the vP, creating a type t meaning. Consistent with the SIH and, scoping at the vP, is interpreted as [and]. The full underlying structure for (1) is:

(6) **Proposed structure for (1)**

\[ TP \text{John}_1 [\&p \text{[v}_1 \text{t saw every student]} \text{[and [v}_1 \text{t saw every professor]]}] \]

Note that there are two kinds of covert syntax in (6). Although John originates within the two vPs, it moves in “across-the-board” fashion out of those vPs to spec-TP, where it is pronounced. This leaves within each vP a covert trace, and it is actually the trace that saturates the argument of the verb. Hence, the first kind of covert syntax: a covert trace. The second kind of covert syntax is the second vP itself, which is unpronounced due to ellipsis: every professor is a remnant of that vP, ellipsis of which deletes saw from the overt string. Since both kinds of covert syntax — traces and ellipsis — are well known in the syntax literature, the proposal “follows for free” from independent syntactic mechanisms. The SIH is viable without new complexity in the syntax.

1.1.1 **Preview of empirical evidence for vP-conjunction**

Evidence that the proposed CR structure is available come from three empirical case studies. The first case involves apparent object DP conjunctions with an adverbial phrase in the second conjunct, as in (7), after Collins (1988). The only possible structure for (7) under the DP analysis is one with yesterday adjoined to the DP Chomsky in the second conjunct. As we saw in the survey in the first chapter, yesterday cannot in general adjoin to DPs, supporting CR over the DP analysis.

(7) **Case 1: sentential adverbs**

Harvard invited Labov and, yesterday, Chomsky.

Case 2 diagnoses the presence of an unpronounced VP in a second apparent DP conjunct by showing that this VP can serve as antecedent to license ellipsis of another elided VP. I consider examples like (8). The clause ten years after Brandeis did contains an elided VP interpreted as invited Chomsky. With CR, the second conjunct in (8) also contains the VP invited Chomsky (Harvard t invited Labov and, ten years after Brandeis invited Chomsky, invited Chomsky), and this VP is the antecedent for Δ. I demonstrate that no antecedent is available under the DP analysis.

(8) **Case 2: VP ellipsis**

Harvard invited Labov and, ten years after Brandeis did Δ, Chomsky.
Case 3 argues that CR is available even when an adverbial is not present by demonstrating that scope readings are observed which CR predicts, but which the DP analysis fails to predict. As we will discuss in detail, *and* in this example takes scope at a different height than the DPs it apparently conjoins, which is not predicted if *and* directly conjoins those DPs as \([\text{and}_3]\).

(9) **Case 3: Split scope**

John refused to visit any city in Europe and any in city in Asia. 

\((\text{and} > \text{refused} > \text{any})\)

1.1.2 **Preview of evidence against DP conjunction with \([\text{and}_3]\)**

Each of Cases 1-3 constitutes an argument that CR is available, but they cannot rule out that direct DP conjunction with \([\text{and}_3]\) is available as well. The last part of the chapter presents data which are most straightforwardly understood if direct DP conjunction with \([\text{and}_3]\) is unavailable:

(10) **Case 4: Frozen scope** (adapted from Rooth & Partee 1982)

Some company hired a maid and a cook. 

\((\text{some} > \text{and}, ?*\text{and} > \text{some})\)

The key fact is that *and* necessarily scopes below the subject existential in (10). If *DP and DP* could be parsed with \([\text{and}_3]\), a *maid and a cook* could be interpreted as a quantifier and, as such, could QR above the subject to derive the unattested reading. Eliminating \([\text{and}_3]\) rules out that parse, and I will show that the pattern is naturally predicted with \(\nu_P\)-level CR. Frozen scope in (10) will reduce to independent properties of \(\nu_P\) conjunction.

1.2 **A note on sum formation**

In addition to previewing what I do claim in this chapter, I want to re-iterate a point made in Chapter 1 about what I do not claim. At least at this point, I do not claim that object DPs can never be directly conjoined with *and*. I claim they can never be directly conjoined with a type-lifted variant of logical *and*. These two statements are not the same, given the possibility that logical *and* is accidentally homophonous with unrelated operators. In Chapter 1, I presented examples such as:

(11) **Collective predication**

a. John and Mary met.

b. *John met and Mary met.

If *and* in (11-a) were interpreted as \([\text{and}_3]\), this example would have a meaning equivalent to (11-b), which is unacceptable. *And* in (11-a) seems to perform a different function, and Link (1983) has proposed that this function is sum formation: *and*, interpreted as (12) of type \(<e,\langle e, e \rangle>\), produces a mereological sum with the individuals *John* and *Mary* as atoms. Whereas \([\text{and}_3]\) is a type-lifted variant of \([\text{and}]\) derivable from the Geach Rule, \([\text{and}_{\text{sum}}]\) is a separate operator altogether. It is compatible with the SIH for *and* to be accidentally homophonous between \([\text{and}]\) and \([\text{and}_{\text{sum}}]\).

(12) **Sum formation operator**

\([\text{and}_{\text{sum}}]^w = \lambda x . \lambda y . x \oplus y\)
Note that something similar to sum formation is possible with quantifiers, as in (13). This data point is interpreted in parallel to (12-b), where there is a single universal quantifier whose restrictor is a predicate of man-woman sums: “for every pair of a man and woman, the atoms of that pair met”.

(13) Sums with quantifiers
   a. Every man and every woman met.
   b. Every man and woman met.

The situation is similar in (14), which would paraphrase: “for every pair of a man and woman such that the atoms of that pair met, the atoms of that pair left”. The relative clause seems to operate on a predicate of man-woman sums to create a narrower predicate of man-woman sums whose atoms met. That predicate of sums then forms the restrictor of a single interpreted universal quantifier. Link (1983) first discovered examples like (14), and called them “hydrams”. Just as the mythical hydra contains too many heads, (14) contains too many quantifiers: every is pronounced twice, but interpreted once.

(14) Link’s (1983) “hydra”
   Every man and every woman who met at the party left.

These cases have been analyzed with another variant of and, which again a different basic meaning, not a type-lifted variant of [and]. As such, these examples do not threaten the SIH. Nonetheless, I want to spell out in a bit of detail exactly how they are parsed, since a parse along these lines should be available for most examples we will see in this chapter. I will assume that (13-a) and (14) are parsed in the same way, and refer to their common analysis as a hydra parse. To fully understand some of the data in this chapter, it will be necessary to control for a hydra parse. So, what exactly is a hydra parse? Illustrating with (13-a), the idea is that every man and every woman can have an LF like (14) (for an explicit syntactic derivation, see Fox & Johnson 2016 and references therein):

(15) Hydra LF

\[ [DP \forall [NP \text{man and woman}]] \]

The conjunction, and, is interpreted as (16), due to Link (1984) (see also Heycock & Zamparelli 2005, Fox & Johnson 2016). \([\text{and}_{\text{link}}]\) operates on two \(<e,t>\) predicates, \(f\) and \(g\), and outputs a predicate of sums, where one member of the sum is in the extension of \(f\) and the other is in the extension of \(g\). In (15), the output is a predicate of man-woman pairs, as in (17).

(16) Interpreting and

\[ [\text{and}_{\text{link}}] = \lambda f_{et} . \lambda g_{et} . \lambda X . \exists y [f(x) \land f(y) \land x \oplus y = X] \]

(17) Interpreting NP in hydra LF

\[ a. \quad [[NP]]^w = [[\text{and}_{\text{link}}]]([[\text{man}}^w][[\text{woman}}^w)] \]
\[ b. \quad = \lambda X . \exists y [\text{man}(x) \land \text{woman}(y) \land x \oplus y = X] \]

The predicate created with \([\text{and}_{\text{link}}]\) forms the restrictor of the quantifier. The DP as a whole, then, is interpreted as universally quantifying over man-woman pairs:
Interpreting DP in hydra LF

a. \[
[Dp]'' = [\text{every}''([\text{NP}''])]
\]
b. \[
= \lambda f. \forall x [\exists y \text{man}(x) \land \text{woman}(y) \land x \equiv y = X] \rightarrow f(X)
\]

Note that a hydra parse is only available in a conjunction like \textit{every man and every woman}, where the very same quantifier appears twice. When there are two different quantifiers, no predicate of sums derives. The follow minimal pair provides one illustration (for more, see Link 1984, Zhang 2007, Fox & Johnson 2016). Hence, the generalization is that for any one quantifier Q, Q N 1 and Q N 2 may have a hydra LF, as in (19). The observed restriction in the availability of a hydra parse will make it possible to control for that parse when necessary.

Hydra parse unavailable

a. A woman and a man who had agreed to dance with each other got on the floor.

b. *A woman and every woman who had agreed to dance with each other got on the floor.

Generalization for Q N 1 and Q N 2

\[
[dp Q [np N 1 \text{and} N 2]]
\]

Overall, then, the chapter claims that an apparent conjunction of object DPs can be parsed with logical [and] through a vP-level CR structure, and cannot be parsed with a type-lifted variant of [and], [ands], directly operating on the DPs. When the string Q N and Q N appears, there may well be a hydra parse, in addition to a CR parse. The hydra parse is something to consider only when it affects argumentation about [and] and [ands]: see Section 5 and Section 6.

One final point regarding sums and hydras is in order. As noted, the existence of [andsum] and [andlink] would be consistent with the SIH. That being said, some have denied their existence and attempted to re-analyze the relevant data with logical and, notably Winter (2001), Champollion (2015), and Schein (2015). If so, these data do become relevant to the SIH, which predicts that they must be analyzed with [and]. I will discuss these ideas briefly when I open discussion to different environments at the end of Chapter 5. Until then, I will remain officially agnostic about how to analyze sums and hydras and, for discussion, will assume the hydra analysis just presented, based on [andlink].

1.3 Roadmap from here

The preliminaries are done, and we are ready to begin the meat of the chapter. First, in Section 2, I present the proposal to revive CR with a vP-level conjunction parse. Then, in Sections 3-5, I present the empirical arguments that CR is available. In Section 6, I present the data which are best understood if [ands] is unavailable. Section 7 concludes.

2 Reviving conjunction reduction

What might the richer underlying syntax of apparent DP conjunction be? My hypothesis represents a revival and modification of an old idea: \textit{Conjunction Reduction} (hence, 'CR'). When CR was considered early on, the structures entertained had and uniformly conjoin full clauses (e.g. Gleitman 1965, Ross
1967, Lakoff & Peters 1969, McCawley 1968, Hankamer 1979). If so, (1) would derive from (21), with struck out material elided. With a CR analysis, and can compose as [and].

(21) **Traditional CR syntax: TP conjunction**

\[ TP \text{John saw every student} \text{ and } TP \text{John saw every professor} \]

Most researchers, however, quickly abandoned CR, in face of empirical evidence that and does not uniformly conjoin full clauses. An underlying structure with full clausal conjunction is plausible in (1) because the equivalence in (22) holds:

(22) **Apparent DP conjunction (a) = clausal conjunction (b)**

a. John saw every student and every professor.

\[ = (I) \]

b. \( \equiv \) John saw every student and John saw every professor.

But, there are cases where an apparent conjunction of non-clausal constituents is not equivalent to any full clausal conjunction. Partee (1970) noted (23). In (23-a), few obligatorily takes scope above and, so the meaning is that few rules are both explicit and easy to read at the same time. (23-a) allows that there may be many explicit rules, so long as few of them are also easy to read. This differs from (23-b) and (23-c), where few scopes below and. These entail that few rules are explicit.

(23) **Apparent constituent conjunction (a) \neq clausal conjunction (b), (c)**

- a. Few rules are explicit and easy to read.
- b. \( \Rightarrow \) Few rules are explicit and few rules are easy to read.
- c. \( \Rightarrow \) Few rules are explicit and it’s easy to read few rules.

A similar point can be made with apparent DP conjunction: whereas (24-a) (adapted from Rooth & Partee 1982) conveys that some one company hired both a maid and a cook (some > and), (24-b) is compatible with separate companies hiring a maid and a cook (and > some). Just as (23-a) cannot derive from the full clausal conjunctions in (23-b) and (23-c), (24-a) cannot derive from (24-b).

(24) **Apparent DP conjunction (a) \neq clausal conjunction (b)**

- a. Some company hired a maid and a cook.
- b. \( \Rightarrow \) Some company hired a maid and some company hired a cook.

While data like (23) and (24) seem to support a move away from CR, I suggest that such a move is premature. Interpreted most conservatively, these data only exclude a particular kind of CR analysis: one based on underlying structures with full clausal conjunction. They do not necessarily undermine the core idea of CR: that and may conjoin constituents of type t that are not surface apparent. So long as the structure contains type t nodes other than the top clausal node, CR may still be viable.

As mentioned in brief in the previous chapter, current ideas in the syntax literature — developed after Partee wrote — predict that there are nodes with a type t meaning other than the full sentence. In particular, the vP is of type t. My main proposal is that the right CR structure for object “DP conjunction” has underlying vP conjunction. vP-level CR allows and to be interpreted as [and], accordant with the SIH, and avoids the empirical pitfalls that classical CR faced.
In the following, I spell out the proposal. First, in Section 2.1, I review the evidence that vPs are of type t. Then, in Section 2.2, I propose a full derivation for apparent object DP conjunction as vP conjunction. The derivation builds on ideas in Wilder (1994), Schwarz (1998, 1999, 2000), and Toosarvandani (2013), and is directly based on Johnson’s (1996, 2009) analysis of another kind of conjunction construction, gapping. By drawing on extant syntactic tools, I show that CR “follows for free” as an epiphenomenon of mechanisms already recognized in syntactic theory. The SIH simplifies the semantics and (at least in this case) does not introduce complexity elsewhere. I will not show yet how the scope data in (23) and (24) are explained with vP-level CR; I defer to Section 7, which addresses scope in depth.

2.1 Step 1: vP are type t

The idea that vPs are of type t follows from the vP-Internal Subject Hypothesis (hence, ‘VPISH’), now well known and widely accepted. According to VPISH, all of the arguments of the verb are saturated within the vP (e.g. Kitagawa 1986, Koopman & Sportiche 1991, Kuroda 1988). An object merges as the complement to V, while a subject merges in the specifier of the vP. A simple transitive clause has the syntax in (25). Spec-TP is not a site of external merger, but a derived position, here resulting from the subject moving from spec-vP.

(25) Structure for a simple transitive

TP
   \[ \lambda 1 vP \]
   \[ t_1 \]
   \[ \nu VP \]
   saw DP

   the student

There are a number of very powerful arguments for vP-internal subjects, and I mention two here, following McCloskey (1997). One piece of evidence comes from scope. Observe that example (26) is ambiguous in the relative scope of the universal and negation:

(26) Evidence from scope

Every student didn’t come.
   a. “No student came.” (every > not)
   b. “Not all students came.” (not > every)

Assuming that scope correlates with syntactic position (as I will argue more later), the availability of not > every shows that the subject can occupy some syntactic position below negation. Adopting a structure similar to (25), wide scope of every results if the subject is interpreted in spec-TP, while narrow scope of every results from optional LF reconstruction of the subject back to its base position in spec-vP. More corroborating evidence comes from quantifier float:
Evidence from quantifier float

a. They all must have been drinking wine.
b. They must all have been drinking wine.
c. They must have all been drinking wine.
d. They must have been all drinking wine.
e. *They must have been drinking all wine.

It is commonly assumed that *they all* forms a constituent. In (27-a), that constituent is surface apparent in spec-TP, but not in (27-b)-(27-d), where *all* appears lower than *they*. Suppose that *they all* originates as a constituent in spec of the lowest vP (*drinking wine*), and moves to spec-TP in successive movement steps through intermediate vPs containing the auxiliaries *have* and *been*. The pattern can then be understood if, at each step of movement, *they all* can move as a constituent, or *they* can move stranding *all*. If so, *all* marks a vP-internal based position of the subject in (27-d), and intermediate landing site for the subject in (27-b) and (27-c). (27-d) is ill-formed, since there is no syntactic position for the subject between V (*drinking*) and its complement (*wine*). The vP-Internal Subject Hypothesis makes it possible to explain the complex pattern of quantifier float.

Now, what are the semantic consequences of the VPISH? The vital result is that the vP has a type t meaning. Consider again the structure in (25) above. For this discussion, I will assume a derivation where the subject does not reconstruct, so (25) is itself the LF. In order for the LF to be interpretable, the variable assignment must map t₁ to an entity. Then, the <e, <e, t>> predicate given by *saw* has its first argument saturated by *the dog* while its second argument saturated by t₁, delivering a type t meaning for the vP. In particular, the vP has the type t meaning in (28), which is assignment-dependent, because t₁ is free within the vP.

Interpreting the vP (type t)

a. \[ [vP]_w^{\omega} = [[saw]_w^{\omega}([\text{the dog}]_w^{\omega}))(t_1)_w^{\omega}] \] (type t)
b. = [saw](\text{the dog})(g(1))
c. = 1 iff g(1) saw the dog

To compute the overall meaning, Predicate Abstraction turns the type t meaning of the vP into an <e, t> predicate, \( \lambda x . x \text{ saw the dog} \), and that predicate is applied to the subject. The result is another type t meaning for the top node, this time assignment-independent, since the trace is bound:

Interpreting the TP (type t)

a. \[ [TP]_w = [\lambda x . x \text{ saw the dog in } w](John) \] (type t)
b. = 1 iff John saw the dog in w

In sum, the syntax literature has posited a richer structure than meets the eye in the verbal domain. The subject originates in spec-vP and moves to spec-TP. At LF, spec-vP either contains the reconstructed

---

2Throughout the dissertation, I will ignore the contributions of tense and aspect. If vPs must compose with tense and aspect, it is not obvious that they are type t. In the composition provided by von Stechow & Beck (2015), for instance, the vP is type <v, t> (i.e. a characteristic function for a set of events). Still, they do posit an aspectual projection just above vP which is of type t and it is possible that and really conjoins that projection in the data I consider, rather than vP. Thanks to Seth Cable, Paul Hagstrom, and David Pesetsky for pointing out this issue.
subject or, if the subject does not reconstruct, a trace, as just shown. This independent covert syntax creates a type t node below the TP, at the vP. My proposal is that apparent object DP conjunction has an underlying vP conjunction structure. And, scoping at the vP, is thus interpreted as [and].

2.2 Step 2: CR as vP conjunction

To build up to vP-level CR, I consider apparent object DP conjunction in light of another kind of conjunction construction: gapping. Gapping is illustrated in (30-a), along with (1):

(30) Gapping vs. CR
    a. John saw every student and Mary every professor.  
    b. John saw every student and every professor. 

There is a basic descriptive parallel between the two constructions: in both gapping and apparent DP conjunction, the verb (saw in (30)) is pronounced just once. The only difference lies in the number of pronounced subjects: two in gapping, but one in apparent DP conjunction. I translate the descriptive parallel into an analytical parallel. Johnson (1996, 2009), following Siegel (1987), proposed that gapping involves vP-conjunction. I extend the derivation to (1), with one modification, based on another independent syntactic mechanism: “across-the-board” (‘ATB’) movement. ATB-movement will account for the single pronounced subject. Overall, CR “follows for free” as an epiphenomenon of independent syntactic mechanisms: the mechanism for gapping, combined with ATB-movement. The analysis builds on Schwarz (1998, 1999, 2000) and Toosarvandani (2013), as I discuss after presenting the proposal.

2.2.1 Gapping

Because Mary every professor in (30-a) is simply not interpretable on its own, gapping must involve more structure than meets the eye. A possible first hypothesis might take gapping to derive from underlying full clausal conjunction — an option which has been posited (for a particularly relevant instance, see Schwarz 1998, 1999, 2000). If so, the underlying syntax for (30-a) would be:

(31) Possible hypothesis: TP conjunction
    \[ [TP \text{John saw every student}] \text{and} [TP \text{Mary saw every professor}]]

Yet, a TP conjunction parse turns out to be empirically inadequate, since it incorrectly predicts a range of scope facts. First, let us consider the scope of the subject in the left conjunct. In example (32), no boy in the left conjunct binds a pronoun in the subject of the right conjunct, his mother. This indicates that the left subject has the right conjunct within its scope. This example in (32) is due to Johnson (2000a,b), and the phenomenon goes back to McCawley (1993).

(32) Surprising scope for a subject
    No boy, joined the navy and his mother, the army.

A parallel effect does not arise in the counterpart full clausal conjunction in (34-a), which is expected from the syntax in (34-b). With TP conjunction, the left subject does not c-command the right conjunct. The scope of the left subject (underlined) comprises material in the same conjunct.
(33) **Gapping ≠ TP conjunction**
   a. *No boy joined the navy and his mother joined the army.*
   b. \[\text{TP no boy joined the navy} \text{ and TP his mother joined the army}\]

In order for the c-command domain of the left subject to include the right conjunct, the subject must take very wide scope, over the conjunction. Wide scope is corroborated by the interpreted scope of the negation in no boy relative to the conjunction. (32) is interpreted as (34-a) \(\text{no} > \text{and}\), not as (34-b) \(\text{and} > \text{no}\). Binding and scope converge: the left subject must scope above the conjunction, which is not predicted with TP conjunction.

(34) **Scope in (32): no > and**
   a. "for no boy \(x\) did \([x\text{ join the navy and } x\text{'s mother join the army}]\)"
   b. "[for no boy \(x\) did \(x\text{ join the navy}\) and \([x\text{'s mother join the army}]\)"

A related fact involves the scope of modals and sentential negation. As an illustration, consider Johnson’s (2014) example with the modal *can* in (35). Observations along these lines go back to Oehrle (1987), Oirouw (1987), Siegal (1984, 1987), and McCalwey (1993).

(35) **Surprising scope for modal**
   X can be true and Y false.

On the one hand, the sentence naturally conveys the meaning in (36-a): that X can be true and Y false simultaneously, as biased by the acceptable continuation in (36-b). On the other hand, the meaning in (37-a) is unavailable. This meaning says that it’s possible X is true and it’s possible that Y is false, though not necessarily simultaneously. (37-a) is weaker than (36-a), and would be compatible with the continuation in (37-b), which is in fact deviant.

(36) **Available: can > and**
   a. \(\Diamond [X\text{ is true} \land Y\text{ is false}]\)
   b. ... because they are logically independent.

(37) **Unavailable: and > can**
   a. \(\Diamond [X\text{ is true}] \land \Diamond [Y\text{ is false}]\)
   b. #... but X can’t be true if Y is false.

The pattern of available readings shows that the modal can and, at least in this example, must scope above the conjunction. Yet, the TP conjunction structure in (38) would yield exactly the opposite pattern: *can* scopes below *and*, not above.

---

3 The full generalization is more involved. When negation or a modal gaps in addition to a subsequent verb, then the negation or modal can scope below the conjunction. Siegel (1987) provides the example: ‘Ward *can’t eat caviar and his guest beans*.’ which may be interpreted ‘Ward can’t eat caviar and his guest can’t eat beans’. As far as I am aware, extent analyses of gapping do not capture the full generalization. The vP-conjunction structure I present predicts that negation should always take wide scope, assuming negation is outside the vP. As we will see, vP conjunction will make correct predictions for the scope of negation in apparent DP conjunction.
Again, gapping ≠ TP conjunction

\[ [TP_{X} \text{ can be true}] \land [TP_{Y} \text{ can be false}] \]

Overall, a TP conjunction parse does not fit the data. To derive exceptional wide scope of the left subject when it binds into the right subject, there must be some available parse other than TP conjunction. Moreover, the inability of the modal to take wide scope in (35) shows that, at least in certain examples, TP conjunction must be unavailable. If TP conjunction co-existed with another parse in (32), it would over-generate \textit{and} \textit{can}. Given the inviability of TP conjunction, Johnson (1996, 2009), following Siegel (1987), argues that gapping must instead involve a conjunction of vPs below a shared T—a move that has been widely adopted in subsequent work (e.g. Coppock 2001, Lin 2001, 2002). Returning to the initial example in (30-a), the base structure is (39). There are two conjoined vPs.

\textbf{Step 1: vPs are conjoined}

\[ [TP_{T} [vp_{\text{John}} \text{ saw every student}] \land [vp_{\text{Mary}} \text{ saw every professor}]] \]

In the second step of the derivation, the subject of the left vP moves to spec-TP, while the subject of the right vP remains in situ. Although movement of the left subject appears to violate Ross’s (1967) Coordinate Structure Constraint (‘CSC’), Johnson claims that A-movement is immune to the CSC (for further discussion of this issue, see Lin 2001, 2002).

\textbf{Step 2: John moves to spec-TP out of the left conjunct}

\[ [TP_{T} [vp_{\text{John}} \text{ saw every student}] \land [vp_{\text{Mary}} \text{ saw every professor}]] \]

These two steps solve the earlier problems. With regard to subject scope, because the left subject is in spec-TP, it correctly takes scope over the conjunction of vPs. With regard to the modal, assuming that modals lie in T, a modal would, likewise, take scope over the conjunction. The remaining steps of the derivation create the surface string: how does some of the underlying syntax end up covert, so that the verb is pronounced just once? Following Coppock (2001) and Lin (2002), I invoke VP-ellipsis in the right conjunct. In Step 3, \textit{every professor} moves out of the VP, and the VP elides in Step 4.

\textbf{Step 3: every professor moves out of the VP}

\[ [TP_{T} [vp_{\text{John}} \text{ saw every student}] \land [vp_{\text{Mary}} [vp_{\text{t2}} \text{ saw every professor}]]]] \]

\textbf{Step 4: the VP elides}

\[ [TP_{T} [vp_{\text{John}} \text{ saw every student}] \land [vp_{\text{Mary}} [vp_{\text{t2}} \text{ saw every professor}]]]] \]

Note that Johnson (1996, 2009) proposes an alternative means of deriving the surface string where \textit{every student} and \textit{every professor} both evacuate their respective VPs to create \textit{saw t}. Those two VPs \textit{saw t} then undergo leftward ATB-movement to a position above the conjunction resulting in a single pronounced VP in the correct linear position. I adopt ellipsis primarily for convenience.

Given an ellipsis analysis, two issues require further comment. First, what is the nature of the movement in Step 3? Following Weir (2015) on other ellipsis constructions, I assume that movement takes place at PF, and is motivated by the constraint in (43). In (30-a), \textit{every student} and \textit{every professor} are

\[ \ldots \]

4I show movement here as rightward, but if movement targets the VP edge, below \textit{Mary}, it could be leftward.
contrastive foci, so *every professor* must evacuate the VP prior to ellipsis to respect *ELIDEDFOCUS*. PF movement has no semantic effects, and we need not consider it further.

(43)  *ELIDEDFOCUS
        A F(ocus)-marked constituent must not be elided.

The second issue has to do with how VP ellipsis is licensed. Licensing of VP ellipsis requires that the linguistic context provide an “appropriate antecedent” for the elided VP, VPe. One idea holds that the antecedent for VPe must have an identical semantic value to VPe under any variable assignment (after Sag 1976, Williams 1977). Recent work has modified this condition in two ways (Rooth 1992a, Tancredi 1992, Fox 1999, Takahashi & Fox 2005). First: parallelism is optionally evaluated not relative to VPe itself, but rather relative to a larger constituent which contains VPe. And, second: the identity requirement is made sensitive to focus. Takahashi & Fox (2005) provide the condition in (44), minimally adapting Rooth (1992a):

(44)  Parallelism condition
        a. VPe can elide if VPe is reflexively dominated by a constituent PD (= parallelism domain), and the linguistic context provides an antecedent AC (= antecedent constituent) for PD which is semantically identical to PD, modulo focused marked constituents.
        b. PD is semantically identical to AC modulo focus if there is a focus alternative to PD, PDAlt, such that for every assignment function g, \([PDAlt] = [AC]\).

I will adopt the Parallelism Condition in (44) and assume that, while ellipsis takes place at PF, the Parallelism Condition is checked at LF (in the spirit of Merchant 2001, 2004). This system correctly predicts ellipsis in (30-a).5 Since ellipsis is checked at LF, the PF operations in Steps 3 and 4 are irrelevant for parallelism. Rather, the relevant structure is the full LF in (45):6

(45)  Full LF for (30-a)
        \[
        [T \ [vP1 every student F \lambda 1 [John F [vP1 saw t1]]
        \[\text{and} \ [vP2 every professor F \lambda 2 [Mary F [vP2 saw t2]]]]]]
        \]

*Every student* and *every professor* are QR-ed and, for exposition, I show John reconstructed into the left vP. All contrastive foci are F-marked. The elided VP is VP2 (saw t1). If VP2 itself were the PD, the AC would have to be VP1 (saw t1). Yet, for any variable assignment g for which g(1) \(\neq\) g(2), saw t1 and saw t2 have distinct semantic values — so, ellipsis would not be licensed. Since VP2 is reflexively dominated by vP2, vP2 is also a possible PD, however. vP2 is informally *Mary saw every professor* and, with *Mary* and *every professor* focused, alternatives to vP2 are propositions of the form *that x saw y*. That John saw

5Note that Coppock (2001) offers an alternative account for licensing of ellipsis in gapping based on the theory of Merchant (2001, 2004). Again, the choice between proposals is not crucial.

6I assume Heim's (1997) constraint in (i). Applied to (45), this prevents the trace of *every student* in the left conjunct and the trace of *every professor* in the right conjunct from being accidentally co-indexed.

(i). No Meaningless Co-indexing
        If an LF contains an occurrence of a variable v that is bound by a node α, then all occurrences of v in this LF must be bound by the same node α.
every student is such a proposition — and is the one expressed by \( vP_1 \). With \( vP_2 \) the PD and \( vP_1 \) the AC, ellipsis is licensed.

To summarize, we have seen that gapping must involve ellipsis, and traced the following logic: a TP conjunction parse seemed plausible, but faced empirical challenges, which were resolved with Johnson’s move to \( vP \) conjunction. I pursue the very same logic for apparent object DP conjunction. The early CR literature considered TP conjunction structures, and those faced irresolvable empirical challenges. I propose instead a derivation with \( vP \) conjunction, based on gapping.

### 2.2.2 From gapping to CR

The mechanism for gapping derives the string in (1) (*John saw every student and every professor*), with one minor modification. The derivation for CR begins as the derivation for gapping did: \( vPs \) are conjoined below a shared T. The difference from the gapping derivation lies in the subject of the \( vPs \). In gapping, the left \( vP \) had *John* as its subject and the right \( vP \) had *Mary*. Here, the subject of both \( vPs \) is *John*.

\[(46) \text{Step 1: } vPs \text{ are conjoined} \]
\[
[T_P \ T \ [v_P \ \text{John saw every student}] \ \text{[and]} \ [v_P \ \text{John saw every professor}]]
\]

This difference in subjects has a consequence at the second step of the derivation. In gapping, *John* moved to spec-TP out of the first conjunct, while *Mary* remained in situ in the second conjunct. In the present derivation, *John* undergoes ATB-movement out of both conjuncts.

\[(47) \text{Step 2: } \text{John ATB moves to spec-TP out of both conjuncts} \]
\[
[T_P \ \text{John}_1 \ T \ [v_P \ \text{t}_1 \ \text{saw every student}] \ \text{[and]} \ [v_P \ \text{t}_1 \ \text{saw every professor}]]
\]

The PF steps of the derivation then proceed exactly as before: *every professor* moves out of the VP in the right conjunct (Step 3), and ellipsis takes place (Step 4). The final output corresponds to the right surface string: the subject *John* is pronounced only once due to ATB movement, and the verb *saw* is pronounced only once due to VP ellipsis, with *every professor* a remnant.

\[(48) \text{Step 3: } \text{every professor moves out of the VP} \]
\[
[T_P \ \text{John}_1 \ T \ [v_P \ \text{t}_1 \ \text{saw every student}] \ \text{[and]} \ [v_P \ \text{t}_1 \ [v_P \ \text{t}_2 \ \text{saw every professor}]]]
\]

\[(49) \text{Step 4: the VP elides} \]
\[
[T_P \ \text{John}_1 \ T \ [v_P \ \text{t}_1 \ \text{saw every student}] \ \text{[and]} \ [v_P \ \text{t}_1 \ [v_P \ \text{saw-t}_2 \ \text{every professor}]]]
\]

Hence, we have achieved a CR derivation with \( vP \) conjunction and, given the VPISH, *and* can compose as \([\text{and}]\) when it scopes at the \( vP \), rendering the derivation consistent with the SIH. At this point, I want to re-iterate a conceptual result: CR “follows for free” from gapping. We have seen that CR involves just one mechanical difference from gapping: ATB movement of the subject. Crucially, ATB-movement is independently attested, for instance in (50):

\[(50) \text{Independent evidence for ATB-movement} \]
\[
\text{Who}_1 \ \text{does John like t}_1 \ \text{and Mary hate t}_1 ?
\]
Because gapping and ATB-movement are each independently attested, CR, which combines the two mechanisms, is an expected epiphenomenon.

2.2.3 Antecedent proposals

My proposal for apparent object DP conjunction owes an important debt to prior authors, in particular Schwarz (1998, 1999, 2000) and Toosarvandani (2013), who advanced related proposals for related constructions. Building on Wilder (1994), Schwarz proposed CR analyses for certain German and English data. For German, he considered examples such as:

(51) **“Odd” coordination in German**
    Äpfel ißt der Hans drei und zwei Bananen.
    apples eats the H. three and two bananas
    ‘Hans eats three apples and two bananas.’

At first, there seems to be a conjunction of DPs, drei Äpfel (‘three apples’) and zwei Bananen (‘two bananas’). Yet, German exhibits verb second word order, and (51) has the notable property that Äpfel has moved into first position. If (51) were a DP-conjunction, movement of Äpfel should violate the CSC, which prohibits an element from extracting out of just one conjunct. Schwarz explains the movement by adopting a richer underlying syntax, where the conjuncts are full clauses, in fact CPs:

(52) **Full clausal conjunction structure for (51)**
    [\&P [CP Äpfel ißt der Hans drei] [und [CP der-Hans-ißt zwei Bananen]]]

The conjunction is now high enough that the landing site of Äpfel is inside the left conjunct. As movement targets a position within the conjunct, not above the conjunction, the CSC is respected. He proposes a similar structure for certain cases of *either* ... *or* coordination in English.

In developing his analyses, Schwarz explicitly links CR to gapping — but, it is important to note that his conception of CR and gapping differ from the one I adopt. Following Ross (1970), Schwarz takes gapping to involve an underlying conjunction of full clauses, and analyzes CR in kind. As we have seen, full clausal conjunction runs into empirical difficulties. If gapping really is vP conjunction, the question remains as to how Schwarz’s data derive. One possibility is that full clausal conjunction and vP conjunction parses are both available in principle, and different languages or constructions invoke different structures. If so, we would need to work out what constrains the distribution of full clausal and vP-level CR parses. At least for the German data, however, Johnson (2002) claims that Schwarz misdiagnosed the structure. Based on scope evidence, he argues that the coordination must be lower, and provides a new definition of the CSC that permits movement of Äpfel. Regardless, Schwarz’s work remains an important predecessor to my own: my proposal represents an amalgam of Schwarz’s idea that CR has the same structure as gapping, and Johnson’s idea that gapping is vP-Coordination.

The most direct antecedent proposal is that of Toosarvandani (2013), who invokes vP-level CR for some (though not all) cases of English corrective *but*, also modeling on gapping:

(53) **Contrastive negation as vP conjunction**
    a. Max doesn’t eat chard, but spinach.
    b. \[\text{TP Max} \text{[NEG [vP t1 eat chard]] [but [vP t1 eat spinach]]}\]
In (53-a), the negation attaches at the vP, but only scopes over the first coordinate: (53-a) entails that Max did not eat chard, and that he did eat spinach. This indicates that the first coordinate is large enough to include the negation, which Toosarvandani achieves with the CR syntax in (53-b). Note that Toosarvandani analyzes but as a coordinator, and there is nothing about his proposed vP-level CR syntax that is contingent on the presence of negation or that particular coordinator. As such, if vP-level CR is available in (53-a), there is a prediction that it should also be available when the coordinator is and. This chapter can be seen as pursuing that prediction.

I cannot attempt an exhaustive review of previous CR proposals here. Rather, I have highlighted those most inspirational for my own in that they link CR and gapping. See Aoun, Benmamoun, & Sportiche (1994, 1999) for a further empirical argument for the availability of CR from first conjunct agreement in Lebanese Arabic, and Yoon & Lee (2005) for a related argument from Korean, among other works.

2.3 Summary and next steps

This section has spelled out my syntactic hypothesis, and provided conceptual motivation from independent constructions. The idea is that apparent object DP conjunction involves underlying vP conjunction. The subject is pronounced once due to ATB-movement, and the verb is pronounced once due to VP-ellipsis. CR “follows for free” from the existence of gapping. Given the VPISH, vPs are of type t, so and is interpreted as [and] — accordant with the SIH. Having established that CR is analytically viable and conceptually expected, I will shift attention to empirical matters. We have already seen some potential evidence from different constructions for CR, and I will provide new evidence for CR in the apparent DP conjunctions of concern here. Since a CR parse with [and] and a direct DP conjunction parse with [and3] are not mutually exclusive, there are, in fact, three possibilities to empirically dissociate:

(54) Three hypotheses

<table>
<thead>
<tr>
<th></th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>And is type-flexible</td>
<td>DP conjunction with [and3] available, CR unavailable</td>
<td>DP conjunction with [and3] available, CR also available</td>
</tr>
<tr>
<td></td>
<td>And is not type-flexible</td>
<td>DP conjunction with [and3] unavailable, CR available</td>
<td></td>
</tr>
</tbody>
</table>

According to H1 and H2, and is type-flexible, in which case direct DP conjunction with [and3] should be available. Under H1, CR is available in addition, while, under H2, CR is unavailable. H2 would amount to a claim that CR syntax, despite being available in principle, is for some reason not deployed in apparent object DP conjunction. According to H3, and is not type-flexible, so direct DP conjunction with [and3] is unavailable, while CR is available. To fully support the SIH we must rule out H1 and H2 in favor of H3. First, in Cases 1-4, I provide a range of data which CR can capture, but which the DP analysis cannot. These data provide evidence that CR must at least be available, ruling out H1, while leaving H2 and H3 as live possibilities. Finally, in Case 5, I report further data which are most naturally
understood under H3. The empirical tests for elided structure which I establish in this chapter will be utilized throughout the rest of the thesis.  

3 Case 1: adverbs

It has been observed that an adverb can linearly proceed a second apparent DP conjunct. I consider, in particular, the example in (55), where the temporal adverb *yesterday* precedes the DP *Chomsky* (after Collins 1988). I argue that this data point can only be analyzed with CR.

(55) **Adverb precedes the second conjunct**

John saw Labov and, yesterday, Chomsky.

Omitting *yesterday*, the DP analysis would assign to (55) the structure in (56), where *and* directly joins the DPs *Labov* and *Chomsky*. There are two ways that the adverb could integrate into the structure — and, as we will see, neither is viable.

(56) **Structure for (55) under the DP analysis (adverb omitted)**

\[ TP [TP [t1 saw [\&P [DP Labov] [and [DP Chomsky]]]]] \]

The first possibility is that *yesterday* adjoins on the clausal spine to the vP or TP, as in (57), where *yesterday* adjoins to the TP. To derive the word order, *Chomsky* would have to extrapose out of the conjunction to adjoin above the adverb.

(57) **Yesterday adjoined on the clausal spine**

\[ TP [TP [tp John saw Labov and t2 yesterday] [DP Chomsky]]2 \]

We can readily exclude this structure on syntactic and semantic grounds. Syntactically, extraposition would violate the CSC, since *Chomsky* asymmetrically moves out of one conjunct. 8 Semantically, the structure in (57) does not derive the right meaning for (55). In (57), *yesterday* takes scope over the entire conjunction, so the predicted meaning is that John seeing Labov and John seeing Chomsky both took place yesterday. That is, (55) should be equivalent to (58), where *yesterday* unambiguously attaches on the clausal spine and both (58-a) and (58-b) are entailed.

(58) **Entailments with yesterday adjoined to TP**

Yesterday, John saw Labov and Chomsky.

a.  \( \Rightarrow \) John saw Labov yesterday.

b.  \( \Rightarrow \) John saw Chomsky yesterday.

Intuitively, however, (55) says that John saw Chomsky yesterday, but leaves open when in the past John saw Labov. The entailment pattern is the one below, different from (58). To capture this, *yesterday* must scope just over the second conjunct in (55), ruling out (57).

---

7I am grateful to Irene Heim and David Pesetsky for help developing the diagnostic in Case 2, and Danny Fox for help developing the diagnostic in Case 3.

8Extraposition is A'-movement, rather than A-movement, so should be subject to the CSC (cf. fn. 3).
Correct entailments for (55)

John saw Labov and, yesterday, Chomsky.

a. ↦ John saw Labov yesterday.

b. ⇒ John saw Chomsky yesterday.

To scope over the second conjunct, yesterday must attach within that conjunct. If the second conjunct is just the DP Chomsky, this means that yesterday must adjoin to that DP. The second possible structure for (55) under the DP analysis is thus (60):

Yesterday adjoined to the DP in the second conjunct

\[TP \text{John}_1 [\text{VP} t_1 \text{ saw } [\&P [\text{DP Labov}] \text{ and } [\text{DP yesterday } [\text{DP Chomsky}]]]]\]

We already saw in Chapter 2 that yesterday is a Class 2 sentential operator and, as such, cannot adjoin to a DP (see the survey in Section 3.1.2). For this reason, the structure in (60) can be ruled out, as well. To make this clear, consider the minimal pair in (61), with (55) repeated as (61-b):

Distribution of yesterday

a. ??John saw yesterday Chomsky.

b. John saw Labov and yesterday Chomsky. = (55)

The deviance of (61-a) recalls the key distributional fact from Chapter 2. Since (61-a) is deviant, a structure like (62), with yesterday integrated into the DP as an adjunct, must be unavailable. If yesterday cannot adjoin to DPs in general, it should not be able to DP-adjoin in (55) either and, accordingly, (60) cannot be the right structure for (55).

Unavailable structure for (61-b)

\[TP \text{John}_1 [\text{VP} t_1 \text{ saw } [\text{DP yesterday } [\text{DP Chomsky}]]]\]

Minimal pairs like (61) making the same point can be constructed in a productive way. Two further examples are given below. In each case, Adv+DP is not a licit constituent in the (a) example and, as such, the adverb must not be DP-adjoined in the (b) example either.

Replicating the minimal triple

a. ?*John saw yesterday you.

b. John saw me and yesterday you.

Replicating the minimal triple

a. ?*John flew off to yesterday Paris.


Note that the (a) examples control for a parse where yesterday attaches on the clausal spine and the DP to its right has extraposed above it. As discussed in Chapter 2, extraposition is not possible with light DPs (such as you and Paris), and it cannot strand a preposition (as extraposition of Paris would in (66-b)).

Thanks to Susi Wurmbrand for bringing the pronoun examples to my attention.
Accordingly, the (a) examples isolate the possibility of a parse where *yesterday* is DP-adjoined, and their deviance shows that such a parse is unavailable.

To summarize so far, the DP analysis does not provide any viable structure for conjunction data with *yesterday* preceding a second apparent DP conjunct. If *yesterday* attaches on the clausal spine, the wrong scope results and, if it attaches within the second conjunct, as it must for scope, it is forced to adjoin to a DP, which it cannot do. Something different is needed.

What we need is a way for the adverb to attach within a second apparent DP conjunct, but not DP-adjoin. CR has precisely this effect. Illustrating with the original example in (55), CR assigns it the syntax in (65), where the vPs *t saw Labov* and *t saw Chomsky* are conjoined, and *yesterday* adjuncts to the vP in the second conjunct. Now, the structure is viable: *yesterday* takes the right scope and, due to the extra hidden syntax from CR, it attaches to a node with a sentential meaning.

(65) **CR structure for (55)**

\[
[TP\text{ John}_1 [\&P\ [vP t1 saw Labov] [and [vP yesterday [vP t1 saw Chomsky]]]])
\]

In turn, all of the contrasts between the (a) and (b) examples in the minimal pairs above are explained. The (a) examples must be parsed with *yesterday* DP-adjoined, while CR in the (b) examples furnishes an extra vP to host the adverb. All of the relevant structures are given together:

(66) **Capturing (61)**

a. \([TP\text{ John}_1 [vP t1 saw [DP yesterday [DP Chomsky]]]])

b. \(\checkmark [TP\text{ John}_1 [\&P [vP t1 saw Labov] [and [vP yesterday [vP t1 saw Chomsky]]]])\)

(67) **Capturing (63)**

a. \([TP\text{ John}_1 [vP t1 saw [DP yesterday [DP you]]]])

b. \(\checkmark [TP\text{ John}_1 [\&P [vP t1 saw me] [and [vP yesterday [vP t1 saw you]]]])\)

(68) **Capturing (64)**

a. \([TP\text{ John}_1 [vP t1 flew off to [DP yesterday [DP Paris]]]])

b. \(\checkmark [TP\text{ J}_1 [\&P [vP t1 flew off to Mars.] [and [vP yesterday [vP t1 flew off to Paris]]]])\)

This section has shown that Class 2 adverbs like *yesterday*, which cannot adjoin to DPs, are able to occur in a second apparent DP conjunct. As a result, the second conjunct must contain additional covert structure, including a node with a sentential meaning able to host *yesterday*. A parse with vP structure fits the bill — arguing that CR is available.

4 **Case 2: VP-Ellipsis**

I now provide a different kind of argument for CR. To re-iterate, the most fundamental syntactic difference between CR and the DP analysis has to do with how much structure is present in a second apparent DP conjunct. For the example in (69), the CR structure is (69-a), with a vP present in the second conjunct, while the structure under the DP analysis is (69-b), with only a DP.
(69) CR (a) vs. DP analysis (b)
Harvard invited Labov and Chomsky.

a. \[TP \text{Harvard}_1 \ T \ [v_P \ t_1 \ invited \text{Labov}] \ [and \ [v_P \ t_1 \ invited \text{Chomsky}]]\]

b. \[TP \text{Harvard}_1 \ T \ [v_P \ t_1 \ invited \ [\&P \ [DP \text{Labov}] \ [and \ [DP \text{Chomsky}]}}]\]

The goal in this section is to directly probe for the presence of vP structure in the second conjunct using an ellipsis-based diagnostic. As discussed in Section 2, VP ellipsis is licensed only when an appropriate antecedent for the elided VP is present in the linguistic context. I will demonstrate that the VP in the second conjunct in the CR structure — invited Chomsky in (69-a) — can serve as antecedent to license ellipsis of another VP. If invited Chomsky is an available antecedent, the extra structure associated with CR must be available.

4.1 VP ellipsis supportive of CR
Consider the data in (70). These examples bear a certain similarity to those in Section 3 in that an adverbial is present in the second conjunct. Instead of a simple adverb, however, there is now a complex adverbial clause. The critical feature is that the VP in the adverbial clause is elided. (70-a) is a minimal counterpart to (69), and (70-b) has a parallel profile.

(70) Introducing an adverbial clause with ellipsis

a. Harvard invited Labov and, ten years after Brandeis did \(\Delta\), Chomsky.
b. John resembles his mother and, though he would rather not \(\Delta\), his father.

Observe that the elided VP in the adverbial clause in each example (\(\Delta\)) is most naturally interpreted as the result of adding a verb to the second apparent DP conjunct. In (70-a), \(\Delta\) is most naturally interpreted as invited Chomsky, and in (70-b), \(\Delta\) is most naturally interpreted as resembles his father. This is captured in the paraphrases in (71) and (72).

(71) Paraphrasing ellipsis in (70-a)
Harvard invited Labov and — ten years after Brandeis invited Chomsky — Harvard invited Chomsky, too.

(72) Paraphrasing ellipsis in (70-b)
John resembles his mother and his father — though he would rather not resemble his father.

Intuitively, the sentences in (70) do not require an extra-sentential antecedent to be present to license ellipsis of invited Chomsky or resembles his father: the sentences are perfectly felicitous out of the blue. It must be, therefore, that an appropriate antecedent is present intra-sententially, just like in gapping.\(^{10}\)

\(^{10}\)Note that there are some examples where ellipsis is licensed without a linguistic antecedent. In most cases, these involve an antecedent that is extremely salient in the extra-linguistic context. This sort of "pragmatically-controlled" VP-ellipsis would arise, for instance, if someone points to a dance floor and says to their partner: "Shall we?" The elided VP may be interpreted as danced, even if danced is not linguistically present. For further data, see Schachter (1977, 1978), Stanley (2000), and Pullum (2001); their data are collected in Merchant (2004:718), along with additional examples. Since the cases I consider will not require a marked context, I take it that the relevant instances of VP ellipsis are linguistically controlled. Moreover, Merchant observed that pragmatic control is very difficult with stative predicates, and I will provide one example constructed with stative resemble. For additional skepticism of the view that VP ellipsis has an antecedence requirement, see Jacobson (2016 and
The VP in the second conjunct in the CR structures for (70-a) and (70-b) provides the necessary intrasentential antecedent. Illustrating with (70-a), the LF according to CR is (73), where the vPs invited Labov and invited Chomsky are conjoined, and the adverbial clause adjoins to the vP in the second conjunct. The internal syntax of the adverbial is shown separately in (73-b).

(73) **CR structure for (70-a)**

a. \[TP \lambda \lambda 1 [\&P [vP t1 invited L] [and [vP [CP ... (73-b) ...] [vP t1 [vP invited Chomsky]]]]]

b. \[CP ten years after [TP BrandeisF \lambda 2 [vP t2 [vP invited Chomsky \Delta]]]]

As in Section 2, I assume that ellipsis is licensed according to the Parallelism Condition in Takahashi & Fox (2005) (cf. (44)). Suppose in (73) that \( \Delta \) itself — invited Chomsky in (73-b) — is the PD. Since invited Chomsky elides in its entirety, given *ELIDEDFOCUS, it must be that invited Chomsky does not contain a focused element. The only focus alternative to invited Chomsky is, therefore, itself. It is no mystery, then, where the AC comes from: the AC is the VP invited Chomsky present in the second conjunct in (73-a). Accordingly, the VP invited Chomsky elides in the adverbial clause under identity with the VP invited Chomsky in the second conjunct. Assuming, as I have, that CR involves ellipsis of the VP in the second conjunct, note that the VP that serves as antecedent for \( \Delta \) is itself elided with Chomsky its only remnant. On the assumption made in Section 2 that parallelism is evaluated at LF and ellipsis takes place at PF, ellipsis does not affect satisfaction of parallelism.

### 4.2 The DP analysis cannot account for observed VP ellipsis

To make clear that data like (70) constitute an argument for CR, we must consider how (70) would be treated under the DP analysis. Omitting the adverbial clause, the DP analysis attributes to (70-a) the structure in (74), repeated from (69-b):

(74) **Structure for (70-a) under the DP analysis**

Harvard [\[vP invited [\&P [DP1 Labov] and [DP2 Chomsky]]]]

As seen in Section 3 with yesterday in (55) (John saw Labov and, yesterday, Chomsky), there are, in principle, two possibilities for where the adverbial clause in (70-a) could attach given the structure in (74): the adverbial clause could adjoin to the vP or TP on the clausal spine, or it could adjoin to the DP in the second conjunct. I argued in Section 3 that neither analysis was viable and the arguments extend to (70-a). Nevertheless, I here provide an additional independent argument against the DP analysis for (70-a) by focusing on the observed ellipsis in the adverbial clause: the interpretation of \( \Delta \) as invited Chomsky is incompatible with the DP analysis. Without the extra structure associated with CR, there is no appropriate antecedent to license ellipsis of invited Chomsky.\(^{11}\)

\(^{11}\)The argument from Section 3.1 against the adverbial clause adjoining on the clausal spine does carry over to (70-a) — but note that it does not carry over to (70-b). To illustrate, suppose that \( \Delta \) could be interpreted as invited Chomsky in (70-a) and resembles his father in (70-b). With the adverbial clause scoping above the conjunction, (70-a) would be predicted to be equivalent to (i), where the adverbial clause takes wide scope.

4.2.1 The adverbial clause adjoins to vP or TP

A possible structure for (70-a) with the adverbial clause adjoined on the clausal spine is (75). The adverbial clause is adjoined to the TP above the conjunction, and Chomsky extraposes out of the conjunction to adjoin above the adverbial clause, deriving the right word order.

(75) Adverbial clause adjoined on the clausal spine
   a. \[TP [TP [TP Harvard invited Labov and t_1] [CP ... (75-b) ...]] Chomsky_1]\n   b. \[CP ten years after [TP Brandeis_2 [VP invited Chomsky_2]]]\n
Ellipsis of invited Chomsky is not licensed in (75). The LF is (76), assuming Chomsky reconstructs into the matrix VP. Take the PD to be the VP invited Chomsky. The only VP present which can serve as the AC is the matrix VP containing the conjunction: the VP invited Labov and Chomsky. Since invited Labov and Chomsky is not semantically identical to invited Chomsky, parallelism is not satisfied. A similar problem arises for any other PD that could be chosen. 12

(76) Corresponding LF
   a. \[TP [TP Harvard [VP invited Labov and Chomsky] [CP ... (76-b) ...]]]\n   b. \[CP ten years after [TP Brandeis_2 [VP invited Chomsky_2]]]\n
Note that there is also an empirical way of showing that ellipsis in (70-a) is incompatible with (75). Consider a counterpart to (70-a) with the adverbial clause clearly scoping above the conjunction:

(77) Adverbial scoping above apparent DP conjunction
Ten years after Brandeis did \(\Delta\), Harvard invited Labov and Chomsky.

There is a strong preference for \(\Delta\) in (77) to be interpreted as invited Labov and Chomsky, rather than just invited Chomsky, contrary to the intuition in (70-a). In fact, there appears to be a generalization that when the adverbial takes scope above the antecedent and the antecedent contains a conjunction, \(\Delta\) is most naturally interpreted as also containing that conjunction. The intuition replicates with vP conjunction:

(i) is not, however, parallel to (70-a): whereas (i) says that Harvard invited each of Labov and Chomsky ten years after Brandeis invited Chomsky, (70-a) leaves open when in the past Labov was invited and says only that Chomsky was invited ten years after his Brandeis invitation. The adverbial clause must take scope in the second conjunct. (70-b) is not, however, clearly distinguishable from (ii):

(ii). Though John would rather not resemble his father, he resembles his mother and his father.

The new argument provided below is illustrated with (70-a), but extends to (70-b): the DP analysis is not, in fact, compatible with \(\Delta\) being interpreted as invited Chomsky in (70-a) or resembles his father in (70-b).

In principle, parallelism could be satisfied if the DP Labov and Chomsky QR-ed out of the matrix VP and the DP invited Chomsky QR-ed out of the VP in the adverbial clause:

(i). a. \[TP_4 [TP_3 [TP_2 [TP Harvard [VP invited t_1]] \lambda_1] [DP Labov and Chomsky]] [CP ... (75-b) ...]]\n   b. \[CP ten years after [TP [TP_5 [TP_4 Brandeis_2 [VP invited t_2]] \lambda_2] [DP Chomsky]]]\n
The PD would be TP_6 (focus alternatives of the form \(\lambda x \cdot y\) invited \(x\), where \(y\) is a replacement of Brandeis) and the AC would be TP_2 (denoting \(\lambda x \cdot Harvard invited x\). If this strategy were available, however, it would be unclear how to constrain the interpretation of \(\Delta\): ellipsis of any VP of the form invited \(y\) could similarly be licensed by QR-ing \(y\) out of the VP in the adverbial clause at LF. I will assume the following correspondence between ellipsis at PF and evaluation of parallelism at LF: any element which elides at PF must be included in the PD at LF. Since Chomsky in (70-a) elides, it must be included in the PD, contrary to the case in (i).
Ten years after Brandeis did A, Harvard fired its president and dismissed its chancellor. This example conveys that Harvard fired its president and dismissed its chancellor ten years after Brandeis did the same two things: there is a preference for A to be fired its president and dismissed its chancellor, rather than just fired its president or just dismissed its chancellor. Given the intuition in (77) and the broader generalization, the readily available interpretation of A as invited Chomsky in (70-a) is not consistent with the adverbial clause taking scope above the conjunction.

4.2.2 The adverbial clause adjoins to the DP in the second conjunct

The remaining possibility under the DP analysis has the adverbial clause adjoin within the second conjunct to DP2. The structure would be:

(79) Adverbial clause adjoined within the second conjunct

- Harvard1 [vP invited [&P [DP1 Labov] [and [[CP ...(79)... [DP2 Chomsky]]]]]
- [CP ten years after [TP Brandeis [VP invited-Chemsky A]]]

Given (79), there is a problem of antecedent containment. To illustrate, suppose that the PD is the elided VP. The only VP available to serve as antecedent for A is again the matrix VP. With the adverbial clause adjoined within the &P, however, the matrix VP is now invited Labov and, ten years after Brandeis did A, Chomsky, and so properly contains A. Since two VPs cannot be identical if one properly contains the other, the matrix VP cannot serve as antecedent for A. The problem of antecedent contained ellipsis (ACE) is familiar from examples like:

(80) Basic example of ACE
Polly visited every town (which) Eric did A.

The elided VP is contained in a relative clause (which Eric did A), which itself is contained within the DP object (every town which Eric did A) in the only VP available to serve as antecedent for the elided VP (visited every town which Eric did A). Since ellipsis is clearly licensed in (80), the grammar must make available ways to resolve ACE. One approach involves movement (e.g. Sag 1976). In (81), the DP every town which Eric did QRs to a position external to the matrix VP (movement is shown as rightward for convenience):

(81) ACE resolves with QR

[T3 [T2 [TP1 PollyF [VP visited t1]] λ1] [DP every town which Eric did A]]

13Note that, at first blush, the observation in (77) appears inconsistent with CR, which attributes to (77) the structure in (i) and thus predicts invited Chomsky to be an available antecedent for A.

(i). a. [TP [CP ... (i-b) ... [TP Harvard1 T [vP t1 invited Labov] [and [vP t1 invited Chomsky]]]]]
   b. [CP ten years after [TP Brandeis T [vP t1 invited-Chomsky A]]]

The broader generalization supported by (78) resolves this apparent objection to CR: even when VPs are overtly conjoined, A includes the conjunction when the adverbial clause takes highest scope.
The relative clause has the structure in (82), where which moves from the complement of visit to its pronounced position. Eric contrasts with Polly in the matrix clause, so is focused.

(82) **Structure for the relative clause**

\[ cp \text{ which } [ T_{P5} \lambda x \ [ T_{P4} \text{EricF } [ v_{P} \text{visited t}_{2} ]] ] \]

Take the PD to be TPs, in which the trace in the elided VP is bound. Focus alternatives to TPs are predicates of the form \( \lambda x \cdot y \text{ visited } x \), where \( y \) is some replacement of Eric. TP3 — the node in the matrix clause just below the landing site of the QR-ed DP — is an appropriate AC, as it denotes the predicate \( \lambda x \cdot \text{Polly visited } x \). ACE is resolved, and ellipsis is licensed.

Is there a way to resolve ACE in (79) to yield an interpretation of the elided VP in the adverbial clause as invited Chomsky? In fact, there is not. It is possible to resolve ACE by moving the DP Labov and, ten years after Brandeis did \( \Delta \), Chomsky to a position external to the matrix VP:

(83) **ACE resolves with QR of the &P**

a. \[ TP \ [ T_{P1} \text{Harvard } [ v_{P} \text{invited t}_{1} ] \] \&P L \[ cp \text{ ten years after } T_{P} \text{BrandeisF } [ v_{P} \text{invited Chomsky } \Delta ] ] \]

b. \[ cp \text{ ten years after } [ T_{P} \text{BrandeisF } [ v_{P} \text{invited Chomsky } \Delta ] ] \]

This is not, however, sufficient to license ellipsis. Supposing that the PD is the elided VP, the only candidate AC remains the matrix VP, which is now invited \( t_{1} \). This is not, however, an appropriate AC, since invited \( t_{1} \) is not equivalent to invited Chomsky under any variable assignment \( g \) where \( g(1) \neq \text{Chomsky} \). A parallel problem arises with any other PD that could be chosen.

4.3 Local summary

The observed VP ellipsis in (70) provides direct support for the availability of CR: the VP in the right conjunct in the CR structure is necessary to provide an intra-sentential antecedent to license the observed VP ellipsis. The DP analysis, regardless of where the adverbial clause attaches — on the clausal spine, or to the DP in the second conjunct — is not compatible with the data.

5 Case 3: Split scope

The previous two cases both consider examples with an adverbial in the second conjunct, either a simple adverbial (yesterday in Case 1), or a complex adverbial clause (ten years after Brandeis did in Case 2). This leaves open the question: is CR generally available in the absence of an adverbial? I now argue that it is using a semantic diagnostic involving scope. I demonstrate that the conjunction and can take scope above an operator while the DPs and apparently conjoins scope below that same operator. These “split scope” readings are predicted with CR, but are not viably derived with direct DP conjunction.

5.1 The split scope signature

The sentence in (84) contains four scope operators: the conjunction (and), the two quantifiers (any city in Europe and any city in Asia), and the intensional predicate refuse. Of particular interest is the scope of the conjunction and the quantifiers relative to refuse.
(84) **Split scope prototype**
John refused to visit any city in Europe and any city in Asia.

The sentence in (84) is acceptable and has two entailments: (i) that John refused to visit any city in Europe, and (ii) that John refused to visit any city in Asia. This paraphrases as (85-a), and is stated more formally in (85-b), where *refuse* is a modal operator quantifying over worlds compatible with what John is willing to do in the evaluation world, ‘W(John)(wₐ)’.

(85) **Scope: and > refuse > any city in Europe, any city in Asia**
   a. John refused to visit any city in Europe and he refused to visit any city in Asia.
   b. \[\neg\exists w' \in W(John)(w₀) [\exists x [x is a CIE in w' \land John visits x in w']] \land \neg\exists w'' \in W(John)(w₀) [\exists y [x is a CIA in w'' \land John visits x in w'']]\]

Under this reading, *and* and the quantifiers scope at different heights relative to *refuse*. On the one hand, *and* scopes above *refuse*, as the modalization contributed by *refuse* occurs separately in each conjunct. On the other hand, the quantifiers scope below *refuse*. This is required for NPI *any* to be licensed, and is clear from (85-b), as the existentials contributed by the quantifiers are interpreted de dicto within the scope of the universal modal in each conjunct. Hence, scope is “split”: *and > refuse > any city in Europe, any city in Asia*. The split scope signature can be stated generally:

(86) **Split scope signature**
   *And* scopes above some operator, which the apparent DP conjuncts scope below.

A further illustration of split scope comes from the example in (87), due to Irene Heim (p.c.). Building on Buring (2007) and Heim (2008), I take *little* to decompose into negation and *much*. To make this explicit, let us start with (88), simplified from (87) by removing the conjunction.

(87) **Replicating split scope**
   This plant is easy to take care of! It needs little water and little sunlight.

(88) **A single conjunct from (87)**
   This plant needs little water.

The sentence in (88) paraphrases as (89-a), and the reading is formally (89-b), where the set of worlds compatible with what the plant needs in the evaluation world are notated ‘N(p)(w₀)’. As is clear in (89-b), the negative component of *little* scopes above *need* and the *much* component below.

(89) **Scope: not > need > much**
   a. It’s not the case that this plant needs much water.
   b. \[\forall w' [w' \in N(p)(w₀) \rightarrow \text{the plant receives much water in } w']\]

The natural reading of the original sentence in (87) is the one paraphrased in (90). *And* scopes above *need*, and within each conjunct, the negative component of *little* also scopes above *need*, while the *much*
component scopes below. That is, and takes widest scope, and the individual conjuncts each have the same profile as (89).

(90) **Scope: and > not > need > much**

a. It's not the case that this plant needs much water and it's not the case that this plant needs much sunlight.

b. \( \neg \forall w' [w' \in N(p(w_0)) \rightarrow \text{the plant receives much water in } w'] \)
\( \wedge \neg \forall w'' [w'' \in N(p(w_0)) \rightarrow \text{the plant receives much sunlight in } w''] \)

This reading instantiates the split scope signature: and scopes above need while the much component of little in little water and little sunlight scopes below: and scopes above an operator and a component of the DPs and seems to conjoin scopes below the same operator.

Rooth & Partee (1982) observe a parallel split scope reading with the disjunction in (91), which can be interpreted as in (92). In (92), or scopes above look for, while a maid and a cook are interpreted de dicto in the scope of look for.

(91) **Split scope in disjunction**

John is looking for a maid or a cook.

(92) **Scope: or > look for > a**

a. John is looking for a maid or he is looking for a cook.

b. \( \forall w' \in L(\text{John}(w_0)) [\exists x [x \text{ is a maid in } w' \wedge \text{John finds } x \text{ in } w']] \)
\( \wedge \forall w'' \in L(\text{John}(w_0)) [\exists y [y \text{ is a cook in } w'' \wedge \text{John finds } y \text{ in } w'']] \)

Rooth & Partee, however, suggested that conjunction does not allow split scope on the basis of (93-a), which lacks the reading in (93-b), where and scopes above hope and a maid and a cook scope below. An empirical contribution here is to demonstrate that conjunction does exhibit split scope. I will explain the absence of a split scope reading in (93) later in the chapter.

(93) **Conjunction example lacking split scope**

a. John hopes that some company will hire a maid and a cook.

b. John hopes that some co. will hire a maid and he hopes that some co. will hire a cook.

5.2 The CR analysis

What is needed to derive split scope is a mechanism for the conjunction to scope at a higher position than the quantifiers. CR offers such a mechanism. Illustrating with (84), for and to scope above refuse, the conjoined vPs are refused to visit any city in Europe and refused to visit any city in Asia.

(94) **Ingredient 1 for split scope: and > refuse**

\[ TP \text{John } \lambda \jmath T [\& p [vP1 t1 \text{ refused PRO to visit any city in Europe} ] \]
\[ \text{and } [vP2 t1 \text{ refused PRO to visit any city in Asia}]]\]
The second ingredient for split scope is for the quantifiers to scope below refuse. This straightforwardly obtains if any city in Europe QRs to a position below refuse in the first conjunct and any city in Asia QRs to a position below refuse in the second conjunct.

(95) **Ingredient 2 for split scope: refuse > any city in Europe, any city in Asia**

$$[TP \lambda 1 \lambda t_1 \text{refused} [TP [\lambda 2 \lambda t_2 \text{PRO to visit } t_2]]$$

$$\text{and } [\lambda t_1 \text{refused} [TP [DP \lambda 3 \lambda t_3 \text{PRO to visit } t_3]]]$$

5.3 The DP analysis is insufficient

The DP analysis assigns the structure in (96), with any city in Europe and any city in Asia directly conjoined. Assuming that any city in Europe and any city in Asia are ordinary quantifiers of type <et, t>, and is interpreted as \( \lambda \text{and}_3 \), and the &P itself has the quantifier meaning in (97).

(96) **Structure for (84) under the DP analysis**

John refused to visit \([\&P [DP \text{any city in Europe}] \text{and } [DP \text{any city in Asia}]]\)

(97) **Quantificational meaning for &P**

a. \( \lambda f_w \exists x [x \text{ is a CIE in } w \land f(x)] \land \exists y [y \text{ is a CIA in } w \land f(y)] \)

b.

Being a quantifier, the &P takes scope via QR. Because the &P contains both and and the DPs and QRs as a constituent, and necessarily scopes at the same height relative to refuse as any city in Europe and any city in Asia. A first possibility is that the &P QRs to a position above refuse, for instance, by adjoining to the matrix TP, as in (98). The predicted interpretation given this structure is the one in (99) where and and the quantifiers all scope above refuse.

(98) **QR targets a position above refuse**

$$[TP [\&P \lambda 1 \lambda t_1 \text{refused} [TP \lambda 2 \lambda t_2 \text{PRO to visit } t_2]]]$$

(99) **Predicted: and > any city in Europe, any city in Asia > refuse**

$$[(98)]^w = 1 \iff \exists x [x \text{ is a CIE in } w \land \exists y [y \text{ is a CIA in } w \land f(x)]] \land \exists y [y \text{ is a CIA in } w \land f(y)]$$

The reading in (99) is distinct from the target split scope reading, as the quantifiers are interpreted de re. Paraphrasing, (99) says that John refused to go to particular places, at least one of which happens to be a city in Europe and at least one of which happens to be a city in Asia. This reading is, in fact, unavailable in (84), due to the licensing requirements of the NPI any. In order to be licensed, any must be in the scope of refuse at LF, and in (98)-(99), any is out of the scope of refuse.

Alternatively, the &P could QR to a position below refuse, as in (100), where the &P adjoins to the embedded TP. The reading is (101), with and and the quantifiers all taking narrow scope.

(100) **QR targets a position below refuse**

$$[TP \lambda 1 \lambda t_1 \text{refused} [TP [\&P \lambda 2 \lambda t_2 \text{PRO to visit } t_2]]]$$
(101) Predicted: refuse > and > any city in Europe, any city in Asia

$$[[\text{W}]](\text{W}) = 1 \iff \exists w' \in W(j)(w) \left[ \exists x \in \text{CIE} \in w' \land \text{John visits x in w'} \right] \land \exists y \in \text{CIA} \in w' \land \text{John visits y in w'}$$

As with split scope, the quantifiers in (101) are de dicto in the scope of refuse. The difference from split scope is that and also scopes below refuse. The result: (101) says that John refused for there to be both a city in Europe that he visits and a city in Asia that he visits together. This is compatible with John being willing to visit a city in Europe so long as he does not also visit a city in Asia, and vice versa. The reading in (101) is logically weaker than the split scope reading — and is again not available due to the licensing requirements of NPI any. In (100)-(101), and takes scope between refuse and the quantifiers, and and is an intervener for NPI licensing (Linebarger 1987, Chierchia 2004, Guerzoni 2006). The intervention effect is established in (102), due to Guerzoni (2006:360):

(102) Intervention in NPI licensing

*I didn’t drink a cocktail and any soda.

In this example, the negation is above the vP, so regardless of whether the string is parsed under the DP analysis, as in (103-a), or with CR as vP conjunction, as in (103-b), negation takes scope over the conjunction, so and intervenes between negation and NPI any. The ungrammaticality of (102) shows that any is not licensed in an intervention configuration.

(103) DP analysis (a) and CR (b) structures

a. I didn’t drink $[\&P [DP \text{ a cocktail}] \land [DP \text{ any soda}]]$

b. $[TP I \text{ didn't} \&P [vP t_1 \text{ drink a cocktail}] \land [vP t_1 \text{ drink any soda}]]$

Hence, whether the &P QRs above refuse or below refuse, the split scope reading is not derived. Either the conjunction and the conjoined DPs all scope above refuse, or they all scope below. Neither reading is split scope and, in fact, neither reading is available due to NPI licensing requirements.

5.4 Ruling out alternatives

Before closing, I want to recognize two ways we might still try to derive split scope without CR. One is a variant of the DP analysis, which interprets the conjuncts not as regular quantifiers, but as even higher-order predicates. The other is based on a parse involving mereological sums of the sort noted at the beginning of the chapter (Section 1.3). Neither alternative, I believe, is truly viable.

5.4.1 Alternative 1: type-lifting the quantifiers

In discussing the DP analysis, I assumed that any city in Europe and any city in Asia are of type <et,t>. Yet, there is another possibility to consider. Type-shifting mechanisms have been proposed which could raise them to a higher type. By lifting the quantifiers, it is possible to derive split scope in accord with the DP analysis — but the derivation runs into other problems. Let us see.

For this discussion, it will be important to represent the world as an argument, rather than a parameter. To do this, I will use the second system presented in Chapter 2, where occurrences of t in the basic quantifier denotations are replaced with <s,t>, i.e. quantifiers are of type <est,st>, as in (104). The
type-raising operation of interest is, then, the Montague Lift, which, adapted for the intensional system, lifts a meaning of type \( \alpha \) to one of type \( \ll \alpha, \text{st}, \text{st} \ll \). The Montague lifted versions of the quantifiers are of type \( \ll \ll \text{est}, \text{st}, \text{st} \ll \). Conjoining these with the appropriate high-type version of \textit{and} results in a conjunction of type \( \ll \ll \text{est}, \text{st}, \text{st} \ll \).

(104) **Basic intensional meanings**

a. \([\text{any CIE}] = \lambda_{f_{\text{est}}} \cdot \lambda w. \exists x [x \text{ is a CIE in } w \land f(x)(w)]\)
b. \([\text{any CIA}] = \lambda_{f_{\text{est}}} \cdot \lambda w. \exists y [y \text{ is a CIA in } w \land f(y)(w)]\)

(105) **Lifted meanings for quantifiers**

a. \([\text{any CIE2}] = \lambda F_{\ll \text{est}, \text{st}, \text{st}} \cdot \lambda w. \phantom{} F([\text{any CIE}](w))\)
b. \([\text{any CIA2}] = \lambda F_{\ll \text{est}, \text{st}, \text{st}} \cdot \lambda w. \phantom{} F([\text{any CIA}](w))\)

(106) **Interpreting direct DP conjunction**

\([\&P] = \lambda F_{\ll \text{est}, \text{st}, \text{st}} \cdot \lambda w. \phantom{} F([\text{any CIE}](w)) \land F([\text{any CIA}](w))\)

Now, the LF in (107) can deliver split scope. The \&P undergoes a first step of QR within the scope of \textit{refuse}, and then QRs above \textit{refuse}. The lower trace (\( t_1 \)) is of type \( e \), which allows it to compose with the verb. The higher trace (\( t_2 \)) is of type \( \ll \text{est}, \text{st} \).

(107) **LF for split scope with lifted quantifiers**

\([\text{TP5} \ W \ [\text{TP4} \ &P \ [\text{TP4} \ &P \ [\text{TP4} \ \text{refuse} \ [t_2 \ [\lambda \ 1 \ [\text{TP1} \ \text{PRO to visit } t_1]]]]]]]]\)

Given this structure, the second step of movement derives a property of quantifier intensions, as in (108), and applying the \([\&P]\) to that predicate in (109) in fact does derive split scope.

(108) **Interpreting \([\text{TP2}]\)**

\([\text{TP2}] = \lambda Q_{\ll \text{est}, \text{st}, \text{st}} \cdot \lambda w. \neg \exists w' \in W(j)(w) [Q(\lambda x. \lambda w'' \cdot \text{John visits } x \text{ in } w'')(w')]\)

(109) **Overall meaning**

a. \([\text{TP4}] = [\&P]\([\text{TP3}]\))
b. \([\text{TP4}] [\text{any CIE}] \land [\text{TP3}] [\text{any CIA}]\)
c. \(= \lambda w. \neg \exists w' \in W(j)(w) [\ll \text{any CIE}](\lambda x. \lambda w'' \cdot \text{J visits } x \text{ in } w''(w'))\)
\(\land \neg \exists w' \in W(j)(w) [\ll \text{any CIA}](\lambda x. \lambda w'' \cdot \text{J visits } x \text{ in } w''(w'))\)
d. \(= \lambda w. \neg \exists w' \in W(j)(w) [\exists x [x \text{ is a CIE in } w' \land \text{John visits } x \text{ in } w']\)
\(\land \neg \exists w' \in W(j)(w) [\exists x [x \text{ is a CIE in } w' \land \text{John visits } x \text{ in } w']\)

The effect is that \textit{and} takes widest scope and, within each conjunct, the original quantifier — \([\text{any CIE}]\) or \([\text{any CIA}]\) — is semantically placed in the position of \( t_2 \), below \textit{refuse}. In this way, \textit{and} takes widest scope and the quantifiers end up de dicto below \textit{refuse}.\(^{15}\)

\(^{15}\)We must assume that NPI \textit{any} is licensed so long as the quantifiers are interpreted in a downward entailing environment, even if they are not in the syntactic scope of the licenser.
Although this derivation does yield split scope, there are reasons to reject it. First, it has been controversially proposed that type-shifting is available only as a last-resort option when required to resolve a type-mismatch (Partee & Rooth 1983). Montague lifting any city in Europe and any city in Asia does not contribute to resolving any type-mismatch, so would be blocked. More compellingly, there is empirical evidence against the derivation. It is critical to note that the derivation disentangles syntactic position and semantic scope: the &P is syntactically above refuse, but type-lifting and high-type traces result in [any CIE] and [any CIA] be interpreted with narrow scope below refuse. Fox (1998, 1999) and Romero (1998) discuss Condition C data which rule out this profile. Fox, for instance, observes the contrast:

(110) **Evidence against semantic reconstruction: Condition C**

a. A new theory by him₁ seems to Quine₁ to be needed.

b. *A new theory by Quine₁ seems to him₁ to be needed.

A new theory by him/Quine has A-moved to the matrix subject position, but is interpreted de dicto below seem and need. The deviance of (110-b), due to Condition C, shows that the quantifier cannot semantically take narrow scope without also syntactically reconstructing below him. A derivation with high-type traces and semantic reconstruction, like that entertained for split scope, would allow the quantifier to be interpreted de dicto, but still be syntactically in the matrix subject position. So, the fact seems to be that a quantifier cannot semantically take narrow scope below an operator without also being syntactically below that operator — ruling out derivations like (107).₁⁶

Given its theoretical and empirical shortcomings, I conclude that (107) is not a viable means of deriving split scope. Overall, then, split scope cannot be derived with direct DP conjunction with variants of logical and: interpreting the quantifiers with their basic type fails to derive split scope at all, and lifting them does not seem to derive split scope in a viable manner.

5.4.2 Alternative 2: sums and homogeneity

Could split scope involve a parse different from any of the ones we have been considering?₁⁷ I noted at the outset of the chapter that a conjunction of quantifiers may be interpreted as if there were a single quantifier taking a set of sums as its restrictor. For any student and any professor, this would amount to existential quantification over the set of pairs of a European city and an Asian city:

(111) **Quantification over sums**

\[
\lambda f_{ref} \cdot \exists X \exists x \exists y \left[ \text{CIE}(x) \land \text{CIA}(y) \land X = x \oplus y \right] \land f(X)
\]

An LF for (84) with the conjunction taking scope below refuse is repeated in (112). If the conjunction is interpreted as (111), this LF corresponds to the meaning in (113). On the face of it, (113) does not capture split scope: it says that John refused for there to be a city in Europe as well as a city in Asia that he visits, but leaves open that he might visit one or the other.

₁⁶Some work subsequent to Fox and Romero has questioned the conclusion that syntactic and semantic scope always correlate, and argued that semantic reconstruction mechanisms must exist (see Sharvit 1998, Lechner to appear, Keine & Poole 2017). All authors agree, however, that semantic reconstruction is not able to feed the binding of a world variable by an intensional predicate to manufacture a de dicto reading. Semantic reconstruction may create a "non-specific de re" reading, where a quantifier reconstructs for scope, but any world variables are still indexed to the actual world.

₁⁷I thank Maribel Romero for bringing this point to my attention.
Possible LF for (84)

\[
[T_P \text{ J refused } [T_P \text{ [\&p any CIE and any CIA] } \lambda 1 [T_P \text{ PRO to visit t_1]}]]
\]

Corresponding meaning

\[
\neg \exists w' \in W(j)(w_0) [\exists x [\exists y [x \text{ is a CIE in } w' \land y \text{ is a CIA in } w' \land X = x \oplus y] \land \forall a <_A' X [\text{John visits a in } w']]]
\]

Yet, there might be more going on. Predicating of a plural entity is known to result in a homogeneity effect. To build up, consider the sentences in (114) with definite plurals:

Homogeneity with definite plurals

a. John visited these cities.

b. John didn’t visit these cities.

These cities denotes the maximal salient plurality of cities. (114-a) is true if John visited all of these cities — but, (114-b) does not simply deny that. Rather, (114-b) is true if John visited none of these cities. If John visited some, but not all of the cities, neither sentence is true or false. This is the homogeneity effect, stated in (115) (from Kriz 2015:4).

The homogeneity effect

A predicate is neither true nor false of a (plural) individual X if it is true of some parts of X and false of other parts of X.

There are different ideas about how homogeneity should be captured, but one traditional approach encodes homogeneity as a presupposition of distributive predicates (Schwarzschild 1993, Löbner 2000, Gajewski 2005).

Now, suppose that a homogeneity presupposition is triggered when visit is predicated of X in (113), and that the presupposition projects universally to yield a global presupposition that for every world w’ in W(John)(w_0), and for every pair X of a European city and an Asian city, John either visited all atoms of X in w’ or no atoms of X in w’. Adding homogeneity to (113) delivers an overall meaning that mimics split scope. (113) says that there is no W-world at which John visits all atoms of X and homogeneity would strengthen this to say that there is no W-world at which John visits any atoms of X, for any X. This is equivalent to: there is no W-world at which John visits a city in Europe and there is no W-world at which John visits a city in Asia — the split scope reading.

While homogeneity may derive a reading mimicking split scope in the particular test example I have used for illustration, this is not a general solution to the problem of split scope. First, as noted, Rooth & Partee (1982) observed split scope with disjunction, which cannot create sums. Moreover, there is conjunction data that cannot be handled with sums. As discussed in Section 1.3, a hydra parse with a single interpreted quantifier ranging over sums is only possible if both DPs in the conjunction show the same quantifier. With distinct quantifiers, the hydra parse is blocked. Even with the hydra parse blocked, however, split scope survives. Consider:18

\footnote{I thank Itai Bassi for drawing my attention to examples like this.}

96
Another case of split scope
John refused to visit Paris and any city in Asia.

Because both conjuncts do not show NPI *any*, this sentence should not have a hydra parse. Yet, scope is still split. The sentence is interpreted as (117) and, furthermore, for the NPI to be licensed, *and* must not intervene between *refuse* and the NPI, meaning *and* must take scope above *refuse* and the NPI below. There must be some mechanism to derive split scope without sums, and CR fits the bill. The CR structure in (118) derives the reading and the NPI is licensed.

Interpretation of (116)

a. "John refused to visit Paris and he refused to visit any city in Asia."
   \( \neg \exists w' \in W(\text{John})(w_0) [\text{John visits Paris in } w'] \\
   \land \neg \exists w'' \in W(\text{John})(w_0) [\exists x [x \text{ is a CIA in } w'' \land J \text{ visits } x \text{ in } w'']] \)

CR parse for split scope

\[ [TP \ \text{John} \ \lambda \ 1 \ T [\&P \ [vP_1 \ t_1 \ \text{refused} \ [TP \ \text{PRO to visit Paris}]] \land [vP_2 \ t_1 \ \text{refused} \ [TP \ \text{PRO to visit any city in Asia}]]]] \]

More evidence that CR is the mechanism for split scope in this profile of data comes from constraints on where split scope is possible. Consider (119) (for similar data, see Chierchia 2013:376):

Another case of split scope

??I doubt that you will visit Paris and any city in Asia.

The licenser for NPI *any* in this case is *doubt*. To avoid an intervention configuration, *and* must scope above *doubt*, while the NPI remains below. The expected interpretation, then, is the split scope reading: "I doubt that you will visit Paris and I doubt that you will visit any city in Asia". CR would derive this reading from the structure in (120):

Hypothetical split scope structure with CR

\[ [TP \ 1 \ \lambda \ 1 \ T [\&P \ [vP_1 \ t_1 \ \text{doubt} \ [CP \ \text{that you will visit Paris}]] \land [vP_2 \ t_1 \ \text{doubt} \ [CP \ \text{that you will visit any city in Asia}]]]] \]

The deviance of (119) suggests that split scope is marginalized or unavailable. While this might look surprising under a CR analysis, let us consider: what makes the CR structure for (116) different from the CR structure for (119)? The most striking difference is that the right conjunct in (116) includes a non-finite clause boundary, while the right conjunct in (119) includes a finite clause boundary. If CR involves the same mechanism as gapping, it should, in fact, be impossible across a finite clause boundary, since gapping has exactly that restriction. Consider:

Gapping across finite and non-finite clauses

a. John refused to visit Paris and Mary London.
   
b. *I doubt that you visited Paris and Mary London.
The conjunction in (121-a) may scope above refuse, with the gapped material is the matrix vP refused to visit. By contrast, in (122-b), the conjunction cannot scope above doubt, with the gapped material the matrix vP doubts that I will visit. The star on (122-b) signifies that it is ungrammatical on that reading, as there is an alternative available reading where the conjunction scopes in the embedded clause and the gapped material is just will visit. Just as matrix gapping is impossible in (121-b), a CR structure for split scope should be impossible in (119), predicting its deviance. With a hydra parse controlled, split scope can arise — and tracks expected constraints on CR.

Overall, while a hydra parse with homogeneity might mimic split scope in some cases, I conclude that genuine split scope is nonetheless attested. CR must be available to derive split scope, and this idea is supported by constraints on where split scope is and is not observed.

5.5 Summary and next steps

This concludes the first wave of empirical arguments: Cases 1-3 have shown that CR is required to host adverbs like yesterday in a non-initial conjunct (Case 1), to license VP ellipsis (Case 2), and to capture split scope (Case 3). Where does this leave us with respect to the an overall argument for the SIH? I have shown that in an environment where the SIH predicts covert syntax, that syntax is at least available. Apparent object DP conjunctions can be parsed with covert structure creating a type t node at which and can scope, interpreted as [and] — in accord with the SIH. It remains an open possibility, however, that CR and the DP analysis co-exist. Next, in Case 4, I will discuss further scope data that are most straightforwardly understood if CR is available — and if direct DP conjunction with [and3] is unavailable. Although, in the terminology from the first chapter, and appears to be a Class 1 operator, on closer examination it really looks like a Class 2 operator. If there is reason to believe that and is not cross-categorial, the fundamental tenet of the SIH as it applies to and is substantiated.

6 Case 4: A missing scope reading

I return to the scope data presented at the outset of the chapter as evidence against classical CR. My aim is to turn these on their head to become some of the strongest evidence for the SIH. Although indeed not predicted under classical CR, I show that certain scope data are also not predicted under the DP analysis. The DP analysis derives unavailable readings, just as classical CR did. While we could introduce stipulations to block unattested readings, I show that the data are directly predicted under the SIH. If the DP analysis is simply unavailable, it cannot over-generate — and a vP-level CR parse, which I have proposed is available, does not over-generate either. Scope facts about apparent object DP conjunction will reduce to independent scope facts about vP conjunction (along with facts about the backgrounded hydra parse, which again does not over-generate).

6.1 A missing scope reading

The example of special concern will be (122), which we have seen in a couple of places (Section 2 and Section 5.1) and, as noted, is adapted from Rooth & Partee (1982). The scope operators are the subject quantifier (some company), the conjunction (and), and the object DPs (a maid, a cook).
A missing scope reading

Some company hired (both) a maid and a cook.

Since the subject and the objects are all existential, their relative scope does not affect interpretation: existentials are commutative with respect to one another. The critical scope relation is between some company and the conjunction. The important observation is that, of the two possible scope orders, only one is attested. There is an available reading where some takes scope above and. On this reading, (122) conveys that some single company hired both a maid and a cook, as in (123).

Available: some > and

a. “Some company hired two people: a maid and a cook.”

b. $\exists x [\text{company}(x) \land \exists y [\text{maid}(y) \land x \text{ hired } y \text{ in } w_0]$
   $\land \exists z [\text{cook}(z) \land x \text{ hired } z \text{ in } w_0]]$

On the other hand, the inverse reading, where and scopes above some, is unavailable. Paraphrased in (124), this reading would convey that some company hired a maid and some potentially different company hired a cook. Rooth & Partee report that reading to be unavailable in the earlier version of the example in (93-a), where (122) is embedded under doubt and moving to a matrix context should not alter the judgment. The vast majority of informants corroborate the intuition (and I will suggest an explanation for the few who do not later).

Unavailable: and > some

a. “Some company hired a maid and some company hired a cook.”

b. $\exists x [\text{company}(x) \land \exists y [\text{maid}(y) \land x \text{ hired } y \text{ in } w_0]$
   $\land \exists x' [\text{company}(x') \land \exists z [\text{cook}(z) \land x' \text{ hired } z \text{ in } w_0]]$

6.2 The DP analysis over-generates

While the DP analysis is able to derive the attested scope order, an over-generation problem arises: the DP analysis can derive both the attested and unattested readings. The basic structure under the DP analysis is (125), where and directly conjoins a maid and a cook.

Structure with the DP analysis

$[TP \text{ some company hired } [\&P \text{ [DP a maid]}] \text{ [DP a cook]]}]$

Quantifier meaning for &P

a. $[[\&P]^w = [[\text{and}_3]]([\text{a maid}]^w)([\text{a cook}]^w)$

b. $= \lambda f_{x'}. \exists x [\text{maid}(x) \land f(x)] \lor \exists y [\text{cook}(y) \land f(y)]$

With and interpreted as $[\text{and}_3]$, the conjunction, a maid and a cook, has a quantifier meaning, and QRs. The two readings derive from different heights of QR. One possibility is for the &P to QR below some company, as in (127-a). In that case, the attested scope reading, some > and, obtains. The unattested scope reading, and > some, obtains just as easily by QRing the &P above some company, as in (127-b). In this way, the DP analysis over-generates the unattested reading.
(127) **QR targeting different positions**

a. \([TP \text{ some company } \lambda_1 T \ [\&P \text{ a maid and a cook } \lambda_2 \ [TP_1 \ t_1 \ hired \ t_2]]]\)

b. \([TP_2 \ [\&P \text{ a maid and a cook } \lambda_1 \ [TP_1 \ \text{ some company } \lambda_2 \ [\&P \ t_2 \ hired \ t_1]]]\)

If the DP analysis is available, QR must be constrained to block the derivation in (127-b). What might such a constraint look like? Since at least certain object quantifiers can take scope above a subject existential (first observed in May 1977, 1985), a general constraint on QR is not supported. The data point in (128), for instance, illustrates the well known possibility for an object universal quantifier to take wide scope. The example is ambiguous:

(128) **Scope ambiguity**

Some company hired every employee.

a. "Some single company hired every employee."  

b. "Potentially different companies hired every employee."

Nonetheless, it has been observed that not all quantifiers have the same scope possibilities (see e.g. Szabolcsi 1994, 1999, Beghelli & Stowell 1997, Liu 1990, de Swart 1998, Collins & Postal 2014, Collins 2016). For an illustration, the data in (129) lack an inverse scope reading (example (129-a) is due to Collins & Postal 2014:173). (129-a) says that everyone is ignorant about physics, not that no physics concept is universally understood. In a similar vein, (129-b) says that there is a picky student, not that most subjects are generally unpopular.

(129) **Scope freezing**

a. Everybody knows nothing about physics.  

b. A student likes few subjects in school.

Does the failure of an &P to take wide scope fit with independent constraints on where inverse scope is possible? As far as I can see, the answer is negative. Collins & Postal (2014) flag one generalization, which unifies the examples in (129): quantifiers that are monotone decreasing, such as no and few, do not allow for inverse scope. Since \([&P]\) is upward montonic, scope rigidity of the &P is not predicted from this generalization. Indeed, with respect to monotonicity, \([&P]\) is identical to every \(N\), which we saw does allow for inverse scope. Beghelli & Stowell (1997) give a richer taxonomy of five types of quantifiers, and propose that each of them moves to a particular functional projection, due to feature-checking requirements. Scope is determined by the height of the corresponding functional projection. The five classes are: interrogative quantifiers (e.g. what, which \(N\)), negative quantifiers (e.g. no), distributive universals (e.g. every \(N\), each \(N\)), counting quantifiers (e.g. few, more), group-denoting quantifiers (e.g. a, some). From the data in (129), the classes which seem not to take wide scope are negative quantifiers and counting quantifiers — and, clearly, \([&P]\) would not fit into either category.

If the DP analysis is available, the only option is to introduce special stipulations to specifically block an &P from QRing above a subject. In Baghelli & Stowell’s system, &Ps would have to bear a special feature specification that matches with some functional projection below the highest position at which the subject can be interpreted. While I cannot rule out such stipulations, they lack independent motivation.
I argue that the data are more naturally understood if the DP analysis is simply unavailable: the scope pattern follows as a prediction of the SIH, given vP-level CR.

Overall, because [and₃] creates a quantificational &P, the properties it should show are those of a quantifier. This subsection has demonstrated that the &P expected with [and₃] does not share a property of at least some quantifiers: their scopal mobility. In effect, where we might expect to see reflexes of [and₃], none exist — and the SIH can predict that.

### 6.3 The CR analysis

Let us suppose that the SIH is correct and the DP analysis is unavailable. Then, vP-level CR makes the following correlational prediction with respect to scope in apparent object DP conjunction:

(130) **Conjunction Scope Prediction (‘CSP’):**
Scope in apparent object DP conjunction should track scope in overt vP conjunction

(modulo any additional scope readings from a hydra parse\(^{19}\)).

Scope in (122) is expected from the CSP. The relevant baseline is (131), which minimally differs from (122), and involves a conjunction of two overt vPs. Just like in (122), scope is rigid, with and necessarily taking wide scope over some. As far as I know, Moltmann (1992) was the first to observe scope rigidity in this configuration. For our purposes, the key result is this: scope in apparent object DP conjunction tracks scope in overt vP conjunction — as the CSP predicts.

(131) **Baseline: overt vP conjunction**
Some company hired a maid and fired a cook.

a. “Some single company hired a maid and fired a cook.”  
   \(\checkmark\) some > and

b. “Some co. hired a maid and some co. fired a cook.”
   \(*\) and > some

How exactly the scope facts are explained analytically is essentially irrelevant. The point is that whatever independently explains scope in overt vP conjunction will also explain scope in apparent object DP conjunction in kind, without any need for new specific stipulations.

Nonetheless, I will offer some analytical remarks. Parsed with CR, (122) involves a conjunction of two vPs, as in (132). Some company occurs in spec-TP, having undergone ATB movement out of both of the conjoined vPs. The available reading is directly predicted: because some company ATB moves, there is a single occurrence of some company, which takes wide scope over the conjunction: some company in spec-TP scopes over and, which conjoins lower vPs.

(132) **CR structure for (122)**

\[TP \text{some company } \lambda t_1 T [\&P [vP_1 t_1 \text{ hired a maid}] [\text{and} [vP_2 t_1 \text{ hired a cook}]]]\]

Given the structure in (132), deriving the unattested reading, and > some, would require the subject to undergo ATB reconstruction back into its base position within each conjunct, as in (133).

(133) **Structure with ATB reconstruction**

\[TP T [\&P [vP_1 \text{ some co. hired a maid}] [\text{and} [vP_2 \text{ some co. hired a cook}]]]\]

\(^{19}\)As I discuss in Section 6.2.1, this caveat is not relevant in this case.
The baseline example in (131) establishes that ATB reconstruction is blocked in this configuration. (131) wears on its sleeve the proposed CR structure for (122), as in (134). Just like with the CR structure for (122), to derive and > some in (133), some company would have to ATB reconstruct to its base position internal to the vPs. Since (131) cannot be interpreted with and > some, it is independently shown that ATB reconstruction is blocked.

\[(134) \quad \text{Structure for (131)} \]
\[\left[ TP \text{ some company } \lambda x \text{ T } [\&p \left[ vP_1 t_1 \text{ hired a maid} \right] [\&p_2 t_1 \text{ fired a cook}]] \right] \]

Constraints on ATB reconstruction have been considered in the literature, notably in Fox (1995, 2000). Applying Fox’s idea to (122) and (131), reconstruction is subject to an economy constraint which allows some company to ATB reconstruct into the two vPs only if, within each vP, some company reverses its scope relative to another operator with which some is non-commutative. The only other scope operator within vP_1 is a maid, and the only other scope operator within vP_2 is a cook. Since two existentials are commutative, the economy condition is not met.

To summarize, the CR proposal is sufficiently restrictive to account for the observed lack of and > some in (122) and reduces this fact to the same lack of and > some in (131). In conjunction with an independently needed theory of when ATB reconstruction can and cannot occur, CR predicts the missing scope reading in (122). Hence, if the SIH is correct, the DP analysis cannot over-generate, since it does not exist — and the proposed CR structure does not over-generate.

6.3.1 Aside: a comment on the hydra parse

Before moving on, I want to meditate on one further aspect of the Conjunction Scope Prediction: that the SIH predicts correlating scope between vP and apparent DP conjunction, modulo additional scope readings from a hydra parse. A hydra parse is presumably available in addition to CR, and we can ask: why does it not contribute a reading mimicking and > some in (122)? On a hydra parse, a maid and a cook would be interpreted as though a single existential quantifier ranged over maid-cook pairs:

\[(135) \quad \text{Quantifying over sums} \]
\[a. \quad \left[ \text{a maid and a cook} \right]^w \]
\[b. \quad = \lambda f_{\eta} . \exists x \exists y \left[ \text{maid}(x) \land \text{cook}(y) \land x \sqsubseteq y \right] \land f(x) \]

With a maid and a cook so interpreted, a reading mimicking and > some could derive from the LF in (136). The structure contains a distributive operator. That operator could attach in different places, depending on exactly how it is defined. I show it counter-cyclically merged between a maid and a cook and the binder index inserted with movement. The distributive operator is interpreted as in (137), and yields the overall meaning: “for some maid-cook pair, every atom of the pair was hired by some (potentially different) company”. This is equivalent to: “some company hired a maid and some company hired a cook” — the unattested and > some reading.
(136) **Possible LF**

\[
\begin{align*}
\text{TP}_4 & \quad & \&P \\
\text{TP}_3 & \quad & \text{TP}_2 \\
& \quad & \lambda 1 \\
\text{a maid and a cook} & \quad & \text{Dist} \\
\text{some company saw } t_1 & \quad &
\end{align*}
\]

(137) **Interpreting Dist**

\[
[\text{Dist}]'' = \lambda f_{et} \cdot (X \cdot \forall a <_{AT} X [f(a)])
\]

(138) **Predicted meaning**

\[
\begin{align*}
\text{a. } [\text{TP}_4]'' &= [\text{a maid and a cook}]''([\text{Dist}]''([\text{TP}_2]'')) \\
\text{b. } &= \exists X \exists x \exists y [\text{maid}(x) \land \text{cook}(y) \land X = x \oplus y] \\
& \land \forall a <_{AT} X [\exists z [\text{company}(z) \land z \text{hired a in w}]]
\end{align*}
\]

While a reading mimicking \textit{and} \textgreater \textit{some} could derive this way in principle, there is independent evidence that the configuration in (136) is unavailable. The sentence in (139), for instance, with a definite plural as the object also lacks a distributive reading for most speakers.

(139) **Baseline: definite plural**

Some company hired these cooks.

\[
\begin{align*}
&\text{a. } \checkmark \text{"Some single company hired each cook."} \\
&\text{b. } \ast \text{"Each cook was hired by a potentially different company."}
\end{align*}
\]

The unavailability of a distributive reading establishes that (139) cannot have an LF like (136), where \textit{these cooks} takes the position of the \&P, QRed above the subject with a distributive operator integrated. Just as that LF is unavailable for (139), it should be unavailable for (122) in kind. While I do not understand \textit{why} the configuration does not arise, the scope fact in (122) is again part of a broader generalization. Since a reading mimicking \textit{and} \textgreater \textit{some} does not arise from a hydra parse, we can set that parse aside in the rest of the discussion.

6.3.2 **Rooth & Partee's (1982) example**

In the remainder of this section, I want to show how the CR analysis can predict scope not only in (122), but rather open up our inquiry to a broader range of related data, and show that CR continually makes correct predictions. To start, I will return to the example in (140), which is the original example from Rooth & Partee (1982), upon which (122) is based. Repeated from (93-a) above (see Section 5.1), this example embeds the sentence in (122) under the matrix verb \textit{hopes}.
Progenitor of (122)

John hopes that some company will hire a maid and a cook.

Note that the addition of hopes does not create any interesting new ambiguities. Being universal, hope is commutative with and. Moreover, the existentials in the embedded clause cannot QR above hope, since QR is well known to be impossible out of finite clauses (May 1985). The point of interest, therefore, remains the relative scope of the now embedded subject, some company, and the conjunction. Rooth & Partee’s judgment is the same as I reported for (122): some company rigidly scopes above and. The sentence can convey that John hopes for a single company to hire a maid and a cook, but not that he hopes a maid and a cook to get hired by potentially different companies.

By combining results from this section and Section 5, we are now in a position to fully see how the scope pattern follows from CR. There are, in principle, two possible CR structures for (140). First, CR could take place entirely inside the embedded clause. Then a structure just like the one we saw for (122) in (132) above occurs underneath hope, as in:

(141) Embedded CR for (140)

\[
\begin{align*}
TP & \text{ John hopes } CP \text{ that } TP \text{ some company } \lambda l \text{ will } [&P [vP t1 \text{ hire a maid}]
\end{align*}
\]

Some company, having ATB-moved to the specifier of the embedded TP, takes scope above the lower conjunction of vPs. As already discussed, ATB reconstruction is blocked in this configuration due to independent constraints. Hence, this structure faithfully predicts the available scope order (some > and), and does not derive the unavailable one (and > some).

The second possible structure has CR to the level of the matrix vP, instead of the level of the embedded vP, as in (142). Since the conjunction is higher than the final position of some company, this structure does have and scope above some without any need for ATB-reconstruction.

(142) Matrix CR for (140)

\[
\begin{align*}
TP & \text{ John1 [&P [vP t1 \text{ hopes that some company will hire a maid}]
\end{align*}
\]

As discussed in Section 5.2.2, however, CR is expected to be constrained in the same ways as gapping, and gapping is impossible across a finite clause boundary. That constraint would render the structure in (142) impossible, since the conjoined vPs each contain the finite clause boundary after hope. Hence, there is simply no viable way to derive the unattested reading for (140) under CR: embedded CR derives and > some, and matrix CR is blocked. CR is empirically adequate.

It is important to highlight a couple of key points in relation to Rooth & Partee. First, the impossibility of and > some is not an argument in favor of the DP analysis. Just like with (122), the DP analysis would over-generate and > some in (140) by allowing a maid and a cook to QR as a quantifier over some company inside the embedded clause. The proposed vP-level CR analysis, with the DP analysis unavailable, offers the most successful account of this data point, as far as I can see. Moreover, recall from discussion in Section 5.2.2 that Rooth & Partee took (140) as evidence against the existence of split scope with conjunction. The discussion here and in Section 5 shows that this was not the right test example: if the mechanism for split scope is vP-level CR, this example would not show split scope,
because the matrix predicate embeds a finite clause. In examples where a non-finite clause is embedded, split scope was observed in Section 5.

6.3.3 Where and > some is observed

So far, we have shown how CR is appropriately restrictive to not derive and > some in cases where that reading is absent. Now, I want to show that CR is also appropriately flexible to derive some > and in cases where that reading is observed. Fox (1998, 2000) observes that a guard can undergo ATB-reconstruction back into the conjoined vPs in the overt vP conjunction in (143).

(143) **Baseline: overt vP conjunction**

A guard is standing in front of every church and sitting beside every mosque.

Fox’s economy constraint allows ATB-reconstruction in this case, since the two conjoined vPs each contain a universal quantifier. The existential a guard is non-commutative with the universals every church and every mosque, so ATB reconstruction is licensed, provided that a guard scopes below every church in the left conjunct and below every mosque in the right conjunct. That is, Fox predicts the sentence in (144) to allow the reading in (145), which he observes is, in fact, available.

(144) **Available reading: and > every church, every mosque > a guard**

\[
\forall x \ [\text{church}(x) \rightarrow \exists y \ [\text{guard}(y) \land y \text{ is standing in front of } x]] \\
\land \forall x' \ [\text{mosque}(x') \rightarrow \exists z \ [\text{guard}(z) \land z \text{ is sitting beside } x']] 
\]

The counterpart example in (145) with apparent DP conjunction similarly allows a reading with a guard taking narrowest scope, consistent with the CSP. The relevant reading is one where for every church, a guard is standing in front of it, and for every mosque, a guard is standing in front of it, where all the guards are potentially different (and > every church, every mosque > a guard).

(145) **Parallel with “DP conjunction”**

A guard is standing in front of every church and every mosque.

Moreover, Zamparelli (2011) observes that and > some is available with the apparent DP conjunction in (146-a): (146-a) can convey that potentially different bullets killed the two victims.

(146) **“DP conjunction” allowing a > and**

A 9mm bullet killed the first victim and the second victim.

The availability of and > a in (146) is not immediately predicted by Fox’s economy constraints, since the existential subject (a 9mm bullet) and the definite objects (the first victim, the second victim) are not commutative. Recall, however, that the point of this section does not crucially depend on economy: the point is that whatever restricts ATB reconstruction in overt vP conjunction, the same observed restrictions obtain with apparent DP conjunction. Example (146) is consistent with this, since the counterpart in (147) with overt vP conjunction similarly allows and > some.
Overall, vP-level CR has made very fine-grained correct predictions about scope: it is adequately restrictive to not derive and > some in (122) and (140), but it also does predict and > some in the data in this subsection; moreover, it was able to derive split scope exactly where it was observed in Section 5. CR, with no interference from an additional DP analysis, predicts scope.

6.3.4 Local summary

We started with a puzzle: that (122) disallows and > some. We then saw that the DP analysis overgenerates the unattested reading: [and3] composes with a maid and a cook to create an object quantifier, and object quantifiers can generally scope above an existential subject. Failure of the quantificational &P to take wide scope does not fit with any independent constraints on QR, so would require specific stipulations. In effect, [and3] creates a quantificational &P, which must be stipulated not to have expected properties of quantifiers without independent support for the stipulations.

I showed that the data are expected under the SIH, given vP-level CR. The DP analysis does not exist — and CR predicts scope from an independent generalization: that overt vP conjunction shows the same scope rigidity that (122) shows. I noted that a hydra parse also does not over-generate, since definite plurals again show the same scope rigidity. The proposed CR analysis predicted a range of additional scope facts, including where wide scope of and was observed.

When I initially presented (122), I noted that a couple of informants did allow for and > some. To re-iterate, the SIH makes a correlational prediction. In so far as there is inter-speaker variation in judgments, the SIH predicts that a given speaker’s judgment with apparent object DP conjunction should track their judgment with overt vP conjunction or a definite plural, not their judgment with quantifiers. All informants seemed to agree that and > some was at least more marginal in (122) than in a quantifier baseline. It would be a fruitful future direction to test the predicted correlations experimentally based on data from a large sample of speakers.

6.4 Ending where we began

Before closing this subsection, I want to look beyond apparent object DP conjunction, and end the chapter where I began it: with Partee’s (1970) original argument against classical CR. As discussed at the outset (see Section 2), Partee (1970) observed that (148-a) is not equivalent to the full clausal conjunctions in (148-b) and (148-c):

(148) Partee’s original paradigm

a. Few rules are explicit and easy to read.
b. Few rules are explicit and few rules are easy to read.
c. Few rules are explicit and it’s easy to read few rules

While this is an argument against classical CR, the existence of type t nodes lower in the structure, along with ideas we have seen in this section, reconcile the data point with the SIH. The profile of these data
closely resembles (122). In the conjunction in (148-a), the subject, *few rules*, necessarily takes wide scope relative to the conjunction: the sentence says that few rules have together the properties of being explicit and easy to read. In (148-b) and (148-c), conversely, *and* scopes above *few*. With the perspective on CR adopted in this chapter, (148-a) may be analyzed as a conjunction of vPs or even APs, as in (149). *Few rules* originates as an argument of *explicit* in the first conjunct, and separately as an argument of *read* in the second conjunct. Both APs are of type t.

(149) **Base structure for (148-a)**

\[\text{[TP } T \text{ are } [\text{AP few rules explicit} \ [\text{and } [\text{AP few rules easy to read few rules}]]]\]

*Few rules* undergoes tough-movement within the second conjunct to the edge of the AP\(^{20}\), and then *few rules* undergoes ATB movement out of both conjuncts to spec-TP.

(150) **LF for (148-a)**

\[\text{[TP few rules T are } [\text{AP t explicit} \ [\text{and } [\text{AP t1 easy to read t1}]]]\]

Due to ATB-movement, the structure in (150) derives the available *few > and* reading. *And > few* cannot derive since *few rules* is blocked from reconstructing. In Fox’s approach, because there is no scopal operator internal to either conjunct, ATB-reconstruction cannot satisfy economy. It thus follows that (148-a) is not equivalent to (148-b) or (148-c). The logic is very much the same as what we saw with (122) under the CR analysis: the subject takes wide scope due to ATB-movement, and narrow scope is missing due to the impossibility of ATB-reconstruction. To re-iterate, the LF in (150) allows *and* to be interpreted as *[and] — accordant with the SIH.*

7 Conclusion

In this chapter, I have undertaken a close study of cases where *and* appears to conjoin object DPs, interpreted as *[and] 3*. The SIH raised the expectation that these data should involve covert syntax creating a type t scope site for *[and] — and this chapter has provided a wide range of arguments that this is so. In particular, I have argued that apparent object DP conjunction involves underlying vP-conjunction, similar to gapping.\(^{21}\) The arguments that a vP-level CR structure is available are recapitulated below. In at least one profile of data, then, the covert syntax which the SIH predicts is available.

- CR “follows for free” from independent syntactic mechanisms: Johnson’s mechanism for gapping, combined with ATB movement of the subject out of the conjoined vPs.
- CR creates a host for sentential adverbs like *yesterday* in a non-initial conjunct.
- The extra vP structure associated with CR is required to furnish an intra-sentential antecedent to license VP ellipsis in certain cases.

---

\(^{20}\)I assume a movement analysis of tough-constructions, but this is controversial and not crucial.

\(^{21}\)Given the possibility of sluicing and fragment answers, which have been taken to involve TP-ellipsis, a question arises as to what prevents a parse with underlying full clausal conjunction. One possibility is that speakers can only access the parse with *and* scoping at the lowest possible type t scope site (cf. Schein 2017). It is important to emphasize that this question arises regardless of whether or not *and* is type-flexible.
CR provides the only viable account of split scope, where *and* scopes above some operator and the DPs that *and* apparently conjoins scope below that same operator.

While the points just listed all argue that CR is available, CR and direct DP conjunction with \([\text{and}_3]\) could both be available, in principle. The final part of the chapter provided direct evidence that a parse with *and* conjoining the DPs as \([\text{and}_3]\) is unavailable. The point was:

- DP conjunction with \([\text{and}_3]\), if available, would over-generate scope readings.

Although the basic data presented in the introductory chapter made it appear as though *and* is able to compose with DPs, the over-generation problem provides a strong argument that *and* actually cannot. Referring to the taxonomy between Class 1 and Class 2 operators, this chapter supports re-categorizing *and* as Class 2, rather than Class 1 — just as the SIH predicted. A richer syntax, imperfectly mapped to phonology, creates the illusion that *and* operates on nodes not of type t.
Chapter 4
Constructing pseudo-clefts

1 Introduction

I now pursue the SIH with a construction where the presence of sentential nodes has been a lively point of controversy: specificational pseudo-clefts. These are built from a copula. In the pre-copular position, there is an XP, which in English shows wh-morphology. Syntactically, XP is either an interrogative CP or a free relative DP. In the post-copular position, there is some YP having the same syntactic category as the pre-copular wh-word. In (1), the pre-copular XP is what Obama approved, what is a DP, and the DP this bill appears post-copular. The post-copular material is called the pivot.

1 A specificational pseudo-cleft

What Obama approved was this bill.

The critical starting observation in this chapter is that the pivot in a pseudo-cleft can be a conjunction. This is illustrated in (2), where this bill is replaced with this bill and that bill.

2 Puzzle for the SIH

What Obama approved was this bill and that bill.

On the face of it, this bill and that bill forms a constituent, but that constituent does not seem to contain any type t nodes. Indeed, this bill and that bill appears to be a direct DP conjunction:

3 Most obvious syntax

\[
[TP \ [XP \text{what Obama approved}] \ [VP \text{was \ [DP this bill and that bill]]}]\]

According to the SIH, the conjunction in (3) cannot be interpreted, at least not with a type-lifted variant of logical and. One possibility might be that and can only be interpreted as an unrelated sum formation operator in this case. Data presented over the course of the chapter, however, will diagnose a parse with logical and. Granting such a parse, we are led to ask: is there covert syntax in (2) creating type t nodes for and to conjoin? One hypothesis might take and to scope very high, above the pseudo-cleft. In that case, the overt material what Obama approved was this bill would be the first conjunct, while that bill would be the remnant of a second clausal conjunct, perhaps another full pseudo-cleft:

4 Hypothesis: wide scope CR

\[
[&P \ [TP_1 \ [DP \text{what Obama approved}] \ [VP \text{was \ [DP this bill]]}] \ [\text{and} \ [TP_2 \ [DP \text{what Obama approved}] \ [VP \text{was \ [DP this bill]]}]])\]

There is, however, a strong basis to reject this idea. The first problem has to do with how pseudo-clefts are interpreted. They are typically interpreted exhaustively, such that (1), for instance, conveys that Obama approved this bill and nothing else. If so, a conjunction of two full pseudo-clefts is predicted to
yield contradictory exhaustivity inferences. The first conjunct in (4) would convey (5-a), and the second conjunct (5-b), contradictorily.

(5) Expected inferences from (4)
   a.  ⇒ Obama approved this bill and nothing else.  \(\text{(from TP}_1\text{)}\)
   b.  ⇒ Obama approved that bill and nothing else.  \(\text{(from TP}_2\text{)}\)

The data point in (6) wears on its sleeve a conjunction of two full pseudo-clefts. Intuitions are somewhat variable, but it is clear that (6) is not the correct paraphrase for (2). For some speakers, (6) is plainly contradictory. For other speakers, the expected exhaustivity inferences seem to be defeasible, but all speakers I consulted report a sense of “squeamishness” in (6) not felt in (2).

(6) String corresponding to (5)
  ?#What Obama approved was this bill, and what he approved was that bill.

In order to explain why there is no hint of contradiction in (2), it appears that (2) must involve a single pseudo-cleft, with and scoping within its post-copular constituent. Though incompatible with the SIH, the original structure in (3) is right in this respect: and is post-copular. A different kind of evidence for and being post-copular comes from a related construction:

(7) Definite description counterpart to (2)
   a.  The things Obama approved were this bill and that bill.
   b.  *The thing Obama approved was this bill and that bill.

The example in (7) is similar to (2), except that the pre-copular constituent is replaced with a regular definite description. This substitution has an instructive morphological consequence. Whereas what Obama approved does not overtly inflect for number, the definite description in (7) does. The key observation is that number is obligatorily plural. If the structure were (8) (analogous to (4)), the overt copula and pre-copular definite description would be in construction just with this bill, and number would be singular. A second copula and definite description, also singular, would be elided.

(8) Candidate CR structure for counterpart
    \([&P \ [TP_1 \ [DP \ \text{the thing Obama approved}] \ [VP \ \text{was} \ [DP \ \text{this bill}]]]]
    \[\text{and} \ [TP_2 \ [DP \ \text{the thing Obama approved}] \ [VP \ \text{was} \ [DP \ \text{that bill}]]]]

The possibility of plural number shows that the conjunction can take scope inside the post-copular constituent, and the necessity of plural shows that it must scope there. Wide scope of and is simply unavailable, likely due to exhaustivity-induced contradiction. On the basis of the semantic and morphological facts just presented, I reject the hypothetical CR structure in (4).

The inescapable conclusion is that (2) is parsed with and inside the post-copular constituent. On the one hand, this leaves us in a difficult situation, since there are not obviously any type t nodes post-copular. On the other hand, having narrowed down the location of and, we can make our question more precise: are there type t nodes post-copular, despite appearances?

The question we have reached by reasoning about conjunction was previously reached in an unrelated context: the well-known puzzle of pseudo-cleft connectivity (Akmajian 1970, Ross 1972, Higgins 1976).
The puzzle is this: the post-copular constituent in a pseudo-cleft behaves as though it were c-commanded by an element within the pre-copular constituent, even though that c-command relation does not hold. Example (9) illustrates with anaphor binding:

(9) **Illustrating the connectivity puzzle**

What Obama\(_i\) approves of is himself\(_i\).

According to traditional Binding Theory, anaphor licensing is a structural matter, dependent on c-command. Condition A says that an anaphor must be c-commanded by a co-indexed DP within the minimal TP containing the anaphor. If the structure for (9) contains just the material that appears overtly, as in (10), himself\(_i\) is apparently licensed without Condition A satisfied: the only potential antecedent is Obama in the pre-copular constituent, and Obama does not c-command himself\(_i\).

(10) **Most obvious syntax**

\[
[TP [XP what Obama\(_i\) approves of] [VP is [DP himself\(_i\)]]]
\]

To solve the puzzle, a number of authors proposed covert syntax (Ross 1972, den Dikken et al. 2000, Schlenker 2003). Their idea was that the overt pivot is the remnant of a full clause, the rest of which is elided. The post-copular constituent in (10) is not just himself\(_i\), but Obama approves of himself\(_i\), as in (11). With this additional clausal structure, a second, unpronounced occurrence of Obama is introduced, which does c-command himself\(_i\).

(11) **Covert syntax**

\[
[TP [XP what Obama\(_i\) approves of] [VP is [TP Obama approves of himself\(_i\)]]]
\]

Whether or not this covert syntax really exists has been a point of controversy. Contrary to the authors above, Heycock & Krock (1999), Jacobson (1994), Sharvit (1999), and Caponigro & Heller (2007) have proposed that the original syntax in (10) is the correct one. The dominant tactic for connectivity on this side of the debate has been to deny the linguistic relevance of c-command and to pursue purely semantic accounts of anaphor binding and related phenomena.

The question of whether there is clausal structure post-copular is closely connected to the question I raised with respect to the SIH: are there type t nodes post-copular? If there is a post-copular clause, then there are type t nodes, and the data reconcile with the SIH. The best candidate SIH-compatible syntax for (2) would be (12). There is a full clause post-copular and that clause has a vP-level CR syntax of the sort proposed in Chapter 3.

---

1. Sharvit (1999) argues that not all pseudo-cLEFTs may be analyzed with ellipsis, but does not rule out the possibility that some may have an ellipsis parse (see also Sharvit 2009). That view is consistent with the syntactic facts I will present in Section 2: they demonstrate that ellipsis is possible, not necessary. Still, the SIH commits to the stronger claim that an ellipsis parse is the only one for any pseudo-cleft with a logical conjunction post-copular, and the semantic analysis I motivate in the second part of the chapter is crucially dependent on there being a full clause post-copular. Caponigro & Heller (2007) argue that pseudo-clefts are never parsed with ellipsis, which is inconsistent with the data in Section 2. I am not certain at present how to fully reconcile Sharvit's and Caponigro & Heller's observations with my own analysis, and I must leave that to future work.
My aim is to develop and defend a particular analysis of pseudo-clefts compatible with the post-copular syntax in (12). Since my primary topic of concern is the SIH, I will confine most of my attention to examples which contain a conjunction. The chapter will proceed in two steps, previewed as follows.

### 1.1 Part 1: Pursuing the SIH

I will provide novel arguments that pseudo-clefts can (at least optionally) be parsed with covert syntax, supporting the aforementioned idea due to Ross (1967), den Dikken et al. (2000), and Schlenker (2003). My strategy will be to extend some of the new diagnostics for covert syntax developed in Chapter 3 to the pseudo-cleft environment. The tests will, in particular, involve placing adverbials into (2):

(13) **Adverb data**

What Obama approved was this bill and, with difficulty, that bill.

I will argue that certain adverb data, including (13), can only be analyzed with a post-copular clause having CR structure. With respect to (13), much as we saw in Chapters 2 and 3 with *yesterday*, the PP *with difficulty* can adjoin on the clausal spine, but not to DPs. The second conjunct in (13) must, therefore, include structure beyond just the overt DP in order to host the adverb.

### 1.2 Part 2: Fleshting out the LF

Having established the key result that the post-copular constituent is a full clause, I will set about to flesh out the LF to create an interpretable structure. The goal of the chapter, then, is not only to defend the SIH, but to fit the associated post-copular structure into a full analysis of pseudo-clefts. In Part II, my proposal will depart quite radically from previous analyses in the covert syntax tradition.

The pervasive intuition is that pseudo-clefts are an intra-sentential pairing of a question with its answer. The pre-copular XP has a question meaning, while the post-copular TP expresses a proposition, the answer to the question. The basic idea is due to Ross (1972), and was refined by Schlenker (2003) who proposed a detailed composition. His idea was that the copula asserts *identity* between the *strongest*
true answer to the pre-copular question and the post-copular proposition. A non-conjunctive pseudo-cleft would be interpreted:

(14) Identity semantics for pseudo-clefts

\[
\text{[[what Obama approved]]} = 1 \text{ iff } [\text{ANS}] (\text{[[what Obama approved]]}) = \lambda w . \text{Obama approved this bill in } w
\]

The identity semantics is successful in this case, and predicts an exhaustivity inference, since the post-copular proposition is the strongest true answer to the question. I will show, however, that an identity semantics runs into difficulties with certain adverb data. The example in (15) provides a particularly clear illustration of the problem:

(15) Adverb data with possibly

What Obama approved was this bill and possibly that bill.

The underlying post-copular clause would express that Obama approved this bill and it's possible he approved that bill. This seemingly can never be identical to the strongest true answer to the pre-copular question, which would be the proposition that Obama approved X for the maximal plurality X of bills that Obama approved. The identity semantics should thus predict a contradictory meaning. Intuitively, the data point is contingent, and conveys that Obama approved this bill, he possibly approved that bill, and he approved no bill other than this or that.

I abandon an identity semantics altogether, and propose a new LF for pseudo-clefts, illustrated in (16). The key idea is that the compositional “glue” semantically connecting the pre- and post-copular constituents is not the copula, but rather a covert operator with a meaning similar to only. This operator, Exh, is familiar from the scalar implicatures literature (e.g. Chierchia 2006, Fox 2007, Chierchia, Fox, & Spector 2012). The copula is inert, and the semantic burden is placed squarely on Exh.

(16) Full LF for a pseudo-cleft

```
                   t
                  /   \
                <st,t>    TP
                  /     \    
              Exh    XP   t
               /    \      
             <stt,<st,t>> <st,t>
             /         /
         what Obama approved
```

When I presented only in Chapter 2, I noted that its interpretation requires a proposition (its ‘prejacent’, p) and a set of propositions, ALT. At that point, I was concerned with the prejacent, and indicated ALT as a parameter, but recognized that it could also be an argument. Here, I take it to be an argument. Overt only and, likewise, covert Exh thus operate on two arguments: ALT and the prejacent. Whereas only
presupposes the truth of the prejacent, Exh asserts it: Exh(ALT)(p) asserts that p is true, and negates excludable alternatives in ALT.

In typical cases, ALT is provided by a covert variable. I observe, however, that pseudo-clefts overtly wear on their sleeve both arguments of Exh. Following Ross, I take the pre-copular XP to have a question meaning and, following Hamblin (1973) and Karttunen (1977), I take a question to be a set of propositions. The intension of the post-copular TP is a proposition. [XP] has the right semantic type to play the role of ALT, and [YP] can play the role of the prejacent. I propose that [XP] and [YP] are both directly arguments of Exh, and play exactly those roles. Like an identity semantics, the proposal predicts exhaustive meanings, but the source of exhaustivity has changed: it is no longer an ANS operator, but rather Exh. This approach handles the example with possibly as well, once facts about how only and Exh interact with possibly are uncovered.

1.3 Roadmap for the chapter

The proposal will posit covert syntax in the post-copular position to reconcile the conjunction data with the SIH, empirically motivate that covert syntax on the basis of adverb data, and build around the post-copular clause a novel LF for pseudo-clefts that offers an empirically adequate analysis of the adverb data. Part I is undertaken in Section 2, and the remainder of the chapter is devoted to Part II. In Section 2, I consider the post-copular constituent in detail, and provide the adverb-based arguments that there is a full clause with CR syntax. In Section 3, I lay out Schlenker’s (2003) identity semantics in detail and then in Section 4 spell out the problem that the adverb data raise. In Section 5 and 6, I introduce my proposal with Exh, and address further data in Section 7. Finally, Section 8 concludes.

2 Diagnosing covert syntax

In this section, I address in detail the syntax of the post-copular constituent in (2). Are there just the overtly detectable DPs, or is there a richer clausal structure with a CR syntax? The DP analysis and the CR structure are repeated side-by-side:

(17) **Structure 1: DP analysis**

\[
\text{TP [DP what Obama approved] [VP was [\&P [DP this bill] [and [DP that bill]]]]}
\]

(18) **Structure 2: CR analysis**

\[
\text{TP [XP what Obama approved] [VP was [TP Obama] [\&P [\text{t1 approved this bill}]
\text{[and [\text{t1 approved that bill}]]]}}
\]

To test for CR, I will home in on the second conjunct, and attempt to diagnose \text{VP} structure within that conjunct by extending the adverb-based tests from Chapter 3. Just as those tests revealed CR structure in basic conjunction data, they will reveal hidden post-copular CR here.

2.1 Test 1: Adverb distribution

The first test for post-copular ellipsis is based on the distribution of adverbs. As we saw in Chapters 2 and 3, certain adverbials can adjoin on the clausal spine, but not to DPs. Here, I show that these
adverbials can be hosted within the post-copular constituent in a pseudo-cleft. This is predicted only if the post-copular constituent contains hidden clausal structure.

2.1.1 Choosing the test adverbial

In Chapter 3, I drew upon one class of adverb with the right distributional profile: temporal adverbs such as yesterday. In constructing examples in this section, I will make use of a different adverbial with the same distribution: the adverbial PP with difficulty. The paradigm in (19) serves to establish the crucial distributional facts:

(19) Distribution of with difficulty
   a. With difficulty, John flew off to Paris.
   b. John, with difficulty, flew off to Paris.
   c. *John flew off to, with difficulty, Paris.
   d. John flew off to Paris, with difficulty.

Examples (19-a), (19-b), and (19-d) are perfectly acceptable and, in these cases, with difficulty can be parsed as adjoining on the clausal spine. Available structures for these data follow:

(20) Available: with difficulty+TP or vP
   a. \([TP [PP \text{ with difficulty}] [TP \text{ John flew off to Paris}]]\) \((=19-a)\)
   b. \([TP \text{ John}, [TP [PP \text{ with difficulty}] [vP \text{ t1 flew off to Paris}]]]\) \((=19-b)\)
   c. \([TP [TP \text{ John flew off to Paris}]] [PP \text{ with difficulty}]\) \((=19-d)\)

The odd one out is (19-c), which lacks such a parse. If the PP adjoined on the clausal spine, the derivation would have to involve extraposition of the DP Paris to a position above the PP. But, extraposition would strand the preposition to, and is blocked for this reason:

(21) Unavailable structure for (19-c)
    \([TP [TP [TP \text{ John flew off} [PP \text{ to t1}]] [PP \text{ with difficulty}] [DP \text{ Paris}]]]\) \(\ast\) for (19-c)

Critically, there is another structural possibility to entertain for (19-c): the PP with difficulty could adjoin to the DP Paris, and that PP+DP constituent would remain in situ, as in (22). The fact that (19-c) is ungrammatical shows that this, too, is not a licit structure.

(22) Unavailable: with difficulty+DP
    \([TP \text{ John flew off} [PP \text{ to } [DP [PP \text{ with difficulty}]] [DP \text{ Paris}]]]\) \(\ast\) for (19-c)

The ill-formedness of (22) suggests that with difficulty cannot integrate into the DP as an adjunct, and thus implicates the distributional profile in (23).

(23) Key distributional profile
    The PP with difficulty can adjoin on the clausal spine, but not to DPs.

\(^2\)Note that I did not include this PP in the survey of sentential operators in Chapter 2, since it is a manner adverb, and thus likely not analyzed as a sentential operator. For our purposes here, its semantic status is irrelevant. All that matters is that it cannot adjoin to DPs, as I demonstrate below, using the first environment from the survey (John flew off to — Paris).
This being so, the distribution of \textit{with difficulty} offers a tool to probe for clausal structure: if \textit{with difficulty} can occur at a given site, there must be clausal structure at that site to host it.

2.1.2 Diagnosing a post-copular TP

Turning now to pseudo-clefts, the key observation is this: \textit{with difficulty} is able to occur within the post-copular constituent in a pseudo-cleft. The example in (24) illustrates by adding the PP \textit{with difficulty} to the test example in (2):

(24) \textbf{Adding \textit{with difficulty} to pseudo-cleft}

What Obama approved was this bill and, with difficulty, that bill.

As we saw at the outset, \textit{and} must scope within the post-copular constituent. The PP, in turn, occurs inside the second conjunct: the PP is flanked by \textit{and} and \textit{that bill}, and it is clear from the meaning of (24) that the PP is interpreted within the second conjunct: (24) conveys (25-b), but not (25-a).

(25) \textbf{Evaluating entailments of (24)}

a. $\Rightarrow$ Obama approved this bill with difficulty.

b. $\Rightarrow$ Obama approved that bill with difficulty.

If the post-copular constituent were just a conjunction of DPs, the only option would be for \textit{with difficulty} to adjoin to the DP in the second conjunct, creating an illicit DP+PP constituent.

(26) \textbf{Structure for (24): DPs post-copular}

$\left[ TP \left[ DP \, \text{what Obama approved} \right] \left[ VP \, \text{was} \left[ \&P \left[ DP \, \text{this bill} \right] \right] \left[ \&P \left[ DP \, \text{that bill} \right] \right] \right] \right]$

Since \textit{with difficulty} can only adjoin on the clausal spine, the post-copular constituent must contain hidden clausal structure capable of hosting \textit{with difficulty}. This is the case if a full TP is post-copular and, within that TP, the conjunction scopes at (or above) the vP. With this CR structure, the PP can adjoin to the vP in the right conjunct — a licit attachment site for the PP.

(27) \textbf{Structure for (24): TPs post-copular}

$\left[ TP \left[ DP \, \text{what Obama approved} \right] \left[ VP \, \text{was} \left[ TP \, \text{Obama}_1 \left[ \&P \left[ VP \, \text{approved this bill} \right] \right] \left[ \&P \left[ VP \, \text{approved that bill} \right] \right] \right] \right] \right]$

In sum, adverbs which cannot adjoin to DPs can attach within a second apparent DP conjunct in post-copular position. This indicates that clausal structure must be present, supporting the idea that the post-copular constituent is a TP, which can have a CR syntax like that proposed in Chapter 3.

2.2 Test 2: VP-ellipsis

Now, I deploy a second test for an elided post-copular TP with CR structure: the test is based on the licensing of VP-ellipsis, and the crucial test example will contain a complex adverbial clause. To sketch the argument, consider again the two candidate structures for our test pseudo-cleft:
(28) **Structure 1: DP analysis**

\[
[TP \{DP \text{what Obama approved} \}[VP \text{was} \{D\{P \text{this bill} \}[D\{P \text{that bill} \}] \}] \]

(29) **Structure 2: CR analysis**

\[
[TP \{DP \text{what Obama approved} \}[VP \text{was} \{TP \text{Obama} \{VP \text{t, approved this bill} \}[\text{and} \{VP \text{t, approved that bill} \}] \}] \]

Focusing on the highlighted portions, a basic difference between the two analyses lies in whether or not the VP *approved that bill* is present in the second conjunct: that VP is present only in (29). I will show that the VP *approved that bill* is available to serve as antecedent to license ellipsis of another VP. From this, I conclude that (29) must be an available parse for the pseudo-cleft.

### 2.2.1 Diagnosing a post-copular TP

Like in Chapter 3, I introduce into the second conjunct a complex adverbial clause with the VP in that clause elided. The adverbial clause is *though he would rather not have* and the elided VP (A) is interpreted as *approved that bill*.

(30) **Ellipsis licensing**

What Obama approved was this bill and, though he would rather not have Δ, that bill.

(31) **Structure for adverbial clause**

\[
[CP \text{though} \{TP \text{he would rather not have} \}[VP \text{approved that bill} \Delta] \]

Where is the antecedent for Δ? Because this example is felicitous out of the blue, the antecedent must be present internal to the sentence itself. As such, the antecedent must be found in one of two places: the pre-copular constituent or the post-copular constituent.

### 2.2.2 No antecedent pre-copular

The pre-copular constituent does contain a VP. The VP begins the derivation as *approved what*, and *what* then moves out of it, leaving *approved t*. This is true regardless of whether *what* is a relative pronoun in a free relative DP, or a question word in an interrogative CP. For illustration, I show the structure with *what Obama approved* a relative clause inside a DP, glossing over structure above the relative clause.

(32) **Full structure for pre-copular free relative**

\[
[DP ... \{CP \{\lambda \{L \{TP \{\text{what Obama} \[VP \text{approved t,}\] \}] \} \}] \]

The question: is *approved t* an appropriate antecedent for *approved that bill*? To assess this, recall the Parallelism Condition on ellipsis, as adopted in Chapter 3:

(33) **Parallelism condition**

a. VP_e can elide if VP_e is reflexively dominated by a constituent PD (= parallelism domain), and the linguistic context provides an antecedent AC (= antecedent constituent) for PD which is semantically identical to PD, modulo focused marked constituents.
b. PD is semantically identical to AC modulo focus if there is a focus alternative to PD, $PD_{Alt}$, such that for every world $w$ and assignment function $g$, $[PD_{Alt}]^{w,g} = [AC]^{w,g}$.

Let us take approved that bill (i.e. $\Delta$) as the PD, and approved $t$ (i.e. the VP in the pre-copular constituent) as the AC. Because approved that bill is elided in its entirety, it must not contain any focused element. The sole focus alternative to approved that bill is therefore itself and, in turn, the Parallelism Condition requires that $[PD]^{w,g}$ be equivalent to $[AC]^{w,g}$ under any variable assignment $g$. The denotations for the PD and AC are:

(34) **Denotations for putative PD and AC**

a. $[PD]^{w,g} = \lambda x \cdot x \text{ approved that bill in } w$

b. $[AC]^{w,g} = \lambda x \cdot x \text{ approved } g(1) \text{ in } w$

Being approved $t$, the AC contains a free variable and has the assignment-dependent meaning. If the variable assignment $g$ is such that $g(1) =$ that bill, $[AC]^{w,g}$ is equivalent to $[PD]^{w,g}$. However, this equivalence does not hold for every choice of $g$: under any variable assignment that maps 1 to an entity other than that bill, $[AC]^{w,g} \neq [PD]^{w,g}$. This means that the Parallelism Condition is not met. The search for an antecedent must continue.

### 2.2.3 A post-copular TP is needed

Since no antecedent for $\Delta$ is to be found in the pre-copular constituent, the antecedent must be post-copular. Without hidden projections, the syntax of the post-copular &P would be (35), which simply adds the adverbial clause to (28).

(35) **DPs post-copular**

a. $[\&P [DP \text{ this bill} \ [and \ [DP \ CP [DP \ that \ bill]]]]]$

b. $[CP \text{ though } [TP \text{ he would rather not have } [\&P \text{ approved that bill} \Delta]]]$

The situation is bleak: if the post-copular &P is just made up of DPs, there is no hope of finding an antecedent for $\Delta$ there. What we need is an antecedent VP, but no VP is present. Positing a post-copular TP with CR syntax resolves the problem:

(36) **TP with CR syntax post-copular**

a. $[TP \text{ Obama} [\&P [CP [VP t1 \text{ approved this bill} \ [and \ [VP \ CP [VP t1 \text{ approved that bill}]]]]]]]$

b. $[CP \text{ though } [TP \text{ he would rather not have } [\&P \text{ approved that bill} \Delta]]]$

The presence of a TP entails the presence of a VP, and a CR syntax means there are two VPs, one in each conjunct. The VP in the right conjunct is approved that bill, which has an assignment-independent meaning clearly equivalent to $[\Delta]$: both denote the predicate $\lambda x \cdot x \text{ approved that bill in } w$. If we take the VP in the right conjunct (red in (36-a)) as the AC, and the VP in the adverbial clause (red in (36-b)) as the PD, the Parallelism Condition is thus satisfied, licensing ellipsis. Hence, we reach our main result, which corroborates the conclusion we reached in the previous section: a parse is available where the post-copular XP is a TP with CR syntax.
2.3 Local summary

I provided two novel arguments that pseudo-clefts may be parsed with a hidden TP as the post-copular XP, and that this hidden TP may have CR syntax. Accordingly, the possibility of apparent DP conjunction in pseudo-clefts is not a counter-argument to the SIH. Rather, the SIH led us to expect hidden syntax in these data — and there is evidence for exactly that hidden syntax.

While positing a full clausal syntax for the post-copular XP has reconciled pseudo-clefts with the SIH and explained a range of data involving adverbs and ellipsis, more work is needed to integrate a clausal analysis into a full theory of pseudo-clefts. The question for us now: how does the post-copular TP integrate with the pre-copular constituent and the copula to compositionally derive the intuitive meaning of a pseudo-cleft? The remainder of the chapter develops a full analysis.

3 Previous compositional analysis

To start, let us set aside adverbs and even conjunction, and look just at the basic pseudo-cleft in (1), which was introduced at the very beginning of the chapter.

(1) Recall: basic pseudo-cleft

What Obama approved was this bill.

Our conclusion in the previous section was that pseudo-clefts can be parsed, at least optionally, with a full TP as the post-copular constituent. Parsed this way, the structure for (1) would be:

(37) Structure for (1) so far

```
TP
   /
  / \
XP   VP
   what Obama approved was TP

Obama approved this bill
```

At this point, the structure is somewhat skeletal. The pre-copular constituent is a nameless XP, and more work is needed to address how that XP, the copula, and post-copular TP compose. In this section, I present a previous idea about how to fill in the skeleton, due to Ross (1972) and Schlenker (2003).

3.1 The semantic desideratum

Before we turn to analytical matters, it is important to re-iterate the semantic desideratum: what does (1) actually mean? Intuitively, there are two inferences: (a) that Obama approved this bill, and (b) that he approved no bill other than this one. I flag these in (38) for easy reference, and refer to them as the basic inference and the exhaustivity inference, respectively.
Some authors have observed that the exhaustivity inference is defeasible (in particular, Collins 1988). Though we might be inclined to take exhaustivity to be an implicature on this basis, most analyses of pseudo-clefts treat it as an entailment, and I will do the same. I take it that the desideratum is, then, to predict (38-a) and (38-b), and to derive both as entailments.\footnote{In addition to the inferences in (38), pseudo-clefts are generally taken to carry an existence presupposition. In (1), this would be a presupposition that Obama approved \textit{something}. This presupposition will not be addressed here.}

### 3.2 Ross's intuition

With the empirical desideratum in place, we can now return to analytical matters. Since Ross (1972), analyses which posit a full post-copular TP have been guided by the intuition that pseudo-clefts are an intra-sentential pairing of a question and answer. The pseudo-cleft in (1) is analogized to the question-answer pair in (39), with the post-copular TP analogous to the TP answer in (39-b).

(39) **Baseline question-answer pair**

a. What did Obama approve?

b. He approved (only) this bill.

Really, (1) is even more closely analogous to the pair of a question and \textit{fragment} answer in (40). Just as post-copular \textit{this bill} is the only pronounced remnant of a full TP, Merchant (2004) analyzes fragment answers as underlyingly full TPs: (40-b) has the same structure as (39-b).

(40) **Closer baseline with fragment answer**

a. What did Obama approve?

b. (Only) this bill.

Recall that pseudo-clefts comprise three syntactic ingredients: (a) a pre-copular XP, (b) a copula, and (c) a post-copular TP. Ross's intuition provides a guide for how each piece should be interpreted. If the post-copular TP is analogous to a fragment answer, the pre-copular XP must be interpreted as question. In turn, the copula must somehow compose with the pre-copular question and the post-copular answer to yield a proposition capturing the basic and exhaustivity entailments.

Note that, in this latter respect, pseudo-clefts distinguish themselves from true question-answer pairs. In (39) and (40), the question and answer constitute separate speech acts and, accordingly, do not need to compose with one another. The situation is different with pseudo-clefts: pseudo-clefts express a single proposition, so the intra-sentential question and answer must compose.

So then, accepting Ross's intuition, there are two questions to answer in order to achieve a full compositional analysis of pseudo-clefts. First, what is the syntax and semantics of the pre-copular XP such that it comes to be interpreted as a question? And, second: how do the pre-copular question and post-copular answer compose? I present previous solutions to each in turn.
3.3 Step 1: the pre-copular XP

The syntactic status of the pre-copular XP has been a point of significant debate. Certain authors take *what Obama approved* to have the internal syntax of an embedded interrogative CP, analogous to (41-a). Others take it to have the syntax of a free relative DP, analogous to (41-b).

(41) **Potential baselines**

a. I wonder \([\text{CP} \ \text{what Obama approved}]\). (*embedded question*)

b. I read \([\text{DP} \ \text{what Obama approved}]\). (*free relative*)

To cache out Ross’s intuition, the pre-copular XP must have the **semantics** of a question. This may make embedded question syntax seem like the more natural idea. Yet, there are strong arguments for a free relative syntax — and this crucially does not undermine the semantic claim, since definite DPs can in general be interpreted as **concealed** questions. The internal syntax of the pre-copular constituent is not a special point of concern here, and I will remain neutral on the debate. I will state my assumptions about how questions are interpreted and briefly sketch how the right meaning could come about under each syntactic option. Then, in the remainder of the chapter, I will adopt CP labeling for the pre-copular constituent, purely as an expositional convenience.

3.3.1 Semantics of questions

Different proposals have been advanced with respect to the semantics of questions. Throughout this thesis, I will adopt the approach of Hamblin (1973) and Karttunen (1977) whereby questions denote sets of propositions. More particularly, I adopt Hamblin’s idea that a question denotes the set of its possible answers. This assumption about question semantics will have an important role to play both in this chapter and the next. Concretely, the question in (42-a) denotes the (characteristic function for) a set propositions *that Obama approved X*, for different contextually salient entities X, as in (42-b). I indicate the set of contextually salient entities here as \(C_e\), though I will omit the restrictor subsequently.

(42) **Hamblin denotation**

a. What did Obama approve?

b. \([\text{(42-a)}]\)’\(w\) = \(\lambda p_{ar}. \exists X \in C_e \ [p = \lambda w'. \forall x <_{AT} X [\text{Obama approved x in w'}]]\)

Note that \(X\) could be an atomic entity, such as this bill, or a plural entity, such as this bill@that bill. In the case of a plural entity, the statement of the Hamblin set in (42-b) yields a distributive reading. That is, the proposition corresponding to this bill@that bill as \(X\) is the proposition that *Obama approved every atom of this bill@that bill* or, equivalently, that *Obama approved this bill and Obama approved that bill*. If there are just two salient entities, this bill and that bill, the Hamblin set contains these three elements:

(43) **Illustrative Hamblin set**

a. \(\{\lambda w. \ \text{Obama approved this bill in w},\) 

b. \(\lambda w. \ \text{Obama approved that bill in w},\) 

c. \(\lambda w. \ \text{Obama approved this bill in w} \land \text{Obama approved that bill in w}\)
In exposition throughout the chapter, I will refer to the entity given by ‘this bill’ as b1, and the entity
given by ‘that bill’ as b2. In kind, I will abbreviate the proposition in (43-a) as ‘φ1’, the proposition in
(43-b) as ‘φ2’, and the proposition in (43-c) as ‘φ1 ∧ φ2’.

3.3.2 Deriving from interrogative CP

Although I have adopted Hamblin’s final semantics, I will put forward a compositional derivation for an
interrogative CP more like Karttunen’s. In this chapter, this is purely an expositional choice, though it
will have more of a role to play in the next chapter. The specific adaptation of Karttunen’s ideas I provide
here is based on von Stechow (1996), Heim (2012), and Fox (2013). The full syntax for the interrogative
CP what Obama approved is:

(44) Pre-copular XP as interrogative CP

\[
\begin{array}{c}
\text{CP}_3 \\
\text{PRO} \\
\lambda 2 \\
\text{CP}_2 \\
\text{what} \\
\lambda 1 \\
\text{CP}_1 \\
\text{C} \\
\text{t}_2 \\
\text{TP} \\
\text{Obama <each> approved t}_1
\end{array}
\]

The interrogative complementizer, defined as (46), asserts that two propositions are identical. Its sister
is a covert pronoun PRO, which moves to adjoin on the clausal spine (modifying Karttunen, following
Heim, Fox). PRO leaves a trace, which is interpreted as a variable of type <s,t>. This variable saturates
the first argument of C. The second argument is saturated by the intension of the TP out of which what
has moved, through Intensional Functional Application. Throughout the chapter, I indicate the intension
of α as [α]_C, after Dowty et al. (1981).

(45) Defining the interrogative complementizer

\[ [C]^w = \lambda p_{st} \lambda q_{st} . p = q \]

(46) Composing [CP1]

\[ a. \quad [\text{CP}_1]^w = [\text{t}_2]^w ([\text{TP}]_C) \] (IFA)
\[ b. \quad = [C]^w ([\text{t}_2]^w g ([\text{TP}]_C) \] (FA)

\[ c. \quad = [C]^w ([\text{t}_2]^w g ([\text{TP}]_C) \]

In stating denotations, I will be fairly scrupulous in including the world parameter to differentiate extension and intension.

The variable assignment I will now show unless necessary.
What is interpreted ex situ as an existential quantifier (over atomic and plural entities) binding its trace, (47). Covert each, (48), attaches at the VP and yields a distributive interpretation. Hence, CP₂ provides the equation \( p = \text{that every atom of } X \text{ got hired} \), with \( X \) existentially bound.

(47) **Defining what**

\[ [\text{what}] = \lambda f \alpha . \exists X \ [f(X)] \]

(48) **Defining covert each (Link 1983)**

\[ [\text{each}] = \lambda f \alpha . \forall x <_{AT} X \ [f(x)] \]

Movement of PRO targets a position just above CP₂. PRO is not itself interpreted, but triggers Predicate Abstraction, binding the propositional variable \( p \). The meaning in (49) for CP₃ is the final meaning for the question, and it is the Hamblin denotation from the previous subsection. The semantic desideratum is achieved compositionally.

(49) **Denotation at CP₃**

\[ [\text{CP₃}] = \lambda p \alpha . \exists X \ [p = \lambda w \ . \forall x <_{AT} X \ [\text{Obama approved } x \text{ in } w']] \]

Note that there are a number of ideas about how a Hamblin denotation comes about. Hamblin himself proposed that what introduces a set of entities, which compose pointwise with higher nodes, propagating up the structure until a set of propositions is derived. A version of this idea was adopted in Kotek (2014). For a number of authors, what Obama approved itself would have a property meaning, and a higher covert morpheme would convert that property to a question meaning (George 2011, cf. Caponigro 2003, Jacobson 1995). Cable (2007, 2010) proposes a hybrid theory mixing these two approaches. I believe the core insights in this chapter could be replicated with any of these compositional tacks. I presented the Karttunen-inspired composition for concreteness, and to set up discussion in the next chapter.

### 3.3.3 Deriving from free relative

Now, suppose the pre-copular free relative is not really an interrogative CP after all, but rather has the syntax of a free relative. There is, in fact, a reasonable body of evidence for a free relative treatment. Sharvit (1999) provides a cross-linguistic argument from Hebrew, where questions and free relatives are morphologically distinct. In embedded questions, there is no overt complementizer, whereas an overt C (Se) does occur in free relatives. Sharvit illustrates with this contrast:

(50) **Baseline embedded question (a) vs. free relative (b)**

- a. Dan berer ma karati
  
  Dan found-out what I-read

  ‘Dan found out the answer to the question: “What did I read?”’

- b. *Dan berer (et) ma Se karati
  
  Dan found-out Acc C that I-read
Berer (‘find out’) requires an embedded question, which it receives in (51-a), where Se does not appear. In (50-b), adding Se yields ungrammaticality, as expected if Se signals a free relative. Sharvit observes that the XP in a pseudo-cleft patterns with free relatives. Se is obligatory in (51).

### Pseudo-cleft patterns with free relative

ma Se hu berer ze im Dan ohev gvina
what that he find-out was whether Dan likes cheese
‘What he found out was whether Dan likes cheese.’

Support for a free relative analysis in English comes from the paradigm in (52), which illustrates an English-internal difference between embedded questions and free relatives, and shows that the pre-copular XP patterns with free relatives. Embedded questions permit which NP in place of what, while free relatives and the XP do not (Sharvit 1999, Caponigro & Heller 2007).

### Distribution of which NP

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>I wonder which bill Obama approved.</td>
<td>(embedded question)</td>
</tr>
<tr>
<td>b.</td>
<td>*I read which bill Obama approved.</td>
<td>(free relative)</td>
</tr>
<tr>
<td>c.</td>
<td>*Which bill Obama approved was this one.</td>
<td>(pre-copular XP)</td>
</tr>
</tbody>
</table>

If the pre-copular XP is syntactically a free relative, how can it be interpreted as a question? In general, free relatives refer to entities (following Jacobson 1995). The baseline example in (41-b) is repeated as (53-a). Read requires an entity as argument, and what Obama approved is interpreted like the referential definite description the thing that Obama approved, per the paraphrase in (52-b).

### Free relatives are usually referential

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>I read what Obama approved.</td>
<td>(free relative)</td>
</tr>
<tr>
<td>b.</td>
<td>I read the thing Obama approved.</td>
<td>(referential definite description)</td>
</tr>
</tbody>
</table>

Yet, as Schlenker (2003) discusses in the context of pseudo-clefts, definite descriptions can be interpreted as questions in certain environments. DPs used this way are known as concealed questions (‘CQ’). This usage was originally observed by Baker (1968) and has been discussed extensively in later work (see Nathan 2006 for an overview). Schlenker illustrates with (54), from Heim (1979):

### Concealed questions

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>John knows the capital of Italy.</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>They revealed the winner of the contest.</td>
<td></td>
</tr>
</tbody>
</table>

On its most prominent reading, (54-a) does not assert that John knows the entity Rome, but rather that he knows that Rome is the capital of Italy. In other words, he knows the answer to the question ‘What is the capital of Italy?’ The intuition is similar in (54-b), which says that they revealed the identity of the winner, i.e. the answer to the question ‘Who won the contest?’ The sentences in (54) thus receive a similar interpretation to those in (55) with true embedded questions:
(55) **Parallel to embedded questions**

a. John knows which city is the capital of Italy.  \(\approx (54\text{-a})\)
b. They revealed which contestant won the contest. \(\approx (54\text{-b})\)

In light of these data, the pre-copular XP in a pseudo-cleft, if a free relative DP, could function as a CQ. This becomes increasingly plausible when we consider independent restrictions on what kinds of nominals allow CQ readings. In general, what Nathan (2006) terms *relational nouns* can form the core of a CQ, while other nouns cannot. *Capital* in (54-a) is a relation between a country and its capital city, and *winner* in (54-b) is a relation between the contest and its winner. A non-relational noun such as *semanticist* does not allow a CQ reading, as the contrast in (56) shows.

(56) **No CQ reading available**

a. *Tell me the semanticist.*  \((*\text{CQ})\)
b. Tell me who the semanticist is.  \((\text{embedded question})\)

Nathan observes, however, that a CQ reading is productively available with any DP that contains a modification structure, regardless of the noun. Unacceptable (56-a) sharply contrasts with (57). Since free relatives are by their very nature modification structures, free relatives are expected to allow CQ readings in kind with (57).

(57) **Modification enables CQ**

Tell me the semanticist who works at MIT. \((\checkmark \text{CQ})\)

One further issue has to do with the environments in which CQ readings can occur for DPs that allow them. It has been known at least since Grimshaw (1979) that CQs are restricted to specific environments. Whereas *know, reveal, and tell* all permit CQs, another interrogative embedding verb, *wonder*, does not. *Wonder* can embed a true interrogative CP, but many speakers reject a CQ DP.

(58) **Dependence on environment**

a. John wonders which city is the capital of Italy.
b. ??John wonders the capital of Italy.

Due to this observation, Grimshaw proposed to separate semantic and syntactic selection. Semantically, *know, reveal, tell*, and *wonder* all require question meanings. Syntactically, the first three select CPs and DPs, while *wonder* just selects DPs. Schlenker proposed that the pre-copular site in a pseudo-cleft is hospitable to DPs. In fact, if the paradigm in (52) (repeated), is conclusive, it may syntactically allow DPs, but not CPs, conversely to the complement of *wonder*.

(52) **Distribution of which NP**

a. I wonder which bill Obama approved.  \((\text{embedded question})\)
b. *I read which bill Obama approved.  \((\text{free relative})\)
c. *Which bill Obama approved was this one.  \((\text{pre-copular XP})\)

\(^5\) See Caponigro & Heller (2007), however, for counter-evidence to the idea that CQ readings are generally available for free relatives. I must leave their data an open puzzle.
The precise composition of CQs is not especially germane, so long as what Obama approved on its CQ reading has the Hamblin denotation seen in the preceding subsections. For convenience, I will present Nathan’s approach, which is very brute force in nature. Nathan proposes separate covert morphemes to derive CQ readings with relational nouns and modified nouns. The covert morpheme tailored to modified nouns is (59). This operator takes two properties (type <e,st>) as arguments and, accordingly, can only compose in a modification structure.

(59) The CQ operator (for modification)
\[
[[CQ]]^w = \lambda Q_{<s,et>} . \lambda P_{<s,et>} . \lambda p_{st} . \exists X [p = \lambda w’ . [P(w’)(X) \land Q(w’)(X)]]
\]

Let us assume the full syntax for the pre-copular XP as a free relative CQ would be (60). The first argument of CQ is saturated, through IFA, by a null nominal denoting the trivial predicate of being an entity. The second argument is saturated, likewise through IFA, by the relative clause modifier. As shown below the tree, the meaning predicted for the NP is the same Hamblin set as we computed from an interrogative CP before — achieving our semantic desideratum. I will leave the D head empty.

(60) Pre-copular XP as free relative CQ

(61) Hamblin denotation predicted at NP
a. \[ [NP]^w = [[CQ]^w([THING]_C([CP]_C))] \]
b. \[ = [[CQ]^w(\lambda w’ . \lambda X . X \in D_e)(\lambda w’ . \lambda X . \forall x <_{AT} X [Obama approved x in w’])] \]
c. \[ = \lambda p_{st} . \exists X [p = \lambda w’ . [X \in D_e] \land \forall x <_{AT} X [Obama approved x in w’]] \]
d. \[ = \lambda p_{st} . \exists X [p = \lambda w’ . \forall x <_{AT} X [Obama approved x in w’]] \]

3.3.4 Concluding Step 1

In sum, there are two syntactic possibilities for the pre-copular XP — it could be an interrogative CP or a free relative DP — and both offer a route to the same meaning: the Hamblin set for the question 'What did Obama approved?'. Purely for expositional convenience, I will adopt the CP label in the remainder of the text, but this choice is not crucial. To reflect this choice, we can update the initial skeletal structure for (1) provided at the beginning of this section as in (62):
3.4 Step 2: composing the question and answer

I have discussed the first step in caching out Ross’s intuition: the pre-copular XP is a CP (by assumption), semantically a question, while the post-copular YP is a full TP expressing a proposition. The second step is to consider exactly how the question and proposition compose. Schlenker (2003) pursues the intuition that the copula asserts identity between the strongest true answer to the pre-copular question and the proposition given by the post-copular TP. That is, the full truth-conditions for (1) paraphrase as in (63). I refer to this idea as the Identity Hypothesis. As we will see, these truth-conditions capture the basic and exhaustivity entailments in basic data such as (1).

(63) **Truth-conditions for (1) (Identity Hypothesis)**

“'The strongest true answer to the question 'What did Obama approve?' is the proposition that Obama approved this bill.'”

Schlenker suggests several different ways to cache out the Identity Hypothesis compositionally. I present one here, which integrates well with the assumptions I have made about how questions are interpreted. The complete LF for (1) is (64), annotated for semantic types.

(64) **Complete LF for (1) (Identity Hypothesis)**

---

 Rather than taking questions to be sets of propositions, Schlenker favors an analysis of the question as denoting an equivalence relation between worlds (as in Groenendijk & Stokhof 1989). I adopt the set of propositions analysis to set up my own proposal, and discussion in Chapter 5. Schlenker does entertain this option, as well.

128
Just above the interrogative CP, an operator, ANS, integrates. Adapted from Dayal (1996), ANS takes as its argument a set of propositions (P). ANS presupposes that P contains a proposition that is both true at w and entails all other propositions true at w. If so, ANS returns that unique strongest true proposition.

(65) **Defining ANS**

\[
[\text{ANS}]^w = \lambda p_{st} : \exists p \ [w \in p \land \forall p' \in P \ [w \in p' \rightarrow p \subseteq p']]
\]

Since ANS yields a proposition, the full pre-copular XP ends up denoting a proposition: the strongest true proposition in the Hamblin set for the question. Now, the pre-copular XP has the same type as the intension of the post-copular TP: both denote propositions. The role of the copula, (66), is to operate on two propositions and assert that they are identical. It composes with the intension of the post-copular TP through IFA, and with the pre-copular XP through FA.

(66) **Defining the copula**

\[
[\text{be}]^w = \lambda p_{st} \cdot \lambda q_{st} \cdot p = q
\]

Putting all the pieces together, the LF in (64) corresponds to the following meaning, which captures the truth-conditions given idiomatically above.

(67) **Predicted truth-conditions**

\[
[[\text{(64)}]]^w = 1 \text{ iff } [\text{ANS}]^w(\lambda p \cdot \exists x [p = \lambda w'. \forall x <_{AT} X \ [\text{Obama approved x in w'}])] = \lambda w' \cdot \text{Obama approved this bill in } w'
\]

Because the proposition that Obama approved this bill is asserted to be a true answer to the question, the basic inference — that this proposition is true — is immediately predicted. The exhaustivity inference comes from the strongest component of the truth-conditions. To illustrate concretely, suppose that the Hamblin set for the pre-copular question contained just three elements, as illustrated earlier in (43): \{\phi_1, \phi_2, \phi_1 \land \phi_2\}. The exhaustivity inference is intuitively consistent with a world, w_1, at which Obama approved only b1. At w_1, the only true proposition in the Hamblin set in (67) is \phi_1, so \phi_1 is clearly the strongest true proposition, as well. Since \phi_1 is identical to the post-copular proposition, the truth-conditions for the pseudo-cleft are satisfied, as desired.

(68) **Interpretation of pre-copular XP at w_1**

a. \[\text{[XP]}^w_1 = [\text{ANS}]^w_1(\lambda p \cdot \exists x [p = \lambda w'. \forall x <_{AT} X \ [\text{Obama approved x in w'}])]\]

b. \[\phi_1\]

c. \,[],c

Now, consider a world inconsistent with the exhaustivity inference: w_2, at which Obama approved b1 and b2. At w_2, all three propositions in the Hamblin set are true, and \phi_1 \land \phi_2 — not \phi_1 — is the strongest, since \phi_1 \land \phi_2 entails both \phi_1 and \phi_2. \phi_1 \land \phi_2 is not identical to the post-copular proposition, so the truth-conditions fail.
(69) **Interpretation of pre-copular XP at w\(_2\)**

a. \([\text{XP}]^{w_2} = [\text{ANS}]^{w_2}(\lambda p . \exists x [p = \lambda w'. \forall x <_{AT} X [\text{Obama approved } x \text{ in } w'])\]

b. \(= \phi_1 \land \phi_2\)

c. \(\not= [\text{TP}_1]_C\)

Extrapolating from \(w_1\) and \(w_2\), the only way for the post-copular proposition (i.e. \(\phi_1\)) to be the strongest true answer to the pre-copular question is for Obama to have approved \(b_1\) and no other — capturing the exhaustivity inference.

### 3.4.1 Concluding Step 2

In sum, the composition goes like this. The pre-copular XP has the semantics of a question and, due to the presence of ANS, denotes the strongest true answer to that question. The post-copular TP expresses a proposition. The copula is defined as an identity relation, and yields these overall truth-conditions: the pseudo-cleft is true just in case the strongest true answer to the pre-copular question is the post-copular proposition. In examples like (1), this captures the basic and exhaustivity inferences. Overall, then, the approach rests on two claims: (a) that the pre-copular XP is semantically a question, and (b) that the pre-copular question and post-copular proposition compose in accordance with the Identity Hypothesis. In the next section, I will show that this package of claims cannot be correct and challenge, in particular, the Identity Hypothesis. A new compositional analysis will be needed.

### 4 New composition puzzle

Despite capturing (1), an identity semantics faces a challenge from data along the lines of those in Section 2, where the post-copular constituent contains an adverb. In at least certain adverb data, the proposition expressed by the post-copular clause can never be equivalent to the strongest true answer to the pre-copular question. An identity semantics thus predicts contradictory truth conditions when, in fact, the data are still intuitively contingent.

#### 4.1 Illustrating with possibly

The clearest illustration of the problem comes from example (70), where the adverb *possibly* is inserted into the second conjunct in the post-copular TP. This example is perfectly felicitous, and is most naturally interpreted with the exhaustive meaning in (71). Assuming the post-copular TP has a CR syntax, the full LF for (70) under the Identity Hypothesis would be (72).

(70) **Pseudo-cleft with possibly**

What Obama approved was this bill and possibly that bill.

(71) **Paraphrase of (70)**

"Obama approved \(b_1\) and possibly \(b_2\), but no bill other than \(b_1\) or \(b_2\)."
To interpret the LF, we require a meaning for possibly. As discussed in Chapter 2, I take possibly to be a modal, analyzed like any modal in Kratzer's (1977) framework as a quantifier over possible worlds. Possibly denotes an existential quantifier and, in all cases we will discuss, quantifies over worlds compatible with what the speaker believes to be true. In other words, possibly is a epistemic modal with existential force. In Kratzer's framework, the restriction for the quantifier is contextually determined. For exposition, however, I write the epistemic restriction right into the lexical entry: \( \text{possibly} \) picks out the set of worlds epistemically accessible to the speaker in context C from world w.

\[
[\text{possibly}]^w_C = \lambda p. \exists w' \in \text{\( \mathcal{S}(\text{SP}(C))(w) \) \[p(w')]}\]

Having established a meaning for possibly, the full post-copular TP in (70) has the intension in (74-a), which I abbreviate as (74-b). In turn, the LF as a whole yields the identity semantics given both formally and idiomatically in (75).

\[
\text{Intension of post-copular TP}
\]

a. \( \llbracket \text{TP}_1 \rrbracket^w_C = \lambda w. \phi_1(w) \land \exists w' \in \text{\( \mathcal{S}(\text{SP}(C))(w) \) \[\phi_2(w')]}\]

b. \(~\sim (abbreviation) \phi_1 \land \diamond \phi_2\)

\[
\text{Overall predicted meaning}
\]

\( \llbracket (70) \rrbracket^w = 1 \text{ iff } \llbracket \text{ANS} \rrbracket^w(\llbracket \text{CP} \rrbracket^w) = \phi_1 \land \diamond \phi_2 \)

"The strongest true answer to the question 'What did Obama approve?' is the proposition that Obama approved b1 and it is epistemically possible that he approved b2".

For the pseudo-cleft to come out contingent, there must be a world at which \( \llbracket \text{TP}_1 \rrbracket^w \) qualifies as the strongest true answer to the pre-copular question. In fact, there is not. Recall that the Hamblin set contains propositions that Obama approved X, where X is some entity:
Recall: Hamblin denotation for question

a. \([\text{CP}][^w] = \lambda_{\text{CP}} \cdot \exists X [p = \lambda x_w' . \forall x <_{AT} X \text{ [Obama approved } x \text{ in } w']]\)

b. \(\approx \{\phi_1, \phi_2, \phi_1 \land \phi_2\}\)

Since none of the propositions in the Hamblin set is epistemically modalized, \([\text{TP}_1][^C] \) is not an element of the Hamblin set. \([\text{TP}_1][^C] \) is not formally an answer to the question, let alone the strongest true answer. For this reason, \([\text{ANS}][[\text{CP}]]\) cannot pick out \([\text{TP}_1][^C] \) at any world. At a given world, \([\text{ANS}][[\text{CP}]]\) returns the proposition that Obama approved \(X\) with \(X\) the maximal plurality of bills that Obama approved at that world. This illustrated in (77) and (78) for two worlds: \(w_1\) (at which Obama approved just \(b_1\)), and \(w_2\) (at which Obama approved \(b_1\) and \(b_2\)).

Strongest true answer at \(w_1\)
\([\text{ANS}][w_1]^w (\text{[CP]}[w_1]^w) = \phi_1 \quad (\neq \phi_1 \land \phi_2)\)

Strongest true answer at \(w_2\)
\([\text{ANS}][w_2]^w (\text{[CP]}[w_2]^w) = \phi_1 \land \phi_2 \quad (\neq \phi_1 \land \phi_2)\)

The strongest true answer to the question is independent of the speaker's epistemic state. Intuitively, \([\text{TP}_1][^C] \) is under-informative compared to the strongest true answer to the question: the answer specifies exactly the bills that Obama in fact approved — not just bills that he possibly approved. Because \([\text{TP}_1][^C] \) can never be the strongest true answer to the pre-copular question, the predicted meaning for the pseudo-cleft is a contradiction.

4.2 Illustrating with a conditional

The problem is not restricted to possibly. Another illustration comes from (79), where the post-copular constituent contains a conditional. This example is again felicitous and contingent, with its most natural meaning paraphrased in (80).

Pseudo-cleft with conditional
What Obama will approve is this bill and, if Biden lets him, that bill.

Paraphrase of (79)
"Obama will approve \(b_1\) and, if Biden lets him, \(b_2\), but not any bill other than \(b_1\) or \(b_2\)."

Following Lewis (1975), Kratzer (1977), and Heim (1982), I analyze the conditional as containing a covert universal modal, with the antecedent clause providing the restrictor for that modal. This being so, the LF just for the post-copular constituent would be (81). The covert modal occurs inside the second conjunct, where if Biden lets him is its restrictor, and the vP is its scope.
(81) **LF for post-copular constituent**

Assuming again that the modal is epistemic, this LF conveys (82), and the overall identity semantics for the pseudo-cleft would be (83).

(82) **Interpreting the LF**

a. $\langle TP \rangle^C = \lambda w. \phi_1(w) \land \forall w' \in E(SP(C))(w) [\text{Biden lets him approve } b_2 \text{ in } w' \rightarrow \phi_2(w')]$

b. $\sim^\omega (\text{abbreviation}) \phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2]$

(83) **Identity semantics for (79)**

$\langle (79) \rangle^w = 1 \text{ iff } \langle \text{ANS} \rangle^w([\text{what Obama approved}]) = \phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2]$

"The strongest true answer to the question 'What did Obama approve?' is the proposition that Obama will approve $b_1$ and, at every epistemically accessible world compatible with what Biden allows, he will approve $b_2.'"

Much as we observed with possibly, $\langle TP \rangle^C$ is not in the Hamblin set for the pre-copular question, due to the modalization associated with the conditional. $\langle TP \rangle^C$ is not equivalent to any proposition that Obama will approve $X$. As a result, $\langle \text{ANS} \rangle([\text{what Obama approved}])$ cannot pick out $\langle TP \rangle$ at any world — and the truth-conditions will never be satisfied.

### 4.3 Illustrating with with difficulty

In Section 2, we discussed the example in (24), with the adverbial PP with difficulty. The natural exhaustive reading of (24) is paraphrased in (84).

(24) **Recall: with difficulty post-copular**

What Obama approved was this bill and, with difficulty, that bill.

(84) **Paraphrase of (24)**

"Obama approved $b_1$ and with difficulty $b_2$, and no other bill."
As argued in Section 2, the post-copular TP has the structure in (85). On the face of it, this example evinces the same problem as the examples with possibly and the conditional. The proposition given by the post-copular clause is, quite informally, (86), which conjoins \( \phi_1 \) and \( \phi_2 \) with a third conjunct, due to the adverb (that Obama found it difficult to approve \( b_2 \)). The expected identity semantics for the pseudo-cleft is, then, (87).

\[
\text{(85) LF for post-copular constituent}
\]

\[
\begin{array}{c}
\text{TP} \\
\text{Obama} \\
\text{vP} \\
\text{t}_1 \text{ approved this bill and} \\
\text{vP} \\
\text{with difficulty} \\
\text{vP} \\
\text{t}_1 \text{ approved that bill}
\end{array}
\]

\[
\text{(86) Intension of post-copular TP}
\]

a. \( [\text{TP}]_C = \phi_1 \land \phi_2 \land \text{Obama found approve } b_2 \text{ difficult} \)

b. \( \sim (\text{abbreviation}) \phi_1 \land \phi_2 \land \text{difficult-} b_2 \)

\[
\text{(87) Identity semantics for (24)}
\]

\[
[\text{(24)}]^w = 1 \text{ iff } [\text{ANS}]^w([\text{what Obama approved}]^w) = \phi_1 \land \phi_2 \land \text{difficult-} b_2
\]

"The strongest true answer to the question ‘What did Obama approve?’ is the proposition that Obama approved \( b_1 \) and \( b_2 \), and found approving \( b_2 \) difficult”.

As before, \( [\text{TP}]_C \) is not an element of the Hamblin set for the pre-copular question, this time due to the final conjunct, contributed by with difficulty. To appreciate the precise nature of the problem, take a world, \( w_1 \), at which the sentence is intuitively true, i.e. a world at which (a) Obama approved \( b_1 \), (b) he approved \( b_2 \), and (c) he found approve \( b_2 \) difficult. At \( w_1 \), the strongest true proposition in the Hamblin set is \( \phi_1 \land \phi_2 \). Because \( \phi_1 \land \phi_2 \) is not identical to \( [\text{TP}]_C \), the pseudo-cleft is false at \( w_1 \).

\[
\text{(88) Denotation of pre-copular XP at } w_1
\]

a. \( [\text{ANS}]^{w_1}([\text{what Obama approved}]^{w_1}) = \phi_1 \land \phi_2 \)

b. \( \neq \phi_1 \land \phi_2 \land \text{difficult-} b_2 \)

Consider the relationship between \( \phi_1 \land \phi_2 \) and \( [\text{TP}]_C \). Because \( [\text{TP}]_C \) has the same two conjuncts as \( \phi_1 \land \phi_2 \), plus a third (difficult-\( \phi_2 \)), \( [\text{TP}]_C \) asymmetrically entails \( \phi_1 \land \phi_2 \). At a world at which the sentence is intuitively true, the post-copular proposition is even stronger than the strongest true answer to the question. Whereas the post-copular proposition was under-informative with possibly, it is over-informative with with difficulty.
4.4 Against a parenthetical solution

A natural impulse for at least certain of the above data might be to take the adverb not to have truth-conditional import, but rather to be a parenthetical, contributing on a non-at-issue dimension. If the adverb does not contribute to the truth-conditions of the post-copular TP, then the composition problem would resolve. A parenthetical treatment may be available for *with difficulty*, but does not seem viable for *possibly* or the conditional. I conclude, therefore, that appealing to parentheticals does not provide a general solution to the composition problem.

Potts (2003, 2005) proposed a multi-dimensional semantic framework, where parentheticals contribute on a non-at-issue dimension separate from other operators. The example in (89) illustrates with an appositive relative clause. The at-issue content is (89-a), while the appositive relative contributes the separate non-at-issue entailment in (89-b).

(89) **Multi-dimensional semantics**

Obama approved this bill, which McCain proposed.

a. Obama approved this bill.  \(\text{(at-issue)}\)

b. McCain proposed this bill.  \(\text{(non-at-issue)}\)

In diagnosing parenthetical status, it will be useful to note a hallmark property of parentheticals: because parentheticals contribute on a non-at-issue dimension, they do not interact with other scope-bearing operators whose contribution is at-issue. This is shown in (90) for an appositive relative, based on the “family of sentences” test:

(90) **Non-at-issue entailments “project”**

a. Did Obama approve this bill, which McCain proposed?

b. Obama didn’t approve this bill, which McCain proposed.

c. If Obama approved this bill, which McCain proposed, I will be happy.

These examples embed the sentence in (89) under other operators: a question-forming operator in (90-a), negation in (90-b), and a conditional in (90-c). The key observation is that embedding impacts the at-issue entailment, but not the non-at-issue entailment. All of (90-a)-(90-c) license the same inference that McCain proposed this bill, just as the original sentence did. The non-at-issue entailment “projects” to take matrix scope.

Having established background on parentheticals, let us first consider *possibly*, returning to the pseudo-cleft in (70). If *possibly* is interpreted as a modal, as we have assumed, removing it from the at-issue content would result in the two-dimensional meaning in (91). The at-issue content would be the proposition expressed by the pseudo-cleft in (91-a), without any contribution of *possibly*. The non-at-issue entailment would be (91-b), due to *possibly*.

(70) **Recall: possibly in pseudo-cleft**

What Obama approved was this bill and possibly that bill.
(91) **Hypothetical parenthetical reading**

a. What Obama approved was this bill and that bill.

b. Obama possibly approved this bill.

Yet, (91-a) and (91-b) do not together capture the meaning of (70). To see this, we need only consider the entailments of the pseudo-cleft component in (91-a), which include both (92-a) and (92-b). While the actual example in (70) does entail (92-a), it does not entail (92-b). (70) conveys that it is possible that Obama approved b2, but does not convey that he actually did. Because (70) lacks an entailment that (91-a) has, (91-a) cannot be part of what (70) asserts.

(92) **Expected entailments**

a. Obama approved b1.

b. Obama approved b2.

Despite the entailment complications, there have been theories of epistemic operators which do take their contribution to be non-at-issue. As Hacquard & Wellwood (2012) put it, “they [might] not [be] modals per se, but rather illocutionary markers which express a speaker’s comment about, or commitment to, the proposition expressed by the prejacent” (p. 2). Drubig (2001) and Westmoreland (1998), for instance, propose that epistemics are a kind of evidential, and evidentials contribute non-at-issue meaning (e.g. Faller 2002, Murray 2010, 2017).

Non-at-issue analyses of epistemics were proposed in large part to explain a celebrated property: that epistemics resist embedding under other operators. As we saw, non-at-issue entailments do not in general interact with other operators. The problem with this approach, however, is that its prediction is too strong: it predicts that epistemics should never embed in environments from which non-at-issue entailments cannot embed — and there are in fact cases where epistemics do embed. Hacquard & Wellwood (2012) review the relevant data, and report natural occurring examples from corpora. Most of the embedding data in the literature, including Hacquard & Wellwood’s corpus data involve the epistemic auxiliaries *might* and *must*. Two illustrative examples are given in (93). Example (93-a) is due to Papafragou (2006), and (93-b) is due to von Fintel & Gillies (2007).

(93) **Embedding epistemic auxiliaries**

a. Must Alfred have cancer?

b. If there might have been a mistake, the editor will have to re-read the manuscript.

The question in (93-a) asks whether or not it is epistemically certain that Alfred has cancer, with the contribution of must interpreted below the question forming operator. Similarly, in (93-b), must is interpreted inside the antecedent of the conditional. (93-b) conveys that the epistemic possibility of an error triggers a re-read. Parallel examples can be constructed with *possibly* itself:

(94) **Embedding possibly**

a. Could you possibly help me?

b. If there was even possibly a mistake, the editor will have to re-read the manuscript.
Hacquard & Wellwood (2012) conclude that epistemic operators must have truth-conditional content and are best analyzed as modals. Their resistance to embedding, they propose, should be attributed to pragmatic factors. I concur with them, and maintain that possibly is a modal, and thus not parenthetical in the pseudo-cleft example. The composition problem remains. The situation with the conditional example in (79) is parallel. Since (79), like its possibly counterpart, does not entail that Obama will approve b2, the conditional, analyzed with a modal, seems to have truth-conditional import.

(79) **Recall: conditional in pseudo-cleft**
What Obama will approve is this bill and, if Biden let him, that bill.

The possibility of embedding is one way to support the idea that conditionals are at-issue. Examples (95)-(96) illustrate with a question and negation. In each case, the continuation in (b) makes clear that the conditional is interpreted below the embedding operator.

(95) **Conditional under a question**
   a. Is it the case that the Republicans will revolt if Obama approves THIS bill?
   b. No. They’ll revolt if he approves THAT bill.

(96) **Conditional under negation**
   a. It is not the case that the Republicans will revolt if Obama approves THIS bill.
   b. They’ll revolt if he approves THAT bill.

I conclude that possibly and conditionals do have truth-conditional import and, accordingly, that the composition problem arises in the corresponding pseudo-clefts. I am less certain about with difficulty. With difficulty has a different profile from possibly and the conditional. On the one hand, introducing possibly and the conditional into the second conjunct in the pseudo-cleft bleeds the inference that Obama approved b2. For a modal to bleed an inference otherwise contributed by the at-issue content, they seemingly must be at-issue. On the other hand, with difficulty adds a new inference, that he approved b2 with difficulty. That new inference could be at-issue or non-at-issue and the correct overall meaning would derive. The pseudo-cleft is repeated below. If with difficulty were non-at-issue, the pseudo-cleft would be interpreted:

(24) **Recall: with difficulty in pseudo-cleft**
What Obama approved was this bill and, with difficulty, that bill.

(97) **Hypothetical parenthetical reading**
   a. What Obama approved was this bill and that bill.
   b. Obama approved that bill with difficulty.

The at-issue pseudo-cleft in (97-a) could be interpreted with an identity semantics and convey that Obama approved b1 and b2 and no other bill, while the non-at-issue entailment in (97-b) adds the inference that approving b2 occurred with difficulty. The prediction accords with intuitions. To assess the plausibility of treating with difficulty as parenthetical, we can attempt to run “the family of sentences” test. For one illustration, consider the example in (98), where a sentence containing with difficulty is interrogated:
With difficulty in a question
Did Obama approve this bill with difficulty?

In this example, with difficulty cannot be parenthetical. If it were, its non-at-issue contribution would project out of the question environment. This would result in (98) having as its at-issue content the question in (99-a), and having (99-b) as an additional non-at-issue entailment. Since (99-b) entails that Obama did approve this bill, it entails an answer to (99-a). As such, (98) would ask a question, while at the same time entailing an answer — which should lead to infelicity.

Hypothetical parenthetical reading for (98)

a. Did Obama approve this bill?
   (at-issue)
b. Obama approved this bill with difficulty.
   (non-at-issue)

In fact, (98) does have a felicitous reading, where the contribution of with difficulty is interpreted as at-issue content within the scope of the question operator, as in the paraphrase in (100-a). If Obama did approve that bill, but found it easy, then (98) could be felicitously answered (100-b).

The observed reading

a. "Did [Obama approve this bill and do so with difficulty]?
   b. No. He approved it, but he did so easily.

Example (101) with negation covers. On the bracketing shown, a parenthetical reading of with difficulty would result in the entailments in (102). Negation operates on the at-issue dimension, but leaves the non-at-issue dimension unaffected. The resultant entailments are contradictory, since (102-a) entails ¬(102-b). There is in fact a non-contradictory reading of (102) on which it naturally conveys the same as (103-a), with with difficulty below negation. The continuation in (103-b) supports the reading.

With difficulty under negation
Obama didn’t [approve this bill with difficulty].

Hypothetical parenthetical reading for (101)

a. Obama didn’t approve this bill.
   b. Obama approved this bill with difficulty.

The observed reading

a. “It’s not the case that [Obama approved this bill and did so with difficulty].”
   b. He approved it, but did so easily.

Deploying the “family of sentences” test shows that with difficulty is not necessarily parenthetical. There is no evidence for an additional parenthetical reading — but also no evidence against one. In the data seen so far, the parenthetical reading comes out pragmatically problematic and, as such, would not be detected. One data point which at first seems to conclusively disprove a parenthetical reading is:

With difficulty under claimed
John claimed that Obama approved this bill with difficulty.
Interpreting *with difficulty* as parenthetical in (104) does not yield an obviously problematic result. The at-issue entailment would be (105-a), and the non-at-issue entailment (105-b), if the contribution of *with difficulty* took matrix scope above John claimed. The two entailments in (105) are consistent. Together, they convey that John claimed that Obama approved this bill and, moreover, that Obama actually did approve it and did so with difficulty.

(105) **Hypothetical parenthetical reading for (104)**

a. John claimed that Obama approved this bill.
b. Obama approved this bill with difficulty.

Despite being not pathological, the reading in (104) does not seem to be available. The only way to interpret (104) is with the contribution of *with difficulty* as part of what John claimed, as in (106). In a case where the parenthetical reading could emerge, it does not, suggesting it does not exist.

(106) **The observed reading**

"John claimed [that Obama approved this bill and that he did so with difficulty]."

Yet, this test is confounded. The environment in (104) is one where true parentheticals also fail to project. In (107), for instance, the non-at-issue entailment that McCain proposed the bill also seems to be interpreted as part of John’s claim, and cannot easily project to the matrix level. The contrast between (107) and the earlier “family of sentences” data in (90) is puzzling, but it means that the behavior of *with difficulty* in (104) is not counter-evidence to a parenthetical parse being possible.

(107) **Baseline with a prototypical parenthetical**

John claimed that Obama approved this bill, which McCain proposed.

Aggregating all of the data discussed, we have seen that *with difficulty* must have a non-parenthetical reading, and we have failed to find evidence for or against an additional parenthetical reading. I remain agnostic on the status of *with difficulty*. Regardless, we have seen that *possibly* and the conditional are most naturally analyzed with truth-conditional import, so the composition problem remains. The solution I propose for *possibly* and the conditional will extend immediately to *with difficulty* as well, if that adverb is not parenthetical. Now, I develop my solution.

4.5 Taking stock

This section has introduced a new composition problem for pseudo-clefts. When extra material is inserted into the post-copular TP, the proposition the TP expresses is not an answer to the pre-copular question. The Identity Hypothesis predicts the data to be contradictory — when, intuitively, they are contingent. I illustrated the problem with *possibly*, a conditional antecedent, and *with difficulty*. To solve the problem, I will abandon the Identity Hypothesis altogether and propose a new LF for pseudo-clefts. For a further attempt to reconcile the adverb data with the Identity Hypothesis, I refer the reader to Appendix A.
5 Establishing a baseline

Common to all of the pseudo-cleft data we have seen is an exhaustivity inference. Even when adverbs like possibly are present in the pseudo-cleft, exhaustivity still naturally arises. The full complement of data, with their associated exhaustivity inferences, are gathered together:

(108) Exhaustivity across examples

a. What Obama approved was this bill.
   "Obama approved b1, and no other bill."

b. What Obama approved was this bill and possibly that bill.
   "Obama approved b1 and possibly b2, and no bill other than b1 or b2."

c. What Obama will approve is this bill and, if Biden lets him, that bill.
   "Obama will approve b1 and, if Biden lets him, b2, and no bill other than b1 or b2."

d. What Obama approved was this bill and, with difficulty, that bill.
   "Obama approved b1 and with difficulty b2, and no bill other than b1 or b2."

I believe the key to the semantics of pseudo-clefts lies in understanding the source of exhaustivity. Under the Identity Hypothesis, the source of exhaustivity was ANS. By returning the strongest true answer to the pre-copular question, ANS predicted exhaustivity in (108-a) — but the account could not extend to (108-b)-(108-d). The source of exhaustivity must be something else.

Although the Identity Hypothesis did not fit the bill, question-answer ('Q-A') pairs remain a useful baseline, since the full range of pseudo-cleft data have a counterpart Q-A pair. The answer in (109-a) is the strongest true answer to the question, while (109-b) contains possibly, (109-c) contains a conditional, and (109-d) contains with difficulty. Just as the pseudo-clefts are interpreted exhaustively, the corresponding answers in (109) are most naturally exhaustive, as well.

(109) Baseline Q-A pairs

What did Obama approve?

a. This bill.
   "Obama approved b1, and no other bill."

b. This bill and possibly that bill.
   "Obama approved b1 and possibly b2, and no bill other than b1 or b2."

c. This bill and, if Biden lets him, that bill.
   "Obama will approve b1 and, if Biden lets him, b2, and no bill other than b1 or b2."

d. This bill and, with difficulty, that bill.
   "Obama will approve b1 and, if Biden lets him, b2, and no bill other than b1 or b2."

A way forward presents itself. If we can figure out how exhaustivity arises in the examples in (109), then we may discover the composition of pseudo-clefts. I pursue that idea: that exhaustivity in (109-a)-(109-c) and exhaustivity in pseudo-clefts have the same source. In this section, I set aside pseudo-clefts,
and derive the exhaustive meanings in (109-b)-(109-d) in a particular way which will shed light on pseudo-clefts when we return to them in the next section.

5.1 A basic exhaustive answer

How *does* exhaustivity come about in answers to questions? One idea might be that this is purely a pragmatic matter. When someone asks the question in (109), they want to know everything that Obama approved. More technically, they want to be able to identify the strongest true proposition in the Hamblin set for the question. Thus, a cooperative responder should provide that proposition, or at least provide as much relevant information as they can. In (109-a), if the responder is assumed to name everything that Obama approved, the exhaustivity inference is explained. The situation in (109-b) would be similar, with the responder in this case providing more information than solicited. In (109-b) and (109-c), the assumption would have to be that the responder names everything that they think Obama *might* have approved. Then, again, the exhaustivity effect would obtain.

But, there is another possibility: that exhaustivity is grammatically encoded in the answer. At least in some cases, this is uncontroversially right. These are cases where the answer contains overt *only*. The simple example in (110) illustrates without the complexities of conjunction and adverbs. When *only* does not appear, as in (111), the covert counterpart Exh could be present, to the same effect as *only*. With *only* or Exh in the syntax, the answer entails exhaustivity.

(110) Question-answer pair with *only*
   a. What did Obama approve?
   b. He *only* approved this bill.

(111) Parallel question-answer pair with Exh
   a. What did Obama approve?
   b. He <Exh> approved this bill.

I will workshop the example in (111) with Exh for illustration. Exh is a two-place operator. One argument is a set of propositions (ALT), and the other is a proposition (the prejacent, p). Exh asserts that p is true, and negates alternatives in ALT not already entailed by p. Exh differs from *only* just in that it asserts the prejacent, while *only* presupposes the prejacent.

(112) Defining Exh
   \[ [\text{Exh}]^\text{w}(\text{ALT})(p) = 1 \iff p(w) \land \forall p' \in \text{ALT} [p'(w) \rightarrow p \subseteq p'] \]

The LF for the answer in (111) is (113). Exh attaches to the TP, and the prejacent of Exh is the intension of that TP: that Obama approved this bill and that bill (i.e. \( \phi_1 \land \phi_2 \)). Exh composes with the TP through IFA. The ALT argument requires more discussion.
(113) LF for the answer in (111)

```
TP₂
  \text{Exh} \quad C₁ \quad \text{TP₁}
```

Obama approved this bill

(114) Prejacent of only in (111)

\[ [\text{TP₁}]_Q = \phi₁ \]

Following Rooth (1992b), I take the ALT argument to come from a covert element, which Rooth indicates as \( C \). \( C \) is a free pronoun and, as such, its semantic value is determined by the contextually given variable assignment. This is shown for an arbitrary assignment \( g \) which contains 1 in its domain:

(115) Interpreting \( C \)

\[ [C₁]^{w,g} = g(1) \]

Because Exh requires a set of propositions as its argument, for composition to proceed, \( g \) must map 1 to a set of propositions. In turn, because the assignment function is determined by context, it is up to the context to provide a salient set of propositions to be taken as \( g(1) \). In the question-answer pair, the question denotes its Hamblin set and, as such, makes that set salient:

(116) Hamblin denotation for question

a. \[ [\text{what did Obama approve}]^w \]
b. \[ = \lambda p_x . \exists X \ [p = \lambda w'. \forall x <_{\text{AT}} X \ [\text{Obama approved } x \text{ in } w']] \]

In our short question-answer dialog, I take it that \( g \) maps 1 to the Hamblin set provided by the question.\(^7\) In this way, \( C \) in the answer comes to be interpreted in exactly the same way as the extra-sentential question. This is achieved via an anaphoric relation between \( C \) and the question, mediated through the variable assignment. To illustrate the current system, it will be most instructive to assume that there are three salient bills — \( b₁, b₂, \) and \( b₃ \) — and focus on just three members of the Hamblin set:\(^8\)

(117) ALT argument of only in (111)

a. \[ [C₁]^{w,g} = g(1) \]
b. \[ = [\text{what did Obama approve}]^w \]
c. \[ \approx \{ \phi₁, \phi₂, \phi₃ \} \]

Now, how do the arguments of Exh conspire to yield an exhaustive interpretation? Referring to the interpretation schema for only in (112) above, we can fill in \([\text{TP}]_Q\) for \( p \), and \([C₁]^{g}\) for ALT. The resultant meaning is computed in (118):

\(^7\)See Appendix B for a more detailed discussion of how the value of \( C \) is determined.

\(^8\)The Hamblin set would, in addition, contain conjunctive alternatives (e.g. \( \phi₁ \land \phi₂ \)), but we can ignore these from this point on without affecting the overall predictions.
By asserting the prejacent, Exh introduces an entailment that Obama approved b1. The rest of the assertive component says that any alternative in C not entailed by \([\text{TP}_1]_C\) is false. \([\text{TP}_1]_C\) entails \(\phi_1\), but it does not entail \(\phi_2\) or \(\phi_3\) (which are logically independent of \([\text{TP}_1]_C\)). Those two alternatives are negated, introducing the entailment that Obama did not approve b2 or b3. Collapsing presupposition and assertion, the predicted inferences are (a) that Obama approved b1, and (b) that he did not approve b2 or b3 — capturing the exhaustive meaning.

To summarize, exhaustivity can be grammaticality encoded in an answer by means of only or its covert counterpart Exh. The key features of the analysis, summarized below, lie in the argument structure of Exh, and how the required arguments are filled.

- The first argument of Exh is a set of propositions (ALT). In a question-answer pair, the ALT argument is indirectly provided by the question: an anaphoric element in the answer picks up the Hamblin set for the question as its value.

- The second argument of Exh is a proposition (its prejacent).

Having spelled out an analysis for exhaustive answers to questions for basic data, we are ready to return to insert adverbs and conditionals, and consider how exhaustive meanings arise in those data.

### 5.2 Possibly and the over-negation problem

The question-answer pair in (119) (repeated) has possibly within the exhaustive answer. I assume, as before, that the source of exhaustivity is Exh.

(119) **Exhaustive answer with possibly**

a. What did Obama approve?

b. This bill and possibly that bill.

(120) **Paraphrase of (119-b)**

"Obama approved b1 and possibly b2, and no other bill."
What meaning is predicted from this LF? The prejacent of Exh is the intension of the TP: the proposition that Obama approved b1, and there is an epistemically accessible world that he approved b2. Assuming that C is again anaphoric to the Hamblin set for the question, the ALT argument is the set \{\phi_1, \phi_2, \phi_3\}. These, then, are the ingredients:

Exh asserts its prejacent and thus, in this case, asserts \(\phi_1 \land \lozenge \phi_2\). Which alternatives in C will Exh negate? Exh negates any alternative not entailed by the prejacent. Clearly, \(\phi_1\) is entailed and \(\phi_3\) is not. Hence, Exh will not negate \(\phi_1\), but will negate \(\phi_3\). A problem arises with respect to \(\phi_2\). The prejacent says that \(\phi_2\) is epistemically possible, but it does not entail that \(\phi_2\) is true: \(\lozenge \phi_2\) does not entail \(\phi_2\). As such, the prediction is that Exh should negate \(\phi_2\). The predicted overall meaning is:

This meaning has some of the right properties: the exhaustive answer does convey the first two conjuncts (that Obama approved b1 and possibly b2) and the last one (that Obama didn’t approve b3), but the third conjunct is clearly wrong. Although not contradictory, \(\neg \phi_2\) is intuitively paradoxical in conjunction with \(\lozenge \phi_2\). The underlined portion would paraphrase:

The conjunction in (125) closely resembles what is known as Moore’s (1942) paradox, illustrated in (126) (see also Wittgenstein (1953), i.a.). Since the sentences in (126-a) and (126-b) may both be simultaneously true, the conjunction in (126-c) is logically contingent. Moore’s paradox is that (126-c), despite being contingent, is perceived as “absurd” or almost contradictory.

---

I assume that the prejacent need not be an element of ALT. For discussion of this assumption, see Appendix B. There, I justify the assumption, but also show that the results in this section do not crucially depend on it. The same issues arise with other conceivable ways of determining the value for C. The analysis of pseudo-clefts will, however, depend on this.
(126) **Moore’s paradox**
   a. It’s raining.
   b. I don’t believe that it’s raining.
   c. #It’s raining and I don’t believe that it’s raining.

The conjunction in (125) has an identical profile to (126-c). In both, one conjunct asserts a proposition, while the other conjunct asserts that this proposition is not epistemically certain. And both have a similar intuitive status: like (126-c), (125) is perceived as absurd or almost contradictory. In effect, then, the prediction is that the exhaustive answer should exhibit Moore’s paradox — which it does not. There is nothing absurd or contradictory about (119-b). To derive the correct meaning for (70), we must prevent Exh from negating $ϕ_2$:

(127) **Target meaning**
   $ϕ_1 \land \Diamond ϕ_2 \land ϕ_3$

As far as I know, the problem discussed here has not been previously observed, and I will refer to it as the over-negation problem, since Exh is predicted to negate too many alternatives. Exh “over-negates” $ϕ_2$. I state the problem as (128) in a general way for Exh and only:

<table>
<thead>
<tr>
<th>(128) <strong>The over-negation problem</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Exh and only are predicted to strengthen ($ϕ \land \Diamond ψ$) to ($ϕ \land \Diamond ψ \land \neg ψ$).</td>
</tr>
</tbody>
</table>

Note that the problem does arise for overt only, as well as Exh. This can be seen in two ways. First, the data in (129) insert only into the exhaustive answer in (119-b). Although some informants find the overt only version to be degraded relative to its counterpart without only, most informants do accept (129-b) with the same meaning as (119-b).

(129) **Replicating with overt only**
   a. What did Obama approve?
   b. (?)Only this bill and possibly that bill.

In addition, the problem arises in other data whose grammatical status seems clear. The dialog in (130) illustrates. Suppose that speaker A hates everyone except for Bill and Sue, and will only go to the party if just one or both of them go and no one else. The conditional in (130-c) seems perfectly felicitous, where only takes scope over [Bill and possibly Sue], and the paraphrase is clearly not (131), diagnosing the over-negation problem.

(130) **Replicating elsewhere**
   A  Who’s going to the party?
   B  Bill and possibly Sue. Certainly no one else.
   A  Well, if only Bill and possibly Sue are going, I’ll go, too.

(131) **Predicted paraphrase for conditional**
   #If [Bill is going, and Sue is possibly going, but actually not going], I’ll go too.
A solution to the over-negation problem is required and, in the following, I develop one. Recall that three ingredients go into a construction with Exh or *only* — the operator, a set of alternatives, and the prejacent. There are, then, three ways we could try to resolve the over-negation problem, each modifying a different one of these ingredients. I follow each avenue in turn and argue that only one is truly viable: we must redefine the operator. I continue to illustrate with Exh, rather than *only*.

### 5.2.1 Option 1: Changing ALT

The most straightforward way to prevent Exh from negating 02 would be to remove 02 from the set of alternatives that Exh takes as argument. In (119-b), the restrictor argument is the Hamblin set for the question. I assumed that three bills were salient — b1, b2, and b3 — and, in turn, took the Hamblin set to be \{\phi_1, \phi_2, \phi_3\}. The salience of b1 and b2 was clearly established by their mention in the prejacent. We could, however, entertain the possibility that the Hamblin set may be restricted to a *subset* of the salient propositions of the form *that Obama approved x*. Then, \phi_2 could optionally be pruned. If Exh took just \{\phi_1, \phi_3\} as ALT, \phi_2 would be spared from negation.

\[
\begin{align*}
(132) & \quad \text{Restrictor with } \phi_2 \text{ pruned (a) and prejacent (b)} \\
& \quad a. \quad \llbracket C_t \rrbracket^\text{Alt} = \{\phi_1, \phi_3\} \\
& \quad b. \quad \llbracket TP \rrbracket^C = \phi_1 \land \diamond \phi_2 \\
(133) & \quad \text{Overall meaning correctly predicted} \\
& \quad \llbracket (119-b) \rrbracket^\text{Alt}^C = \phi_1 \land \diamond \phi_2 \land \neg \phi_3
\end{align*}
\]

This approach is, however, too general. Variability in the composition of the *first* argument of Exh (the restrictor) should be independent of the *second* argument (the prejacent). Thus, if \phi_2 could be pruned from ALT, it should be possible for Exh not to negate \phi_2 regardless of its prejacent. In fact, the availability of an over-negation reading crucially depends on the prejacent. While \phi_2 is not intuitively negated in (119-b) where \diamond \phi_2 is a conjunct in the prejacent, \phi_2 is negated when that conjunct is removed. The dialog in (134) clearly establishes the point:

\[
\begin{align*}
(134) & \quad \text{But, pruning is not generally available} \\
& \quad a. \quad \text{What did Obama approve?} \\
& \quad b. \quad \#I'm not sure if he approved that bill, but he only approved this bill.
\end{align*}
\]

\[
\begin{align*}
(135) & \quad \text{LF for second conjunct in (134-b)} \\
& \quad [\text{only(C)} \{\text{he approved this bill}\}]
\end{align*}
\]

In this case, overt *only* occurs with just \phi_1 as its prejacent. If \phi_2 could be pruned from that Hamblin set to yield \{\phi_1, \phi_3\}, (134-b) should express a natural thought: the first conjunct establishes that the speaker is uncertain about \phi_2, while the second would assert \(\phi_1 \wedge \neg \phi_3\). Intuitively, however, (134-b) does have the status of Moore's paradox: the second conjunct entails \(\neg \phi_2\) (in addition to \phi_1 and \neg \phi_3), and the first conjunct asserts that \neg \phi_2 is not epistemically certain, yielding perceived absurdity. This indicates that the Hamblin set must be \{\phi_1, \phi_2, \phi_3\}, no pruning possible.
On this basis, I conclude that pruning alternatives from the restrictor of Exh is not the right mechanism to solve the over-negation problem, and set about to develop a different approach. Whatever suppresses negation of $\phi_2$ in (70) must link to possibly $\phi_2$ being in the prejacent.

5.2.2 Option 2: Changing the prejacent

Exh will negate $\phi_2$ unless $-\phi_2$ contradicts the prejacent. To solve the over-negation problem by changing the prejacent, we would have to re-analyze the proposition expressed by (136) such that it entails $\phi_2$. I concluded earlier in Section 4.4 that possibly is best analyzed as a modal with truth-conditional import, in which case the prejacent asserts $\Diamond \phi_2$ and thus not entail $\phi_2$. I do not pursue Option 2 further.

5.2.3 Option 3: Changing Exh

A single option remains: the solution to the over-negation problem must come from re-defining Exh itself. The current definition of Exh is repeated below:

(112) **Current definition of Exh**

\[
[\text{Exh}]^w(\text{ALT})(p) = 1 \text{ iff } p(w) \land \forall p' \in \text{ALT} [p'(w) \rightarrow p \subseteq p']
\]

When Exh negates alternatives, it seeks to avoid a contradiction, and does so in a particular way. When it decides whether or not to negate an alternative $p'$, it looks at whether $-p'$ is logically consistent with the prejacent (which Exh asserts in a separate conjunct). Exh negates $p'$ unless to do so would logically contradict the asserted content of the prejacent. This can be shown more directly in an equivalent re-formulation of Exh:

(136) **Equivalent re-formulation**

\[
[\text{Exh}]^w(\text{ALT})(p) = 1 \text{ iff } p(w) \land \forall p' \in \text{ALT} [p \cap -p' \neq \emptyset \rightarrow -p'(w)]
\]

The over-negation problem is a demonstration that Exh as defined negates too many alternatives: $-\phi_2$ being logically consistent with the prejacent is not a sufficient condition for Exh to negate $\phi_2$ (though it may be a necessary condition). I maintain the underlying idea that Exh seeks to avoid contradictions, but I take the over-negation problem to show that we have not yet uncovered all the contradictions that Exh seeks to avoid. For inspiration, let us return to Moore's paradox, repeated in (137). Although different accounts of the paradox have been advanced, Moore’s own idea had to do with the inferences that are drawn from an assertion. Specifically, he noted that a speaker, by asserting a certain proposition, implies something about their epistemic state, namely that they believe that proposition. It then follows that (137), though logically contingent, is perceived as absurd or contradictory: the first conjunct implies that the speaker believes it is raining, and the second conjunct contradicts this implication.

(137) **Recall: Moore's paradox**

It's raining and I don’t believe that it’s raining.

I propose that Exh avoids a contradiction not just between the prejacent and $-p'$ (for some alternative $p'$), but also between the prejacent and the proposition that the speaker believes $-p'$ — i.e. the proposition that would be implied if $-p'$ ended up as part of the assertion. The revised definition for Exh follows.
I refer to this as *epistemic sensitive* Exh and, beginning in (139), abbreviate the proposition that the speaker believes \( \neg p' \) as '\( \text{B}(\neg p') \)'.

(138) **Epistemic sensitive Exh**

\[
\text{Exh}_{ep}\sqcup_c (\text{ALT})(p) = 1 \iff \\
p(w) \land \forall p' \in \text{ALT} \ [p \cap \neg p' \cap \lambda w'. \forall w'' \in \delta(\text{SP}(C))(w') \ [\neg p'(w'')] \neq \emptyset \rightarrow \neg p'(w)]
\]

(139) **Epistemic sensitive Exh (abbreviated)**

\[
\text{Exh}_{ep}\sqcup_c (\text{ALT})(p) = 1 \iff p(w) \land \forall p' \in \text{ALT} \ [p \cap \neg p' \cap \text{B}(\neg p') \neq \emptyset \rightarrow \neg p'(w)]
\]

Directly comparing Exh and Exh_{ep}, it is clear that Exh_{ep} will negate fewer alternatives than Exh. In effect, for Exh to negate an alternative \( p', p' \) must pass one test: \( \neg p' \) must be consistent with the prejacent. Now, for Exh_{ep} to negate \( p', p' \) must pass two tests: the test imposed by Exh and the additional test that \( \text{B}(\neg p') \) be consistent with the prejacent. Not all alternatives which pass the original test will pass the new test in addition. In this way, the stage is set to bring closure to the over-negation problem. The final LF for the exhaustive answer in (119-b) is (140), with Exh_{ep} in lieu of Exh. Which of the three Hamblin alternatives, \( \{\phi_1, \phi_2, \phi_3\} \), will Exh_{ep} negate?

(140) **Final LF for (119-b)**

\[
\text{TP} \quad \text{Exh}_{ep} \quad \text{C} \quad \text{TP} \\
\text{Obama \lambda [t\ approved this bill] and [possibly [t\ approved that bill]]}
\]

Like Exh, Exh_{ep} will not negate \( \phi_1 \), since the asserted content of \( \neg \phi_1 \) itself contradicts the prejacent, as in (141). Moreover, again like Exh, Exh_{ep} will negate \( \phi_3 \). \( \phi_2 \) passed the original test that its negation not contradict the prejacent, and it passes the new test just as well: \( \text{B}(\neg \phi_2) \) does not contradict the prejacent either, as in (142). I show the prejacent in red, and underline contradictory conjuncts.

(141) **Testing whether \( \phi_1 \) is excludable**

\[
\phi_1 \land \lnot \phi_2 \land \lnot \phi_3 \land \text{B}(\neg \phi_1)
\]

(142) **Testing whether \( \phi_3 \) is excludable**

\[
\phi_1 \land \lnot \phi_2 \land \neg \phi_3 \land \text{B}(\neg \phi_3)
\]

The upshot of moving to Exh_{ep} is realized with \( \phi_2 \). As we saw with Exh, \( \phi_2 \) passes the original test for negation, since \( \neg \phi_2 \) is consistent with the prejacent. Crucially, it does not pass the new test: \( \text{B}(\neg \phi_2) \) contradicts the prejacent. \( \text{B}(\neg \phi_2)' \) says that all worlds epistemically accessible to the speaker are \( \neg \phi_2'-
worlds, while possibly in the prejacent says that that some $\phi_2$-word is epistemically accessible. To avoid the contradiction, $\text{Exh}_{ep}$ does not negate $\phi_2$ — and the target meaning derives. The problem is solved.

(143) **Testing whether $\phi_2$ is excludable**

$\phi_1 \land \Diamond \phi_2 \land \neg \phi_2(w) \land B(\neg \phi_2)(w)$

(144) **Overall predicted meaning**

$\phi_1(w) \land \phi_2 \land \neg \phi_3$

Note, further, that failure to negate $\phi_2$ is now directly linked to $\Diamond \phi_2$ being part of the prejacent. Assuming that only is defined in parallel with epistemically accessible Exh, consider again (134) (repeated from discussion under Option 1). In (134), neither $\neg \phi_2$ nor $B(\neg \phi_2)$ contradicts the prejacent, $\phi_1$. With $\Diamond \phi_2$ removed from the prejacent, $\phi_2$ qualifies for negation — and the two conjuncts in (134) together express Moore’s paradox.

(134) **Recall: where $\phi_2$ is negated**

a. What did Obama approve?

b. I’m not sure if he approved that bill, but, he only approved this bill.

Epistemically sensitive only cannot introduce Moore’s paradox internal to the only statement itself, but its output can be a player in Moore’s paradox together with a separate conjunct. At least in the core data, the proposal predicts where over-negation readings are — and are not — observed.

5.2.4 **Concluding possibly**

This subsection has presented a theory of how exhaustivity arises in answers to questions when the answer contains possibly. The source of exhaustivity remains a covert Exh. In order to capture the data, as well as related examples with overt only, a revision to the semantics of Exh/only was required: they must be epistemically sensitive, so that they negate an alternative $\phi$ only if both $\neg \phi$ and $B(\neg \phi)$ are logically consistent with the prejacent.

5.3 **Adding a conditional**

We are ready to turn to the question-answer pair in (145) (repeated), where the exhaustive answer contains a conditional. The source of exhaustivity is, I take it, again a covert $\text{Exh}_{ep}$.

(145) **Exhaustive answer with conditional**

a. What will Obama approve?

b. This bill and, if Biden lets him, that bill.

(146) **Paraphrase of (145-b)**

“Obama will approve b1 and, if Biden lets him, b2, and no bill other than b1 or b2.”
(147) **LF for (145-b)**

![Diagram](image)

O λ1 [t1 approved this bill] and [[if Biden lets him] □] [t1 approved that bill]

What meaning is predicted now? Recall that the prejacent expresses the proposition in (148): that Obama will approve b1, and in all epistemically accessible worlds compatible with what Biden allows, he will approve b2. Taking C to be anaphoric to the question, the restrictor argument remains the same as in the last two sections: the Hamblin set \(\{\phi_1, \phi_2, \phi_3\}\), as in (149).

(148) **The prejacent of Exh**

a. \(\lbrack TP \rbrack^C_C\)

b. \(= \lambda w . \phi_1(w) \land \forall w' \in \delta'(SP(C))(w) [\text{Biden lets him approve } b_2 \text{ in } w' \rightarrow \phi_2(w')]\)

c. \(\sim\) (abbreviation) \(\phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2]\)

(149) **The ALT argument**

\([C_1]^{\sim\subseteq} = \{\phi_1, \phi_2, \phi_3\}\)

Clearly, Exh asserts the prejacent \(\lbrack \phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2]\rbrack\) — and negates \(\phi_3\), but not \(\phi_1\). \(\phi_3\) passes both tests for excludability, since the prejacent is consistent with both \(\lnot \phi_3\) and \(\text{B}_{\lfloor \phi_3 \rfloor}\). On the other hand, \(\phi_1\) fails the original test for excludability, since \(\lnot \phi_1\) itself contradicts the prejacent.

Now, what about \(\phi_2\)? A problem seems to arise again: \(\phi_2\) should get negated, since \(\lnot \phi_2\) and \(\text{B}(\lnot \phi_2)\) are both consistent with the prejacent. To make this clear for \(\text{B}(\lnot \phi_2)\), consider a world \(w_1\) such that (a) \(\phi_1\) is true at \(w_1\), (b) \(\delta'(s)(w_1)\) contains no world at which \(\phi_2\) is compatible with Biden’s wishes, and (c) all worlds in \(\delta'(s)(w_1)\) are \(\lnot \phi_2\)-worlds. The prejacent is true at \(w_1\): \(\phi_1\) is true (given (a)), and the conditional is vacuously true, since its restrictor is empty (given (b)). Given (c), \(\text{B}(\lnot \phi_2)\) is true, as well. Hence, \(\phi_2\) passes the tests for excludability, and the overall answer should mean:

(150) **Wrongly predicted meaning**

\(\phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2] \land \lnot \phi_2 \land \lnot \phi_3\)

The third conjunct, negating \(\phi_2\), is incorrect, just as it was in the possibly data. The underlined portion would paraphrase as (151), which seems intuitively “absurd” or paradoxical and clearly is not what the exhaustive answer conveys:

(151) **Paraphrase of underlined portion**

“If Biden lets him, Obama will approve b2, but he actually won’t approve b2.”

150
The problem is very general, and replicates with overt only in (152). The answer in (152-b) is perfectly felicitous, and has a parallel meaning to the counterpart without only.

(152) Replicating in a question-answer pair

What will Obama approve?

a. This bill and, if Biden lets him, that bill.
b. Only this bill and, if Biden lets him, that bill.

We have arrived at a second over-negation problem and, from discussion so far, it appears as if the move to Exh_{ep} which solved the over-negation problem with possibly may not be enough to solve the problem with the conditional.

(153) Another over-negation problem

Exh and only are predicted to strengthen \([\phi \land (\psi \rightarrow \xi)]\) to \([\phi \land (\psi \rightarrow \phi) \land \neg \psi]\).

In fact, the move to Exh_{ep} does solve the problem, once we re-evaluate the semantics of the conditional. It is important to flag a notable feature of the world \(w_1\) at which the prejacent and \(B(\neg \phi_2)\) are both true: that the restrictor of the conditional is empty at that world. Universal quantifiers do not tolerate empty restrictors. To make this visible to Exh_{ep}, I assume that this constraint is not purely pragmatic, but arises from a semantic presupposition: the conditional in (153) presupposes a non-empty restrictor. Projecting that presupposition from the second conjunct results in the prejacent of Exh_{ep} being revised to:

(154) The prejacent (revised)

\[
\begin{align*}
A: & \quad \phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2] \\
P: & \quad \Diamond \phi_{\text{Biden}} 
\end{align*}
\]

The semantic presupposition creates a partial function. The proposition expressed by the prejacent is true iff there is an epistemically accessible \(\phi_{\text{Biden}}\)-world (satisfying the presupposition), and the assertion is true. The proposition is false iff there is an epistemically accessible \(\phi_{\text{Biden}}\)-world, and the assertion is false. At a world like \(w_1\) from which no \(\phi_{\text{Biden}}\)-world is epistemically accessible, the proposition is now undefined. The situation is summarized:

(155) Conditions for the prejacent to be ...

\[
\begin{align*}
a. & \quad \Diamond \phi_{\text{Biden}} \land \phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2] & \quad \text{... true} \\
b. & \quad \Diamond \phi_{\text{Biden}} \land \neg[\phi_1 \land [\phi_{\text{Biden}} \rightarrow \phi_2]] & \quad \text{... false} \\
c. & \quad \neg \Diamond \phi_{\text{Biden}} & \quad \text{... undefined}
\end{align*}
\]

Assuming that two propositions are "consistent" only if both can be simultaneously true, \(B(\neg \phi_2)\) is no longer consistent with the prejacent. The prejacent is only true at worlds where (155-a) holds and conjoining \(B(\neg \phi_2)\) with (155-a) yields a contradiction, as shown in (156). From worlds in which (155-a) holds, there is an epistemically accessible Biden-world. Since the conditional, \(\phi_{\text{Biden}} \rightarrow \phi_2\), says that \(\phi_2\) is true at every accessible Biden-world, it follows that there is also an epistemically accessible \(\phi_2\)-world. That is, \(\Diamond \phi_2\) follows — and \(\Diamond \phi_2\) and \(B(\neg \phi_2)\) contradict.
Testing whether $\phi_2$ is excludable

a. $\Diamond \left( \text{Biden} \land \phi_1 \land \text{Biden} \rightarrow \phi_2 \right) \land \neg \phi_2 \land B(\neg \phi_2)$
b. $\leftrightarrow \Diamond \left( \text{Biden} \land \phi_1 \land \text{Biden} \rightarrow \phi_2 \right) \land \Diamond \phi_2 \land \neg \phi_2 \land B(\neg \phi_2)$

Hence, if the universal modal in the conditional semantically encodes an existence presupposition, $\phi_2$ does not pass the tests for excludability with $\text{Exhep}$ defined exactly as before. The right overall meaning is now derived for the exhaustive answer in (145-b):

Correctly predicted meaning

A: $\phi_1 \land \left[ \phi_{\text{Biden}} \rightarrow \phi_2 \right] \land \neg \phi_3$

P: $\Diamond \phi_{\text{Biden}}$

The prejacent is asserted and $\phi_3$ is negated, but $\phi_1$ and $\phi_2$ are not. This results in the target meaning: Obama will approve $b_1$ and, if Biden lets him (which is a possibility), $b_2$ (from the prejacent, given its presupposition), but not $b_3$ (from negation of $\phi_3$).

5.4 Adding with difficulty

This brings us to the final adverb: *with difficulty*. I will give *with difficulty* the benefit of the doubt and assume that it does have truth-conditional import. The correct exhaustive meaning for the answer in (158-b) follows from the system exactly as it stands.

Exhaustive answer with *with difficulty*

a. What did Obama approve?

b. This bill and, with difficulty, that bill.

Paraphrase of (158-b)

"Obama approved $b_1$ and, with difficulty, $b_2$, and no bill other than $b_1$ or $b_2$."

LF for (158-b)

The prejacent of Exh expresses the proposition in (161) and the ALT argument remains, once again, the set $\{\phi_1, \phi_2, \phi_3\}$, assuming that C is anaphoric to the Hamblin set for the question.
(161) The prejacent of Exh
   a. $\text{TP}_{\text{C}} = \phi_1 \land \phi_2 \land \text{approving b2 occurred with difficulty}$
   b. $\sim (\text{abbreviation}) \phi_1 \land \phi_2 \land \text{difficult-b2}$

(162) The ALT argument
   $[C_1]^\text{wg} = \{\phi_1, \phi_2, \phi_3\}$

Exh asserts the prejacent and straightforwardly negates $\phi_3$ (which is logically independent of the prejacent), but not $\phi_1$ (which the prejacent entails). The situation with respect to $\phi_2$ is equally straightforward this time. As noted in Section 4.4, the effect of introducing with difficulty into the prejacent is very different from the effect of possibly or the conditional. Whereas possibly and the conditional bleed $\phi_2$ as an entailment of the prejacent, with difficulty does not. The prejacent in (161) still entails $\phi_2$, along with the extra entailment of difficult-b2. Hence, $\phi_2$ fails the original test for excludability: the prejacent is not consistent with $\sim \phi_2$:

(163) Testing whether $\phi_2$ is excludable
   $\phi_1 \land \phi_2 \land \text{difficult-b2} \land \sim \phi_2 \land B(\sim \phi_2)$

The correct exhaustive meaning for the overall answer follows, as in (164): Obama approved b1 and b2, and approved b2 with difficulty (all from the prejacent), and he did not approve b3 (the exhaustive entailment from negating $\phi_3$).

(164) Overall predicted meaning
   $[(158-b)]^\text{gc} = \phi_1 \land \phi_2 \land \text{difficult-b2} \land \sim \phi_3$

5.5 Local summary

This section began with the observation that exhaustive meanings with adverbs exactly like those in pseudo-clefts are observed in answers to questions. The section then diagnosed how exhaustivity comes about in answers — with and without adverbs. The source of exhaustivity was a covert Exh. By considering the adverb data, we discovered a new over-negation problem for Exh/only. The solution came from a move to epistemically accessible Exhep. Whereas Exh negates an alternative $\phi$ just in cases $\sim \phi$ is consistent with the prejacent, Exhep is more constrained: it negates $\phi$ just in case both $\sim \phi$ and $B(\sim \phi)$ are consistent with the prejacent. Exhep predicted the observed exhaustive meanings with possibly and conditionals, as well as with difficulty. Our task now is to develop an analysis of pseudo-clefts based on the independent results of this section.

6 Proposal for pseudo-clefts

The pseudo-cleft data which we intend to capture are repeated in (165), along with their paraphrases, which are perfectly identical to those of the exhaustive answers from the last section.
The pseudo-cleft data

a. What Obama approved was this bill. = (1)
   "Obama approved b₁, and no other bill."

b. What Obama approved was this bill and possibly that bill. = (70)
   "Obama approved b₁ and possibly b₂, and no bill other than b₁ or b₂."

c. What Obama will approve is this bill and, if Biden lets him, that bill. = (79)
   "Obama will approve b₁ and, if Biden lets him, b₂, and no bill other than b₁ or b₂."

d. What Obama approved was this bill and, with difficulty, that bill. = (24)
   "Obama approved b₁ and with difficulty b₂, but no bill other than b₁ or b₂."

Before presenting the Identity Hypothesis for pseudo-clefts, we had a skeletal structure — which is repeated in (166). I maintain Ross's idea that the pre-copular constituent is interpreted as a question (derived syntactically either from an interrogative CP or free relative DP with a concealed question operator). The post-copular constituent is a TP expressing a proposition.

Skeleton of a pseudo-cleft

The Identity Hypothesis filled in the skeleton one way. It added ANS to derive from the pre-copular question its strongest true answer, and the copula was interpreted as an identity predicate equating that strongest true answer with the post-copular proposition. I accept the basic skeleton in (166), but propose a new way of filling it in. The key idea is that the compositional "glue" allowing the pre- and post-copular constituents to compose with one another is a covert Exhep — a proposal I develop over the following subsections. The effect will be that pseudo-clefts have an LF which, in all important respects, is identical to their exhaustive answer counterpart. The interpretation of all the pseudo-clefts in (166) will "follow for free" from the proposals in the last section for how exhaustive answers are interpreted.

6.1 The basic proposal

To present the proposal, I home in on the simple example in (1), without an adverb. The key observation is this: the pseudo-cleft "wears on its sleeve" the arguments of Exhep. The pre-copular constituent is semantically a question and the intension of the post-copular constituent is a proposition. There is a set of propositions (the Hamblin set for the pre-copular question), and a proposition (from the post-copular TP). Based on this observation, I propose the following LF:
The main operator is covert Exhep, and it composes directly with the overt pre- and post-copular material. The post-copular TP expresses the proposition in (168), and that constitutes the prejacent of Exhep. Although the copula is in the scope of Exhep, I take the copula to be semantically inert.

(168) The prejacent of Exhep

\[ \llbracket TP \rrbracket_C = \phi_1 \]

The ALT argument comes from the pre-copular constituent. As we have seen repeatedly, its Hamblin denotation is a set of propositions that Obama approved X, illustrated \( \{\phi_1, \phi_2, \phi_3\} \).

(169) The ALT argument

\[ \llbracket \text{what Obama approved} \rrbracket^w = \{\phi_1, \phi_2, \phi_3\} \]

Now, everything functions in a familiar way. Exhep asserts the prejacent, and negates excludable alternatives in ALT. In this case, Exhep negates \( \phi_2 \) and \( \phi_3 \) (since \( \neg \phi_2 \) and \( B(\neg \phi_3) \) are both consistent with the prejacent and, likewise, for \( \neg \phi_3 \) and \( B(\neg \phi_3) \)), but not \( \phi_1 \) (\( \neg \phi_1 \) contradicts the prejacent). The correctly predicted meaning is the proposition in (170): that Obama approved b1, but not b2 or b3. Exhaustivity is derived without ANS, through Exhep. This is all there is to the proposal, and its main features are summarized in the box below the predicted meaning.

(170) Overall predicted meaning

a. \( \llbracket (1) \rrbracket^w_C = [\llbracket Exhep \rrbracket^w_C(\llbracket \text{what Obama approved} \rrbracket^w)(\llbracket TP \rrbracket_C)] \)

b. \( = 1 \iff \phi_1(w) \land \neg \phi_2(w) \land \neg \phi_3(w) \)

(171) Summary of proposal for pseudo-clefts

a. Covert Exhep composes with the pre- and post-copular constituents.

b. The proposition given by the post-copular TP = the prejacent.

c. The Hamblin set for the pre-copular question = ALT.

With the proposal in place, pseudo-clefts are interpreted in exactly the same way as exhaustive answers. For direct comparison, the LF for the exhaustive answer counterpart to (1) is repeated:

155
Recall: question-answer pair
a. What did Obama approve?
b. Obama approved this bill and that bill.

LF for (111-b)

The TP in the answer is syntactically and semantically identical to the post-copular TP in the pseudo-cleft, so the prejacent of Exhep is the same in both structures. The ALT argument is determined in different ways — but they are two routes to the same semantic end. In a question-answer pair, recall that ALT is provided by a covert pronominal element, C. C is interpreted anaphorically and picks up the Hamblin set for the extra-sentential question as its value. In this way, C in the LF for (111-b) gets interpreted as [[what did Obama approve]], the Hamblin set for the question in (111-a). In the pseudo-cleft, the question is present intra-sententially and directly provides the ALT argument. Whether it comes about indirectly via anaphora with an extra-sentential question (as in (111-b)) or from an intra-sentential occurrence of the question (as in (1)), the ALT argument is the Hamblin set for 'What did Obama approve?'. Given that the prejacent and ALT are semantically identical in answers and pseudo-clefts, the overall interpretation is identical.

Summary pseudo-clefts vs. exhaustive answers

<table>
<thead>
<tr>
<th>Main operator</th>
<th>Prejacent</th>
<th>ALT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo-clefts</td>
<td>Exhep</td>
<td>Post-copular TP</td>
</tr>
<tr>
<td>Answers</td>
<td>Exhep</td>
<td>Overt TP</td>
</tr>
</tbody>
</table>

The proposal maintains a parallel between pseudo-clefts and question-answer pairs. But, the parallel is different from the one drawn in the traditional approach. Traditionally, pseudo-clefts are conceptualized as containing a question and an answer, and they assert that the answer is the strongest true answer to the question. The entire question-answer pair is wrapped up into the pseudo-cleft. In my proposal, the pseudo-cleft is not conceptualized as a full question-answer pair. Rather, it is parallel to an exhaustive answer, interpreted as such an answer would be in the context of a particular question.

6.2 Adding adverbs

Because pseudo-clefts are semantically parallel to answers, the adverb data “follow for free” from the results of Section 5. I will not provide the full compositions, since they are rather transparent at this point, but I will provide the LFs. The pseudo-cleft with possibly has the LF in (173), which asserts that
Obama approved b1 and possibly b2 (the prejacent), but not b3 (φ3 is negated). As discussed at length in Section 5.2, φ1 is not negated, since ¬φ1 is inconsistent with the prejacent, and φ2 is not negated, since B(¬φ2) is inconsistent with the prejacent.

(173) Proposed LF for (70) (with possibly)

```
TP
  /    \
Exhp   CP
  |      |  what Obama approved
  |      |  Obama ⊢ [t1 approved this bill] and [possibly [t1 approved that bill]]
```

(174) Corresponding meaning

\[ [[(70)]]^C = \phi_1 \land \Diamond \phi_2 \land \neg \phi_3 \]

The pseudo-cleft with the conditional has the LF in (175), which asserts that Obama will approve b1 and, if Biden let him (which is a possibility), b2 (the prejacent), but not b3. φ3 gets negated, but φ1 does not (since ¬φ1 is inconsistent with the prejacent), nor does φ2 (since B(¬φ2) is inconsistent with the prejacent, given an existence presupposition in the modal), as detailed in Section 5.3.

(175) Proposed LF for (79) (with conditional)

```
TP
  /    \
Exhp   CP
  |      |  what Obama approved
  |      |  O ⊢ [t1 approved this bill] and [[[if Biden lets him] □] [t1 approved that bill]]
```

(176) Corresponding meaning

A: \( \phi_1 \land \phi_{\text{Biden}} \rightarrow \phi_2 \land \neg \phi_3 \)

P: \( \Diamond \phi_{\text{Biden}} \)

Finally, the LF containing with difficulty is (177), which that Obama approved b1 and, with difficulty, b2 (the prejacent), but not b3. Once again, φ3 gets negated, but neither φ1 nor φ2 does, since ¬φ1 and ¬φ2 are both inconsistent with the prejacent. This derivation mirrors the one in Section 5.4.
(177) Proposed LF for (24) (with *with difficulty*)

```
TP
  └── Exhep
  │    └── CP
  │         └── what Obama approved
```

(178) Corresponding meaning

\[
[[\text{(24)}]]_C = \phi_1 \land \phi_2 \land \text{difficult-}\phi_2 \land \neg \phi_3
\]

To re-iterate, nothing new has happened in these derivations. They are exactly like the ones from Section 5, except that the ALT argument for Exhep comes from the pre-copular question, rather than covert C anaphoric to the same extra-sentential question.

6.3 Taking stock

The core data are now captured. Section 2 provided novel evidence that the post-copular constituent is a TP, on the basis of adverb data. The proposal has fit that TP into a full LF which predicts correct, contingent meanings for a range of adverb data. The proposal is based on Exhep, which was independently motivated for exhaustive answers. Note that the crucial properties of exhaustive answers with Exhep were mirrored with overt *only*, so the analysis with Exhep did not involve special stipulations unique to a covert operator either.

7 A further puzzle

This section expands the range of data under consideration to probe further into the the over-negation problem. Although certain data will initially appear troubling for Exhep, I believe that the system as it stands can in fact handle them.

7.1 Examples with *believe*

To begin, the example in (179) appears consistent with the proposal. *I believe* occurs in the second conjunct and the paraphrase, given in (180), indicates that \(\phi_2\) does not get negated.

(179) Replacing *possibly* with *I believe*

What Obama approved was this bill and I believe that bill.

(180) Paraphrasing (179)

“Obama approved b1 and I believe b2, but no bill other b1 or b2.”
The most obvious syntax for the prejacent is (181-a), where \textit{I believe} embeds a TP expressing \( \phi_2 \). Given this analysis, the prejacent expresses the proposition in (181-b):

(181) \textbf{Prejacent in (179)}

\begin{enumerate}
  \item \( [TP_3 \text{ Obama approved this bill} \quad \text{and} \quad [TP_2 \text{ I believe } [TP_1 \text{ Obama approved that bill}]]] \)
  \item \( [TP_1] = \lambda w . \phi_1(w) \wedge B_w(\phi_2) \)
\end{enumerate}

If the main operator in (179) were \textit{Exh}, this example would exhibit over-negation: the prejacent does not entail \( \phi_2 \), so \textit{Exh} would erroneously negate \( \phi_2 \), creating Moore’s paradox. Just like with (70), the move to \textit{Exhep} solves the problem: \( B(\neg \phi_2) \) directly contradicts \( B(\phi_2) \) in the prejacent, so negation of \( \phi_2 \) is blocked. So far so good. Kai von Fintel (p.c.) points out, however, that a puzzle arises if the example is changed in a minimal respect: by replacing \textit{I believe} with \textit{Mary believes}. The resultant sentence is ambiguous between two readings.

(182) \textbf{Replacing I believe with Mary believes}

What Obama approved was this bill and Mary believes that bill.

On one reading, (182) conveys \( \neg \phi_2 \) and establishes that Mary mistakenly believes \( \phi_2 \). If interpreted this way, the sentence would paraphrase as (183). This reading is not very salient in (182), but it can readily be brought out by replacing \textit{and} with \textit{but}, as in (184).

(183) \textbf{Reading 1: paraphrase}

“O. approved b1 and no other; Mary wrongly believes he approved b2.”

(184) \textbf{Bringing out Reading 1}

What Obama approved was this bill, but Mary believes that bill.

Reading 1 is straightforwardly derived. There is, for instance, a viable parse where the conjunction takes wide scope above the pseudo-cleft. The left conjunct is the pseudo-cleft \textit{what Obama approved was this bill}, while the right conjunct is \textit{Mary believes (Obama approved) that bill}, as in (185). The pseudo-cleft asserts that Obama approved this bill and no other, and the right conjunct separately asserts Mary’s belief. Given exhaustivity in the pseudo-cleft, that belief must be mistaken.

(185) \textbf{LF to derive Reading 1}

More interesting for our purposes is the second reading, which is the dominant reading in (182) itself. On this reading, (182) does \textit{not} convey \( \neg \phi_2 \) and thus does not condemn Mary’s belief to be mistaken.
Rather, the speaker reports that \( \phi_1 \) is true, and puts forward the notion that \( \phi_2 \) is true on the basis of Mary’s belief. The exhaustivity entailment is then that Obama did not approve any bill other than b1 or b2. This reading paraphrases as (186).

(186) **Reading 2: paraphrase**

“Obama approved b1 and I suggest b2 on the basis of Mary’s belief, but certainly no bill other than b1 or b2.”

The best candidate LF to derive this reading is (187), where the conjunction scopes back within the post-copular constituent. However, a new layer of the over-negation problem now arises: even with the main operator Exhep, \( \phi_2 \) is predicted to get negated.

(187) **Apparent LF for Reading 2**

\[
\begin{align*}
\text{Exhep} &\quad \text{what Obama approved} \\
\text{TP}_2 &\quad \text{Obama approved this bill and} \\
\text{TP}_1 &\quad \text{Mary believes Obama approved that bill}
\end{align*}
\]

(188) **Prejacent in (188)**

\[
[\text{TP}_1]_C = \lambda w . \phi_1(w) \land \forall w' \in \delta(Mary)(w) [\phi_2(w')]
\]

To assess the possibility of negating \( \phi_2 \), we run our two tests. Since Mary believing \( \phi_2 \) does not entail \( \phi_2 \), the prejacent is compatible with \( \phi_2 \), and the first test passes. The second test looks for consistency between the prejacent and \( B(-\phi_2) \), where \( B(-\phi_2) \) is, as we discussed, an abbreviation for the speaker believing \( -\phi_2 \). Since Mary believing \( \phi_2 \) (in the prejacent) is compatible with the speaker believing \( -\phi_2 \), \( B(-\phi_2) \) is consistent with the prejacent and the second test passes, too. Hence, as far as Exhep is concerned, \( \phi_2 \) is excludable and should be erroneously negated.

Before the reader succumbs to new skepticism about the core proposal, I re-iterate again that the problem arises independent of pseudo-clefts in question-answer pairs. In the question-answer pair in (189), the answer may be interpreted in a way precisely parallel to Reading 2 of the pseudo-cleft (of course, in addition to allowing Reading 1, as well).

(189) **Baseline question-answer pair for (182)**

a. What did Obama approve?

b. This bill and Mary believes that bill.

Despite not challenging the core proposal, the problem we have reached does require a solution, and I suggest one in the following. If the solution I propose is correct, this new layer of the over-negation problem is really much less pernicious than it might appear at this point.
One natural first idea might be that we should perform additional surgery on Exh, modifying it further from Exh_{ep} to avoid negating \( \phi_2 \) in the Mary believes example. This is, however, difficult to achieve. The intuition about why \( \phi_2 \) is not negated is this: by mentioning that Mary believes \( \phi_2 \) to be epistemically possible, the speaker seems to indicate that they themselves take \( \phi_2 \) to be epistemically possible, based on Mary’s belief. This extra inference conflicts with \( B(\neg \phi_2) \) and that prevents \( \phi_2 \) from getting negated. The problem is that the bridge between an assertion about Mary’s belief and an extra inference about the speaker’s belief seems as though it would have to be purely pragmatic. As mentioned briefly earlier (and as I will address analytically in the next section), pseudo-clefts presuppose that the pre-copular question is already salient. If the pseudo-cleft is a response to the question What did Obama approve?, we might infer that Mary’s belief about \( \phi_2 \) is only relevant if the speaker endorses that belief, at least as possible. However, since Exh_{ep} is a syntactic operator, its interpretation should not be affected by pragmatic inferences, but rather only logical properties. Deriving a pragmatic inference involves reasoning about the pseudo-cleft’s semantic interpretation, and deriving the semantic interpretation itself involves composing Exh_{ep}. By their nature, pragmatic inferences come “too late” to be visible by Exh_{ep} — and it is thus not feasible to re-define Exh_{ep} not to negate \( \phi_2 \) based on such an inference.

Although changing the analysis of the prejacent seemed like a lost cause with possibly, this seems a very promising direction for Mary believes. The role for Mary believes assumed thus far has Mary believes take the TP expressing \( \phi_2 \) as its complement and compose with \( \phi_2 \) to derive a modalized meaning that does not entail \( \phi_2 \). However, there likely is another parse: one where Mary believes is instead analyzed as parenthetical.

To muster initial support for this idea, let us consider the distribution of \( x \) believes outside of pseudo-clefts. Like a prototypical parenthetical, \( x \) believes can appear at different positions in the sentence, as illustrated first with I believe in (190). When I believe occurs in non-initial positions, it is prosodically offset, again like any parenthetical.

(190) Parenthetical I believe

a. I believe Obama approved this bill in the senate.

b. Obama, I believe, approved this bill in the senate.

c. Obama approved this bill, I believe, in the senate.

d. Obama approved this bill in the senate, I believe.

Like I believe, Mary believes can also show the distribution of a parenthetical and, in that case, the intuition is very much like Reading 2 of the pseudo-cleft. Suppose someone asks the question What did Obama do? and the interlocutor responds with one of the sentences in (191):

(191) Parenthetical Mary believes

What did Obama approve?

a. Mary believes Obama approved this bill.

b. Obama approved this bill, Mary believes.

These responses are natural if the speaker does not have personal knowledge of Obama’s activities, but has learned of his activities from Mary, who he takes to be reliable. The speaker seems to put forward the proposition that Obama approved this bill, based on Mary’s report.
We are now ready to return to the pseudo-cleft in (182). Because the post-copular TP is elided, it is impossible to directly replicate the paradigm in (191) to demonstrate that *Mary believes* can be a parenthetical in this environment. But, if a parenthetical analysis is available in general, it should be available in the pseudo-cleft: the post-copular constituent is, after all, just an ordinary TP. Supportive of this conclusion, prototypical parentheticals can occur post-copular, as in (192):

(192) **Post-copular as-parenthetical**
What Obama approved was this bill and, as Mary said, that bill.

On this basis, I suggest that (182) has a parse with *Mary believes* as parenthetical, and this parse yields Reading 2. As mentioned earlier, Potts (2003) proposes a multi-dimensional semantic framework whereby parentheticals contribute secondary entailments on a separate dimension from the primary asserted content. Given this approach, the contribution of *Mary believes* is secondary, and the primary assertion of the prejacent is simply (193), with no modalization at all.

(193) **Primary assertion of prejacent**
\[ \phi_1 \land \phi_2 \]

I assume that Exh<sub>ep</sub> is exclusively concerned with the primary assertion of its prejacent. This fits Exh<sub>ep</sub> in with a wide range of operators. The conditional in (194-a) and the negation in (194-b), for instance, do not interact with the as-parenthetical. (193-a) carries a secondary entailment that Mary said he approved that bill, and (193-b) has a reading on which it does, as well.

(194) **Operators transparent to secondary entailments**
   a. If Obama approved this bill, as Mary reported, I will be upset.
   b. Obama didn’t approved this bill, as Mary reported.

With these pieces in place, the over-negation problem in (182) immediately resolves. Since (194) entails \( \phi_2 \), \( \neg \phi_2 \) is itself incompatible with the primary assertion of the prejacent, and \( \neg \phi \) does not pass the first test for excludability. \( \neg \phi_2 \) is correctly predicted not to get negated.

### 7.2 Examples with *say*

The problem raised with *Mary believes* in the preceding subsection replicates with a range of different expressions, and the solution replicates, as well. Consider, for instance, *Mary said:*

(195) **Mary said in the second conjunct**
What Obama approved was this bill and Mary said that bill.

If *Mary said* contributed to the primary assertion and embedded a TP expressing \( \phi_2 \), as an (196-a), an over-negation problem would arise. The proposition expressed is (196-b), with a universal modal over worlds compatible with the content of Mary’s speech report (‘S(m)(w)’). This prejacent is consistent with both \( \neg \phi_2 \) and B(\( \neg \phi_2 \)), so negation of \( \phi_2 \) would take place.
The prejacent in (195)

\[ \begin{align*}
\text{a. } & \quad [[TP \Diamond \text{ approved this bill } \land [TP [TP \text{ Mary said } [TP \Diamond \text{ approved that bill}]]]]] \\
\text{b. } & \quad \models (196-a) = \lambda w. \phi_1 \land \forall w' \in S(m)(w) [\phi_2(w')]
\end{align*} \]

Yet, intuitively, (195) has a reading analogous to Reading 2 of (182), by which \( \phi_2 \) is not negated: “Obama approved this bill and I put forward that he approved that bill, based on Mary’s report”. As with Mary believes, I suggest that (195) has a parse where Mary said is parenthetical and thus not relevant to the primary assertion. The primary assertion, (197), is then inconsistent with \( \neg \phi_2 \).

Proposal: primary assertion of prejacent

\[ \phi_1 \land \phi_2 \]

In sum, it is clear what Mary believes and Mary says have in common with respect to the overall proposal: unlike possibly and I believe their quantify over something other than worlds epistemically accessible to the speaker: Mary believes quantifies over worlds epistemically accessible to Mary, and Mary says quantifies over worlds compatible with the content of her speech report. Because \( \text{Exh}_{ep} \) checks whether the speaker believing the negation of an alternative is compatible with the prejacent, possibly and I believe can prevent an alternative from being negated, but Mary believes and Mary says cannot. To prevent \( \phi_2 \) from being negated in data with Mary believes \( \phi_2 \) and Mary says \( \phi_2 \), there must be an analysis where \( \phi_2 \) itself is asserted, as in the parenthetical analysis I have suggested. By extension, such an analysis must be available for any instance of the over-negation problem involving an intensional operator not quantifying over the speaker’s belief-worlds.

8 Conclusion

In the first part of the chapter, I provided novel empirical arguments that the post-copular constituent in a pseudo-cleft is a full TP and, by doing so, brought apparent DP conjunction in pseudo-clefts into line with the SIH: within the post-copular TP, there are type t nodes for and to conjoin. In the second part of the chapter, I built a complete LF around the TP. The key idea was that the compositional “glue” allowing the pre- and post-copular constituents to compose with one another is a covert \( \text{Exh}_{ep} \). The primary empirical testing ground were data with adverbials within the post-copular constituent.

A Pursuing the identity semantics

In response to the composition problem posed by the adverb data, I abandoned the Identity Hypothesis altogether. In this appendix, I ask whether there are moves we could have made to analyze adverbs and still maintain the Identity Hypothesis. I will consider just the example with possibly:

Recall: pseudo-cleft with possibly

What Obama approved was this bill and possibly that bill.

Recall the problem that the identity semantics faced. The identity semantics equates the strongest true proposition in the Hamblin set for the pre-copular question with the proposition expressed by the post-
copular TP. In (197), the Hamblin set contains propositions of the form *that Obama approved X*, where X is some entity. The proposition expressed by the post-copular TP is not, however, of that form, due to the modalization contributed by *possibly*. As a result, the pseudo-cleft predicts contradictory truth-conditions. The relevant semantic ingredients are repeated.

(198)  **Hamblin denotation for pre-copular question**

a. \[ [\text{what Obama approved}]^w \]

b. \[ = \lambda_Y \exists X [p = \lambda w' . \forall x <_\text{AT} X \text{[Obama approved x in } w']] \]

(199)  **Intension of post-copular TP**

\[ [[\text{TP}]_Q^C = \phi_1 \land \Diamond \phi_2 \]

(200)  **Identity semantics for pseudo-cleft**

\[ [[(70)]^w = 1 \iff [\text{ANS}]^w([\text{what Obama approved}]^w) = \phi_1 \land \Diamond \phi_2 \]

To have any chance at resurrecting an identity semantics, one of the other ingredients must change, either \([\text{what Obama approved}]^w\) or \([\text{TP}]_Q^C\). I concluded earlier that *possibly* should be interpreted as a modal with truth-conditional import. As such, the only option is to try to change the denotation for the question so that its Hamblin set includes modalized propositions. In the following, I consider where that tactic would lead. First, I modify the question semantics to achieve the desideratum that the Hamblin set contains \([\text{TP}]_Q^C\). I show that doing so re-creates a problem reminiscent of the over-negation problem which my proposal with Exhep faced. Unlike with Exhep, however, the problem does not link to independent data, and the potential solutions I can envision are either stipulative, or empirically inadequate for other data. The proposal with Exhep is, I believe, better motivated.

**A.1 Re-analyzing the question**

Suppose we re-analyze the pre-copular question so that the variable in its denotation does not range over entities X, but rather over quantifiers Q. The new Hamblin set will be a set of propositions *that Obama approved Q*, for different quantifiers Q. A parse along these lines was proposed for independent reasons in Spector (2008). Earlier in the chapter, we entertained two compositional means for deriving the original Hamblin denotation, differing in their syntax: one took the XP to be an interrogative CP, and the other a concealed question DP. To streamline discussion, I will adopt here just the interrogative CP analysis. Before, *what* encoded an existential quantifier of type \(<et,t>\). Now, I take *what* to denote a higher-order quantifier over basic quantifier meanings:10

(201)  **What ranging over quantifiers**

\[ [[\text{what}]'] = \lambda Y F_{<\text{et},t>,st} . \lambda w . \exists Q [F(Q)(w)] \]

The pre-copular CP has the LF in (202). *What* moves in two steps, first to a position below C, and then to spec-CP. Both steps leave a trace, but the traces have different types. The lower trace \((t_1)\) is of type e, as is required to furnish an argument for *approve*. The higher trace \((t_2)\) is of type \(<est, st>\).

10It will facilitate exposition here to move to intensional types for the rest of the appendix.
step of movement triggers abstraction over quantifiers and the result is the Hamblin denotation in (203): a set of propositions of the form that Obama approved Q.

(202) **New LF for question**

\[
\begin{align*}
\lambda_3 & \quad \text{CP} \\
\lambda_2 & \quad \text{what} \\
\lambda & \quad \text{CP} \\
C & \quad t_3 \\
\text{TP} & \quad t_2 \\
\lambda_1 & \quad \text{TP} \\
\text{Obama approved } t_1
\end{align*}
\]

(203) **Alternative Hamblin denotation**

a. \[\text{[what Obama approved']}\]
b. \[= \lambda p_{st} . \exists Q \{ p = \lambda w . Q(\lambda x . \lambda w' . O \text{ approved } x \text{ in } w')(w)\}\]

Now, what propositions does the new [what Obama approved'] contain? Let us assume that the salient bills are b1 and b2. First, it contains the same propositions as the original Hamblin set: \{\phi_1, \phi_2, \phi_1 \land \phi_2\}. To illustrate with \phi_1, (204) is a quantifier corresponding to the Montague Lift of b1. Filling in that quantifier for Q in (203) yields the proposition that Obama approved b1, \phi_1.

(204) **Quantifier meaning**

\[\lambda f_{est} . \lambda w . f(b1)(w)\]

(205) **Corresponding element of Hamblin set**

a. \[\lambda w . [\lambda f_{est} . \lambda w' . f(b1)(w')](\lambda x . \lambda w'' . \text{Obama approved } x \text{ in } w'')(w)\]
b. \[= \lambda w . \text{Obama approved } b1 \text{ in } w\]

The conjunction, \phi_1 \land \phi_2, is in the new Hamblin set via the quantifier in (206). According to the SIH, this quantifier is not expressible in natural language, at least not with a conjunction of DPs. Yet, it is still a meaning of type <est,st> and thus a possible value for Q. Filling in that quantifier for Q yields \phi_1 \land \phi_2. By extension, every proposition of the form that Obama approved X is equivalent to some proposition of the form that Obama approved Q.

(206) **Quantifier meaning**

\[\lambda f_{est} . \lambda w . f(b1)(w) \land f(b2)(w)\]

(207) **Corresponding element of Hamblin set**

a. \[\lambda w . [\lambda f_{est} . \lambda w' . f(b1)(w') \land f(b2)(w')](\lambda x . \lambda w'' . O \text{ approved } x \text{ in } w'')(w)\]
b. \[= \lambda w . \text{Obama approved } b1 \text{ in } w \land \text{Obama approved } b2 \text{ in } w\]
Now, we reach the important point: that the new Hamblin set is not equivalent to the original set, but rather is a proper subset of it. There are many propositions of the form that Obama approved $Q$ which do not have an equivalent with $X$. Crucially, this includes propositions that are epistemically modalized. The meaning in (208) is modalized and of type $<\text{est},\text{st}>$, and filling it in for $Q$ yields the proposition that Obama possibly approved $b_1$.

(208) Quantifier meaning
\[ \lambda f_{est}. \lambda w. \exists w' \in \delta'(s)(w) [f(b1)(w')] \]

(209) Corresponding element of Hamblin set
a. $\lambda w. [\lambda f. \lambda w'. \exists w'' \in \delta'(s)(w') [f(b1)(w'')]](\lambda x. \lambda w'' . O \text{ approved } x \text{ in } w''')(w)$
b. $= \lambda w. \exists w'' \in \delta'(s)(w) [\text{Obama approved } b_1 \text{ in } w'']$

As we have seen, the Hamblin set contains modalized, as well as non-modalized, propositions, and contains conjunctions. Putting these pieces together, the post-copular proposition in the pseudo-cleft in (70) is an element of the Hamblin set. The relevant quantifier is (210), and filling that quantifier in for $Q$ yields the proposition that Obama approved $b_1$ and possibly approved $b_2$, i.e. $\llbracket TP \rrbracket$.

(210) Quantifier meaning
\[ \lambda f_{est}. \lambda w. f(b1)(w) \land \exists w' \in \delta'(s)(w) [f(b2)(w')] \]

(211) Corresponding element of Hamblin set
a. $\lambda w. [(210)](\lambda x. \lambda w'' . \text{Obama approved } x \text{ in } w'')(w)$
b. $= \lambda w. \text{Obama approved } b_1 \text{ in } w \land \exists w'' \in \delta'(s)(w) [\text{Obama approved } b_1 \text{ in } w'']$

The are, of course, a great many propositions of the form that Obama approved $Q$ on top of the ones given here. These would include, for instance, negative proposition such as $\neg \phi$. The meaning in (212) is a quantifier, and filling it in for $Q$ delivers the proposition that Obama didn’t approve $b_1$. Spector (2008) assumed that $Q$ is restricted to upward entailing quantifiers, and I will make that assumption here, too. Negative quantifiers such as (211) are downward entailing.

(212) Quantifier meaning
\[ \lambda f_{est}. \lambda w. \neg f(b1)(w) \]

(213) Corresponding element of Hamblin set
a. $\lambda w. [\lambda f_{est}. \lambda w. \neg f(b1)(w)](\lambda x. \lambda w'' . \text{Obama approved } x \text{ in } w'')(w)$
b. $= \lambda w. \text{Obama didn’t approve } b_1 \text{ in } w$

To circumscribe discussion, I will assume that the Hamblin set contains just the non-modalized propositions in (214-a), the modalized propositions in (214-b), and all conjunctions thereof. Note that the modals in (214-b) are possibly $p$ and certainly $p$, and I take certainly $p$ to be veridical, as in (215). The veridicality assumption which is not crucial, but will simplify exposition.
Elements of the Hamblin set

- $\phi_1, \phi_2$
- $\diamond \phi_1, \diamond \phi_2, \Box \phi_1, \Box \phi_2$

Defining $\Box$

$\Box \phi \iff \lambda w \cdot \phi(w) \land \forall w' \in \delta(s)(w) [\phi(w')]$

Since $\llbracket TP \rrbracket$, under current assumptions, is a conjunction with both non-modalized and modalized conjuncts, the Hamblin set I consider is the most restricted it can be and still contain $\llbracket TP \rrbracket$. We will see that narrowing the Hamblin set as much as possible gives an identity semantics for the pseudo-cleft its best shot at predicting a sensible meaning. Now that $\llbracket TP \rrbracket$ is a formal answer to the question, let us consider: does an identity semantics, predict the correct meaning for the pseudo-cleft?

### A.2 Applying to the pseudo-cleft

The identity semantics is repeated in (216). ANS picks out the strongest true element of the Hamblin set and, to assess the reasonableness of the meaning, we must evaluate the conditions under which $\llbracket TP \rrbracket$ would qualify as this strongest true element.

Truth-conditions for pseudo-cleft

$\llbracket (70) \rrbracket(w) = 1 \text{ iff } \text{ANS}(\llbracket \text{what Obama approved'} \rrbracket)(w) = \phi_1 \land \diamond \phi_2$

Because the Hamblin set contains non-modalized propositions, modalized propositions, and their conjunctions, its strongest true element will say which bills Obama approved, and which bills the speaker believes it is possible or certain that Obama approved. Consider, for instance, a world $w_1$. At $w_1$, Obama approved $b_1$, but not $b_2$. Moreover, the speaker believes it is possible that he approved $b_1$, and is certain he did not approve $b_2$. At $w_1$, the strongest true element is:

Applying ANS at $w_1$

- $\text{ANS}(\llbracket \text{what Obama approved'} \rrbracket)(w_1) = \phi_1 \land \diamond \phi_2$

Now, consider again $\llbracket TP \rrbracket$. $\llbracket TP \rrbracket$ asserts $\phi_1$, but it does not assert that the speaker believes $\phi_1$ to be possible or certain. Conversely, $\llbracket TP \rrbracket$ asserts $\diamond \phi_2$, but it does not assert $\phi_2$ itself. For $\llbracket TP \rrbracket$ to be the strongest true proposition in the Hamblin set, therefore, the following must hold. First, $\phi_1$ must be true, but the speaker must believe that $\phi_1$ is impossible. And, second, the speaker must believe that $\phi_2$ is possible, but $\phi_2$ must be false. By asserting that $\llbracket TP \rrbracket$ is the strongest true proposition, the pseudo-cleft thus introduces two paradoxical entailment patterns. I present each in detail.

### A.3 Paradox 1

To see the first paradox, compare $\llbracket TP \rrbracket$ to the proposition in (218), which is also in the Hamblin set. Because (218) adds to $\llbracket TP \rrbracket$ the additional modalized conjunct, $\diamond \phi_1$, (218) is stronger than $\llbracket TP \rrbracket$. For $\llbracket TP \rrbracket$ to be the strongest true proposition in the Hamblin set, therefore, (218) must be false — and the only way for (218) to be false if $\llbracket TP \rrbracket$ is true is for $\diamond \phi_1$ to be false.
Proposition stronger than [TP]
\[ \Diamond \phi_1 \land \phi_1 \land \Diamond \phi_2 \]

In this way, we arrive at a paradoxical element pattern: the pseudo-cleft should entail both that Obama approved b1 (because [TP] is true), and that Obama approving b1 is epistemically impossible (because (218) is false). The prediction is paradoxical ('\(\phi_1 \) but I believe \(\phi_1\) is impossible').

This first paradox is not fatal, but rather can be avoided quite easily by strengthening the asserted content of the post-copular TP. We have assumed that the first conjunct is not modalized: the TP asserts \(\phi_1\), but does not assert the speaker’s beliefs about \(\phi_1\). By uttering the pseudo-cleft, the speaker does actually convey that they believe \(\phi_1\). This could be a pragmatic inference, but it also possible that the first conjunct asserts both \(\phi_1\) and that the speaker believes \(\phi_1\). The latter can be achieved by parsing the first conjunct with a covert universal modal:

Alternative LF for post-copular TP

Because \(\Box\) is veridical, the TP now expresses the proposition that Obama approved b1, that the speaker believes so, and that Obama possibly approved b2. I refer to this proposition as [TP']. The identity semantics for the pseudo-cleft revises to (221), equating the strongest true proposition in the Hamblin set with [TP'].

Corresponding meaning for TP

[TP'] = \(\Box \phi_1 \land \Diamond \phi_2\)

Revised truth-conditions for pseudo-cleft

\[ ((70))(w) = 1 \text{ iff ANS(‘what Obama approved’) } = \phi_1 \land \Diamond \phi_2 \]

The pseudo-cleft cleft no longer entails that the proposition in (218) is false. In fact, it entails (218) itself. Whereas (218) was stronger than [TP], it is weaker than [TP']. \(\Box \phi_1\) in [TP'] entails \(\Diamond \phi_1\) in (218) and, with \(\Box\) veridical, it entails \(\phi_1\) also. Hence, by modifying the assertion of the post-copular TP in an intuitive way, the paradoxical entailment of \(\neg \Diamond \phi_1\) is avoided.

A.4 Paradox 2

The second paradox resembles the over-negation problem, and is the more difficult one to solve. With the post-copular TP analyzed as in the last subsection, the paradox arises by comparing [TP'] to the
proposition in (222), which is again in the Hamblin set. This proposition is stronger than [TP'], since it adds to [TP'] the non-modalized conjunct, \( \phi_2 \).

(222) \begin{equation}
\text{Proposition stronger than [TP']}
\Box \phi_1 \land \Diamond \phi_2 \land \phi_2
\end{equation}

For [TP'] to be the strongest true proposition in the Hamblin set, (223) must be false. If [TP'] is true, the only way for (223) to be false is for \( \phi_2 \) to be false. Hence, the second paradoxical entailment pattern: the pseudo-cleft is predicted to entail both \( \Diamond \phi_2 \) (since [TP'] is true) and \( \neg \phi_2 \) (since (222) is false). This is identical to the over-negation problem which arose for my proposal with Exh: the pseudo-cleft should convey ‘Obama possibly approved b2, but he didn’t actually approve b2’.

How might we resolve the paradox? Taking a cue from Paradox 1, we could try modifying the asserted content of the post-copular TP so that it is equivalent or stronger to (222). Then, the pseudo-cleft would not entail that (222) was false. This approach is, however, a non-starter. Not only does the pseudo-cleft not intuitively entail that (222) is false, but it does not entail that (222) is true either. The speaker commits to \( \phi_2 \) being possible, but not to \( \phi_2 \) being true.

Maintaining [TP'], the only option I see is to attempt to restrict the Hamblin set so that the proposition in (222) is not included in it. That way, [TP'] could be the strongest true element, independent of whether (222) were true or false. In the following, I sketch how this would work, but the required restriction does not seem well-motivated.

A.4.1 Option 1: contextually restricting the Hamblin set

In the revised meaning for the post-copular TP, both conjuncts are modalized. Accordingly, for [TP'] to be an element of the Hamblin set for the pre-copular question, it is not crucial to include non-modalized propositions in the set. As Irene Heim points out, Paradox 2 would be avoided if the Hamblin set were optionally contextually restricted to include just the modalized propositions (\( \Diamond \phi_1, \Diamond \phi_2, \Box \phi_1, \Box \phi_2 \)) and their conjunctions. The complete restricted Hamblin set is enumerated in (223). Whereas [TP'] is still in the Hamblin set, the proposition in (222) is not, because the final conjunct (\( \phi_2 \)) is non-modalized. Because (222) is not an element of the Hamblin set, the over-negation problem does not arise.

(223) \begin{equation}
\text{Restricting Hamblin set to modalized propositions}
\begin{align*}
a. \{ & \Diamond \phi_1, \Box \phi_1, \Diamond \phi_2, \Box \phi_2, \\
b. & \Diamond \phi_1 \land \Diamond \phi_2, \\
c. & \Diamond \phi_1 \land \Box \phi_2, \\
d. & \Box \phi_1 \land \Diamond \phi_2, \\
e. & \Box \phi_1 \land \Box \phi_2 \}
= [TP']
\end{align*}
\end{equation}

In general, the strongest true proposition in the restricted Hamblin set will say every bill that Obama certainly or possibly approved, but will not be forced to say anything about what bills Obama actually approved. Assuming that the epistemic accessibility relation is reflexive, asserting \( \Box \phi \) entails \( \phi \), so inferences about what Obama actually approved may be drawn from assertions about what he certainly approved. However, asserting \( \Diamond \phi \) will not give rise to any entailments about what Obama actually approved. This last point is crucial to resolving the over-negation problem.
Concretely, $[TP']$ asserts that Obama certainly approved $b_1$, and possibly approved $b_2$. The only proposition in the Hamblin set stronger than $[TP']$ is now (224), which asserts that Obama certainly approved $b_1$, and certainly approved $b_2$. For $[TP']$ to be the strongest true proposition in the Hamblin set, all that must hold is that the speaker is not certain that Obama approved $b_2$, falsifying the second conjunct in (224).

(224) **Proposition stronger than $[TP']$**

$$\square \phi_1 \land \Diamond \phi_2$$

By asserting that $[TP']$ is the strongest true proposition, therefore, the pseudo-cleft would entail $\square \phi_1$ and $\Diamond \phi_2$ (since $[TP']$ is true), and it would entail $\neg \square \phi_2$ (since (224) must be false). In prose, Obama certainly approved $b_1$, and possibly but not certainly approved $b_2$. That $\phi_1$ is true follows from $\square \phi_1$, but there is no entailment about the actual truth or falsity of $\phi_2$. The over-negation problem is avoided — and the predicted meaning is contingent and intuitively reasonable. To get a feel for the analysis, I will flag several particular worlds.

First, take a world, $w_1$, at which it is epistemically certain that Obama approved $b_1$, and epistemically possible that he approved $b_2$. At $w_1$, the true propositions in the Hamblin set are:

(225) **True propositions at $w_1$**

a. $\Diamond \phi_1, \square \phi_1, \Diamond \phi_2$

b. $\Diamond \phi_1 \land \Diamond \phi_2$

c. $\square \phi_1 \land \Diamond \phi_2$

Because (225-c) is stronger than any of the other true propositions, ANS will return (225-c) — and (225-c) is just $[TP]$. Hence, an identity semantics for the pseudo-cleft comes out true at $w_1$. Note that the prediction is independent of whether or not $\phi_2$ itself is true at $w_1$, since $\phi_2$ is not included in the Hamblin set.

(226) **ANS([what Obama approved])(w_1)**

a. $= \square \phi_1 \land \Diamond \phi_2$

b. $= [TP']$

In addition to predicting the pseudo-cleft to come out true under intuitively correct conditions, the pseudo-cleft also comes out false where it should. For one illustration, suppose a world $w_2$ which is like $w_1$ in that every epistemically accessible world is a $\phi_1$-world, but unlike $w_1$ in that there is no epistemically accessible $\phi_2$-world. In other words, at $w_2$, the speaker believes that $\phi_2$ is impossible. Then, the true propositions in the Hamblin set are just (227-a) and (227-b). Since (227-b) is stronger, ANS would return it, and the identity semantics would come out false, since (227-b) is not equivalent to $[TP]$. The prediction for $w_2$ is intuitively correct.
(227) True propositions at \( w_2 \)
   a. \( \Diamond \phi_1 \)
   b. \( \Box \phi_1 \)

(228) \( \text{ANS(\{what Obama approved\})(w_1)} \)
   a. \( = \Box \phi_1 \)
   b. \( \neq [\text{TP}'] \)

Moreover, the exhaustive entailment associated with the pseudo-cleft is captured. Suppose another bill, \( b_3 \), is salient in addition to \( b_1 \) and \( b_2 \). If \( b_3 \) is salient, (229) is among those propositions added to the original Hamblin set in (223) above. The exhaustive meaning of the pseudo-cleft conveys that (229) is false: Obama definitely approved \( b_1 \) and it's possible that he approved \( b_2 \) but it is not possible that he approved \( b_3 \).

(229) Another proposition in the Hamblin set
\( \Box \phi_1 \land \Diamond \phi_2 \land \Diamond \phi_3 \)

With (229) in the mix, let us return to the world \( w_1 \) above at which \( \phi_1 \) is epistemically certain and \( \phi_2 \) epistemically possible and assume, moreover, that \( \phi_3 \) is epistemically impossible. At \( w_1 \), (229) is false, and the true propositions in the Hamblin set are the same as in (225) above, repeated in (230), and \( \text{ANS} \) returns (230-c), equivalent to \([\text{TP}]\), just as before. When \([\text{TP}]\) is true and (229) false, the identity semantics comes out true, as it should.

(230) True propositions at \( w_1 \)
   a. \( \Diamond \phi_1, \Box \phi_1, \Diamond \phi_2 \)
   b. \( \Diamond \phi_1 \land \Diamond \phi_2 \)
   c. \( \Box \phi_1 \land \Diamond \phi_2 \)

(231) \( \text{ANS(\{what Obama approved\})(w_1)} \)
   a. \( = \Box \phi_1 \land \Diamond \phi_2 \)
   b. \( = [\text{TP}'] \)

Now, what happens with a world \( w_3 \) at which \( \phi_1 \) is epistemically certain and \( \phi_2 \) and \( \phi_3 \) are both epistemically possible? At \( w_3 \), (229) is true in addition to all the propositions in (230). Because (229) is stronger than (230-c), \( \text{ANS} \) returns it rather than (230-c). Because (229) is not equivalent to \([\text{TP}]\), the identity semantics rightly comes out false.

(232) \( \text{ANS(\{what Obama approved\})(w_3)} \)
   a. \( = \Box \phi_1 \land \Diamond \phi_2 \land \Diamond \phi_3 \)
   b. \( \neq [\text{TP}] \)

The cases we have discussed make transparent that the only way for \([\text{TP}]\) to be the strongest true answer to the pre-copular question is if Obama certainly approved \( b_1 \), possibly, but not certainly, approved \( b_2 \), and certainly did not approve \( b_3 \) — capturing reasonable truth-conditions.
Although the system presented here makes correct predictions, the restriction upon which it is based does not seem well-motivated: the contextual restriction to just modalized propositions. For this restriction to come about, it would have to be that the modalized propositions count as salient or relevant, while the non-modalized propositions do not. It is difficult to see, however, how $\Box \phi$ or $\Diamond \phi$ could be salient or relevant, if $\phi$ is not also salient or relevant, as well.

A.4.2 Option 2: restriction inside ANS

Before jettisoning the approach, however, we could entertain a different way of bringing about the restriction, which does not involve modifying the Hamblin set itself. For this purpose, let us return to the earlier Hamblin set, which included non-modalized and modalized propositions, and their conjunctions. The idea I entertain is to modify ANS to incorporate a restriction into its definition in a perhaps more principled way. The familiar definition is repeated in (233). ANS applies to a set of propositions (Q) and a world (w), and returns the strongest true proposition in Q at w.

(233) Original definition of ANS

\[
\begin{align*}
\llbracket \text{ANS} \rrbracket(Q)(w) \\
a. & \quad \text{A: } \top p \ [w \in p \in Q \land \forall p' \in Q \ [w \in p' \rightarrow p \subseteq p'] ] \\
b. & \quad \text{P: } \exists p \ [w \in p \in Q \land \forall p' \in Q \ [w \in p' \rightarrow p \subseteq p'] ]
\end{align*}
\]

I entertain modifying ANS so that it returns the strongest true proposition which is in Q and which the speaker believes to be true. To cache this out, I define in (234) an operator BS (for ‘belief state’), which is indexed to an individual x, and applies to a world w to return the set of propositions which are true according to x’s beliefs at w.

(234) Defining BS

\[ BS_x = \lambda w . \lambda p . x \text{ believes } p \text{ in } w \]

ANS may, then, be redefined as (235). In addition to Q and w, ANS takes as argument BS, which I take to be indexed to the speaker s in the relevant data. It is written into ANS that Q is intersected with BS_s(w), and ANS(\text{BS}_s)(Q)(w) thus returns the strongest proposition in Q that is also in \text{BS}_s(w). The difference between ANS and ANS’ can be brought out with informal paraphrases. On the one hand, we might paraphrase ANS(Q)(w) as: “the answer to the question Q at w is ...”. On the other hand, we might paraphrase ANS'(BS_s)(Q)(w) as “my best answer to the question Q at w is ...”.

(235) Revised definition of ANS

\[
\begin{align*}
\llbracket \text{ANS}' \rrbracket'(Q)(w) \\
a. & \quad \text{A: } \top p \ [w \in p \in \{Q \cap BS_s(w)\} \land \forall p' \in \{Q \cap BS_s(w)\} \ [w \in p' \rightarrow p \subseteq p'] ] \\
b. & \quad \text{P: } \exists p \ [w \in p \in \{Q \cap BS_s(w)\} \land \forall p' \in \{Q \cap BS_s(w)\} \ [w \in p' \rightarrow p \subseteq p'] ]
\end{align*}
\]

The meaning expressed by the pseudo-cleft updates to (236), with ANS’ in place of ANS. The pseudo-cleft asserts that \llbracket TP' \rrbracket is the strongest true proposition of the form that Obama approves Q which the speaker believes to be true — \llbracket TP' \rrbracket is my best answer to the pre-copular question.
Truth-conditions for pseudo-cleft

\[ [\{\psi\}](w) = \text{iff } \text{ANS}'([\text{what Obama approved}](w) = \Box \phi_1 \land \Diamond \phi_2) \]

With ANS re-defined, the over-negation problem evanesces, even if the Hamblin set contains non-modalized, as well as modalized, propositions. Recall that the over-negation problem stemmed from the proposition in (237), which is stronger than \[\text{TP}'\].

Proposition stronger than \[\text{TP}'\]

\[ \Box \phi_1 \land \Diamond \phi_2 \land \phi_2 \]

With the pseudo-cleft defined in terms of ANS, it was wrongly predicted to entail that (238) is false and, accordingly, that \( \phi_2 \) is false. If (237) were true, \([\text{TP}']\) could never be the strongest true answer to the pre-copular question. The re-definition of the pseudo-cleft bleeds the problematic entailment. Even if (237) were true, \([\text{TP}']\) still could be my best answer to the pre-copular question.

Concretely, take a world, \( w_1 \), at which both \([\text{TP}']\) and (237) are true. At \( w_1 \), it is epistemically certain that Obama approved \( b_1 \) and epistemically possible (but not certain) that he approved \( b_2 \), and at which Obama actually did approve \( b_2 \). To assess the truth-value of the pseudo-cleft, we must provide \( \text{BS}_s(w_1) \), which includes the propositions in (238), and all conjunctions thereof:

\[ \text{BS}_s(w_1) \approx \{ \phi_1, \Diamond \phi_1, \Box \phi_1, \Diamond \phi_2, \ldots \} \]

\([\text{TP}']\) is an element of \( \text{BS}_s(w_1) \) — but (237) is not. Because \( \phi_2 \) is not epistemically certain, \( \phi_2 \) is not an element of \( \text{BS}_s(w_1) \) and (237), which contains \( \phi_2 \) as a conjunct, is thus not an element either. Although (237) is the strongest true proposition in the Hamblin set at \( w_1 \), \([\text{TP}']\) is the strongest true proposition in the Hamblin set that is also in \( \text{BS}_s(w_1) \) — and that is now enough to make the pseudo-cleft true. Since the pseudo-cleft is compatible with (237), the unwanted entailment that \( \phi_2 \) is false does not arise, solving the over-negation problem.

To re-iterate, the solution presented here, while different in letter, is similar in spirit to the one in the previous section. In Option 1, the Hamblin set itself was restricted to contain just modalized propositions, eliminating the problematic proposition in (237) from the Hamblin set. In Option 2, an additional set is defined (the belief state) and the Hamblin set is intersected with that set. The problematic proposition is not necessarily contained in the intersection and 'strongest' in the new definition of ANS is evaluated just over the intersection.

Although the move to \( \text{ANS}' \) avoids the over-negation problem, \( \text{ANS}' \) does not in general seem to be a viable definition for that operator. ANS occurs not just in pseudo-clefts, but in all embedded questions, and perhaps indeed in all questions. In the embedded environment in (239), \( \text{ANS}' \) generates a faulty prediction.

Embedded question

I know what Obama approved.

If the embedded question is parsed with \( \text{ANS}' \), (239) should have the trivial meaning that I know my own best answer to the pseudo-cleft, i.e. I know exactly what I believe Obama approved. In actual fact, however, (240) seems to convey something stronger: it conveys that I know the actual strongest true
answer to the question. To derive the right meaning here, ANS is required, rather than ANS' and, on this basis, I do not believe the Option 2 solution is viable.

A.5 Local summary

This section laid out in detail what an identity semantic account of pseudo-clefts would have to say to reconcile itself with the possibly data point. For the post-copular proposition to be an element of the pre-copular question, the variable in the Hamblin set must be of quantifier type. That alone is not, however, enough to resolve the puzzle. Rather, it creates two paradoxes, the second of which resembles the over-negation problem my own account with Exh faced. I considered two solutions to the over-negation problem as manifest with an identity semantics and argued that neither is viable. Option 1 involved a contextual restriction on the Hamblin set, and Option 2 restricted the Hamblin set by modifying the definition of ANS in a way that was inappropriate in other environments. The Exh approach linked the over-negation problem to a more general puzzle with Exh and only independently requiring a solution. For this reason, I believe that the Exh approach is overall the more principled proposal.

B More on ‘ALT’ in question-answer pairs

The ‘ALT’ argument for Exhep in a question-answer pair is a covert pronoun, C. I assumed that C is anaphoric to the Hamblin set for the question in all question-answer pairs. In this appendix, I offer a more thorough discussion of how C is determined, and justify that choice. In addition, I show that the main results of Section 5 replicate even if the value for C were determined in a different way.

B.1 Assumptions about how C is determined

To start, let us consider the question-answer pair in (240), where the answer does not contain possibly. The LF that I presented in the text for (240-b) is (241).

(240) Recall: basic question-answer pair
   a. What did Obama approve?
   b. This bill.

(241) Recall: LF for (240-b)

Rooth (1992b) proposed that the make-up of C is constrained, with constraints introduced as presuppositions of a covert operator, ~. Taking ~ into account, the LF for (240-b) would revise to (242). Exhep and ~ each introduce an occurrence of C, with their respective C arguments co-indexed.
C introduces three presuppositions about the make-up of C. The first two are that C must contain the prejacent, and some element other than the prejacent. The third has to do with focus. To streamline the main text, I ignored an important feature of the answer in (240-b): that the constituent corresponding to what, i.e. this bill, is focused. The contribution of focus is to introduce alternatives. In Rooth's system, focus alternatives are computed by replacing the focused element with any meaning of the same semantic type. When this bill is focused, the focus alternatives to the TP are propositions just like that Obama approved this bill, but with this bill replaced by different entities:

(243) **Focus alternatives to TP**

a. \( \lambda w . \text{Obama approved} \; b_1 \; \text{in} \; w \),

b. \( \lambda w . \text{Obama approved} \; b_2 \; \text{in} \; w \),

c. \( \lambda w . \text{Obama approved} \; b_1 \; \text{and} \; b_2 \; \text{in} \; w \), ...

Rooth encodes focus alternatives on a separate dimension of meaning. Every node in the structure has an ordinary semantic value and a focus semantic value. The values for the TP are:
Multi-dimensional meaning for TP

a. \[\llbracket\text{TP}\rrbracket^o = \lambda w. \text{Obama approved b}_1 \text{ in w}\]

b. \[\llbracket\text{TP}\rrbracket^f = \{\lambda w. \text{Obama approved b}_1 \text{ in w},\]
\[\lambda w. \text{Obama approved b}_2 \text{ in w},\]
\[\lambda w. \text{Obama approved b}_3 \text{ in w}, \ldots\}\]

The third presupposition of \(\sim\) is that \(C\) must have as its ordinary value a subset of the focus value of the sister to \(\sim(C)\). For the case at hand, then, the three presuppositions \(\sim\) introduces are:

Presuppositions of \(\sim\)

a. \(\llbracket\text{TP}\rrbracket^o \in \llbracket C\rrbracket^o\)

b. \(\exists p \ [p \in \llbracket C\rrbracket^o \land p \neq \llbracket\text{TP}\rrbracket^o]\)

c. \(\llbracket C\rrbracket^o \subseteq \llbracket\text{TP}\rrbracket^f\)

In (240-b), all three conditions are met if \(C\) is the Hamblin set for the question, \(\{\phi_1, \phi_2, \phi_3, \ldots\}\). That set contains \(\llbracket\text{TP}\rrbracket^o\) (i.e. \(\phi_1\)), another element (e.g. \(\phi_2\)), and is a subset of the \(\llbracket\text{TP}\rrbracket^f\) (cf. (244)).

Now, let us consider one of the adverb examples. For illustration, take the possibly example, in particular. The question-answer pair is repeated in (246), and an LF is provided with \(\sim\) included.

Recall: basic question-answer pair

a. What did Obama approve?

b. This bill and possibly that bill.

LF with \(\sim\) for (246-b)

In Section 5.2, I took the value for \(C\) to again be the Hamblin set for the question. If so, are the three presuppositions of \(\sim\) met? The first is not, since the Hamblin set does not contain modalized propositions. The second clearly is. In Rooth's system, the third presupposition should be met, as well. To make this clear, we must determine the focus value of the TP. Since this bill, possibly, and that bill all survive ellipsis, I take it that each of them is focused. As a result, each of them is replaced with different meanings of the same semantic type to create focus alternatives. This bill and that bill are replaced by different entities. Possibly is of type \(<st,st>\) so should be replaced by other meanings of that type. One such meaning is the identity function which maps a proposition to itself \((\lambda p. p)\). As a result, the propositions in (248) are elements of the \(\llbracket\text{TP}\rrbracket^f\) — and they are equivalent to \(\phi_1, \phi_2,\) and \(\phi_3\) in the Hamblin set.
In (248-a), for instance, this bill is replaced with itself, that bill is replaced with this bill, and possibly is replaced with the identity function, bleaching it.

\begin{equation}
(248) \quad \text{Included in } [\text{TP}]^f
\end{equation}

\begin{itemize}
\item a. \( \lambda w . \text{Obama approved } b1 \text{ in } w \land \text{Obama approved } b1 \text{ in } w \) = \( \phi_1 \)
\item b. \( \lambda w . \text{Obama approved } b2 \text{ in } w \land \text{Obama approved } b2 \text{ in } w \) = \( \phi_2 \)
\item c. \( \lambda w . \text{Obama approved } b3 \text{ in } w \land \text{Obama approved } b3 \text{ in } w \) = \( \phi_3 \)
\end{itemize}

The first presupposition of C does not in general play a crucial role, and I suggest that we abandon it, so that C only introduces the two presupposition in (249). Those two remaining presuppositions are satisfied in (246-b) with \( [C]^o \) the Hamblin set for the question — explaining the choice made in the text to take that set as the value for C.\(^{11}\)

\begin{equation}
(249) \quad \text{Presuppositions of } \sim
\end{equation}

\begin{itemize}
\item a. \( \exists p [p \in [C]^o \land p \neq [\text{TP}]^o] \)
\item b. \( [C]^o \subseteq [\text{TP}]^f \)
\end{itemize}

**B.2 A different value for C**

A key point of discussion in Section 5 was the over-negation problem. Because the Hamblin set includes \( \phi_2 \), if the Hamblin set is the value for C, then Exh, as originally defined, should negate \( \phi_2 \), erroneously introducing the entailment \( \neg \phi_2 \). It is worth noting that the over-negation problem arises under different theories of how C is determined as well. A different take for C in (247) might look like the following. Although, in Rooth's system, focus alternatives are computed by replacing the focused element with any other meaning of the same semantic type, more constrained theories of alternative computation have been proposed. Horn (1972) proposed logical operators are on scales. Focus alternatives, then, may be computed by replacing the focused element just with its scale-mates. Suppose possibly were on a scale <necessarily, possibly>. Then, all of the elements of \([\text{TP}]^f\) would contain either possibly or necessarily, as illustrated with the sampling in (250):

\begin{equation}
(250) \quad \text{Including in } [\text{TP}]^f
\end{equation}

\begin{itemize}
\item a. \( \{ \lambda w . \text{Obama approved } b1 \text{ in } w \land \text{Obama possibly approved } b2 \text{ in } w, \)
\item b. \( \lambda w . \text{Obama approved } b1 \text{ in } w \land \text{Obama necessarily approved } b2 \text{ in } w, \)
\item c. \( \lambda w . \text{Obama approved } b2 \text{ in } w \land \text{Obama possibly approved } b2 \text{ in } w, \)
\item d. \( \lambda w . \text{Obama approved } b2 \text{ in } w \land \text{Obama necessarily approved } b2 \text{ in } w, \ldots \}
\end{itemize}

This set does include \([\text{TP}]^o\), but the Hamblin set for the question is not a subset of it. Thus, the C argument would have to be an accommodated set of propositions like those in (250). Importantly, that set should include (250-c). The prejacent in (247) does not entail (250-c), so the original Exh would negate that alternative. Given the prejacent, negating (250-c) would introduce the entailment \( \neg \phi_2 \), re-creating the over-negation problem. Hence, regardless of whether \([C]^f\) is the Hamblin set for the question, or an accommodated set such as (250), the over-negation problem arises.

\(^{11}\)Note that pseudo-clefts may similarly require jettisoning the first presupposition of \( \sim \). The post-copular constituent in a pseudo-cleft displays a similar focus structure to the answer to a question, and \( \sim \) should be in the pseudo-cleft LFs, as well.
Chapter 5
Coordinating questions

1 Introduction

This chapter is concerned with coordination of questions, both conjunction and disjunction. It is well known that questions can be conjoined (e.g. Krifka 1999), and it has controversially been suggested that they can be disjoined, as well (Groenendijk & Stokhof 1989, Ciardelli et al. 2015, pace Szabolcsi 1999, Krifka 1999, Haida & Repp 2013). I will report new data which converge with Ciardelli et al.'s (2015) recent conclusion that disjunction is attested. First, I use conjunction and disjunction — both analyzed in accord with the SIH — as a vehicle to probe into the syntax and semantics of questions. I propose that questions are constructed in such a way that their LF contains type t nodes, at which logical and or can scope without type-lifting. Second, armed with a basic LF, I undertake a close study of the disjunction data, and propose a detailed analysis of this controversial construction. In this way, this chapter has a similar make-up to the previous chapter on pseudo-clefts: it pursues the SIH, and situates the results within a full analysis of the relevant construction. In the remainder of the introduction, I sketch the key ideas of the chapter.

1.1 Part I: Pursuing the SIH

From the perspective of the SIH, question coordination looks deeply puzzling. The SIH requires that and only be interpreted as \([\text{and}]\), which operates on truth-values. With respect to disjunction, the SIH should make a similar claim: that or be interpreted as \([\text{or}]\), which encodes the \(\lor\) connective of propositional logical and, as such, again operates on truth-values.

\[
(1) \quad \text{Recall } [\text{and}] \\
[\text{and}] = \lambda p_r . \lambda q_r . p \land q
\]

\[
(2) \quad \text{Defining } [\text{or}] \\
[\text{or}] = \lambda p_r . \lambda q_r . p \lor q
\]

When declarative clauses are coordinated, it is clear that the LF contains truth-value denoting nodes. The example in (3), for instance, is straightforwardly interpreted as (4), as we have already seen in earlier chapters. Each TP has a truth-value, and \([\text{and}]\) can operate on those truth-values.

\[
(3) \quad \text{Conjunction of declarative sentences} \\
[TP_1 \text{ Mary is the mother}] \text{ and } [TP_2 \text{ John is the father}].
\]

\[
(4) \quad \text{Meaning predicted with } [\text{and}] \\
a. \quad [\text{and}](\langle TP_1 \rangle^w)(\langle TP_2 \rangle^w) \\
b. \quad = 1 \text{ iff Mary is the mother in } w \land \text{ John is the father in } w
\]

The situation with questions, however, looks very different. Questions do not obviously provide a true or false description of the world, but rather solicit such a description. One way to model the meaning of a question, already discussed in Chapter 4, is as a set of propositions. (5-a) is modeled as (the characteristic function for) a set of propositions of the form that \(x\) is the mother, as in (5-b).

179
(5) **Hamblin denotation (b) for simple question (a)**

   a. Who is the mother?
   b. \( \lambda x \exists p [p = \lambda w . x \text{ is the mother in } w] \) 

It is not at all clear, therefore, that the LF for a question contains truth-value denoting scope sites: assuming a semantics like (5-b), all we know for sure is that the LF contains a node of type \(<st,t>\). Our departure point is the skeleton in (6), where the CP has the meaning in (5-b).

(6) **Skeletal LF for question**

\( [_{CP} \text{ who is the mother}] \)

My tactic will be to take the SIH seriously, and flesh out the LF for questions in such a way as to incorporate nodes of type \( t \). This investigation will illustrate, once again, how the SIH turns coordination into a powerful tool for probing the analysis of specific constructions. Concretely, I will propose that there are type \( t \) nodes both *above* and *below* the CP in (6). The higher type \( t \) node will come from a covert performative verb (as in Ross 1970), together with an analysis of performative clauses as truth-value denoting declaratives.

(7) **Introducing a higher type \( t \) node**

\( [_{CP} \text{ I-ask } {}_{CP} \text{ who is the mother}] \)

The lower type \( t \) node will come from a particular compositional approach to deriving the Hamblin denotation. The syntax-semantics mapping I adopt will draw inspiration from Karttunen (1977), and directly follows Heim (2012), and Fox (2013). As a brief preview, the interrogative C in (8) is a relation between propositions, and the CP just above it is of type \( t \).

(8) **Introducing a lower type \( t \) node**

\( [_{CP} \text{ I-ask } {}_{CP} \text{ ANS } \lambda p [_{CP} \text{ who } \lambda x [_{CP} [C p] [_{TP} x \text{ is the mother}]])] ] \)

The higher type \( t \) node will provide a suitable scope site for [and], but will yield pragmatically inappropriate readings with [or]. Conversely, the lower type \( t \) node will provide a suitable scope site for [or], but will yield pragmatically inappropriate readings with [and].

1.2 **Part II: Dissecting question disjunction**

While question conjunction is, I believe, quite readily understood, question disjunction requires close study. Groenendijk & Stokhof (1989) report the intuition that a disjunction \( Q_1\ or \ Q_2 \) offers the speaker a choice of which question to answer: they can provide an answer to \( Q_1 \) or an answer to \( Q_2 \). This pattern is illustrated in (9), modified slightly from Ciardelli et al. (2015):

(9) **Choice answering pattern**

\[ \text{Where might we rent a car, or who might lend us one?} \]

   a. We might be able to rent a car at Avis. \( (\text{addresses } Q_1) \)
   b. John might lend us one. \( (\text{addresses } Q_2) \)

I pursue a different idea, based on a descriptive parallel between certain disjoined questions and a particular kind of simple question, namely *mention-some* questions. Mention-some questions also can be answered
in different ways, and leave the choice up to the responder. An example is given in (10), where (10-a) and (10-b) are both cooperative answers:

(10)  **Mention-some question**
Where can we get gas?
   a. We can get gas at Shell.
   b. We can get gas at Petro-Canada.

I propose that disjoined questions should be analyzed as mention-some questions. I reject Groenendijk & Stokhof’s idea that $Q_1$ or $Q_2$ offers up a choice between questions. Rather, $Q_1$ or $Q_2$ is a single mention-some question: the choice is not between *questions*, but between different *answers* to one question. I show that this analysis follows for free from scoping [or] at the lower type t node in (8). Disjunction at that level creates a single question whose Hamblin set is the union of the Hamblin sets for $Q_1$ and $Q_2$, i.e. a question whose possible answers include all answers to $Q_1$ and all answers to $Q_2$. This Hamblin set — together with an independent theory of the distribution of mention-some readings (due to Fox 2013) — will predict the choice answering pattern. The key contribution of Part II, then, is to fit question disjunction into the broader interrogative typology as a kind of mention-some question.

### 1.3 Roadmap for the chapter

Part I will be undertaken in Section 2. There, I flesh out the LF for questions with a *higher* type t node at which *and* can scope, and a *lower* type t node at which *or* can scope. The remainder of the chapter is dedicated to Part II, on disjunction. In Section 3 and 4, I develop the semantic proposal whereby a question disjunction is interpreted as a single mention-some question. In Section 5, I discuss the pragmatic properties of mention-some questions, and show that disjunctions share these properties. Taking into account the pragmatics helps explain additional puzzling examples. In Section 6, I compare the overall proposal to alternatives, and Section 7 concludes.

### 2 Pursuing the SIH

Our first task is to build up the LF for a question, taking conjunction and disjunction as diagnostic probes for the presence of type t nodes, in accordance with the SIH. Recall our point of departure: the skeletal LF in (6), where the CP has a Hamblin denotation.

(6)  **Skeletal LF for question**

    [CP who is the mother]

To develop the structure, I consider conjunction and disjunction in turn. Although coordinating questions may seem at odds with the SIH, it is very possible to construct an LF for a question with the requisite type t nodes and, in fact, analyses have been proposed with just the right profile.

#### 2.1 Question conjunction

The example in (11) shows that questions can be conjoined, and illustrates how a responder should answer a question conjunction. To fully resolve the conjunction, the responder must resolve both the question in the
left conjunct \((Q_1; \text{’Who is the mother?’})\) and the question in the right conjunct \((Q_2; \text{’Who is the father?’})\). The conjunction in (11-a) may be felicitously answered as (11-b), while neither ‘Mary is the mother’ nor ‘John is the father’ would constitute a full answer on its own.

(11) **Question conjunction**

a. Who is the mother and who is the father?

b. Mary is the mother and John is the father.

If *and* in (11-a) directly operated on the Hamblin sets for the two conjoined questions, this data point would require interpreting *and* not as \([\text{and}]\), but as a higher-type variant, which I will call \([\text{and}_5]\). \([\text{and}_5]\) operates on two \(<st,t>\) meanings to create a new meaning of that same type.

(12) **Higher-type and**

a. \([\text{and}_5](\frac{\lambda Q_{str} \cdot \lambda Q’_{str} \cdot \lambda p_{str} \cdot Q(p)(Q’(p)))}{\lambda Q_{str} \cdot \lambda Q’_{str} \cdot \lambda p_{str} \cdot Q(p) = Q’(p) = 1}\)

b. \([\text{and}_5](\frac{\lambda p_{str} \cdot \{Q_1\}''(p) \land \{Q_2\}''(p)}{\lambda p_{str} . \exists x [p = \lambda w . x \text{ is the mother in } w] \land \exists y [p = \lambda w . y \text{ is the father in } w]}\)

On closer inspection, however, this cannot be the right approach. The computation in (13) illustrates what would be predicted if \([\text{and}_5]\) did apply to the respective Hamblin sets for \(Q_1\) and \(Q_2\).

(13) **Predicted meaning with \([\text{and}_5]\)**

a. \([\text{and}_5](\frac{\{Q_1\}''(\{Q_2\}''(p))}{\{Q_1\}''(p) \land \{Q_2\}''(p)})\)

b. \([\text{and}_5](\frac{\lambda p_{str} \cdot \{Q_1\}''(p) \land \{Q_2\}''(p)}{\lambda p_{str} . \exists x [p = \lambda w . x \text{ is the mother in } w] \land \exists y [p = \lambda w . y \text{ is the father in } w]}\)

On closer inspection, however, this cannot be the right approach. The computation in (13) illustrates what would be predicted if \([\text{and}_5]\) did apply to the respective Hamblin sets for \(Q_1\) and \(Q_2\).

The output is the characteristic function for the set of propositions which are *both* in the Hamblin set for \(Q_1\) and in the Hamblin set for \(Q_2\). In other words, \([\text{and}_5]\) outputs the intersection of \([\{Q_1\}]\) and \([\{Q_2\}]\). Since \([\{Q_1\}]\) contains propositions of *that x is the mother* and \([\{Q_2\}]\) contains propositions *that y is the father*, no proposition is in both sets, and the predicted \([\{11-a\}]\) characterizes the empty set. Clearly, this is not an appropriate meaning for (11-a). *And* must conjoin something other than the Hamblin denoting CPs. So, where does *and* scope?

I propose that *and* scopes *higher* than the CP, above an unpronounced performative prefix. To build up, consider again the simple question in (5), repeated as (14-a). Its communicative impact parallels that of (14-b), which exhibits an overt performative *I hereby ask you*. Informally, both establish ‘Who is the mother?’ as the question under discussion.

(14) **Simple question (a) vs. performative utterance (b)**

a. Who is the mother? = (5)

b. I hereby ask you who the mother is. (performative)

I attribute to (5) an underlying structure exactly like the one that (14-b) wears on its sleeve: the CP *who is mother* is the complement of a performative verb *ask*. Their common syntax is (15). *I (hereby) ask you* is overt in (14-b), but elided in (5).

---

1Note that I use \([Q]\) to indicate the Hamblin denotation for the question \(Q\) throughout the chapter. Once the LF for a question is fully articulated, the highest node in \(Q\)’s LF will not be associated with the Hamblin denotation (there will be material above the Hamblin-denoting CP). Still, I will use \([Q]\) as a convenient shorthand for the Hamblin denotation.
Common underlying syntax

[CP I ask you [CP who the mother is]]

The idea that sentences of natural language rampantly contain elided performatives was originally proposed in Ross (1970) as the Performative Hypothesis. He provides a range of arguments for this move. The paradigm in (16) illustrates one, adapted to the question case at hand.

Support for performative

a. *Should physicists like himself play golf?
b. Should physicists like myself play golf?
c. Should physicists like yourself play golf?

The like-phrase tolerates the reflexives myself or yourself, while himself is ungrammatical. The data in (17) show that reflexives are possible with like only if their antecedent is present intra-sententially in a c-commanding position. As such, the contrast in (16) argues that I and you are syntactically represented in the question LF, as expected with the performative layer.

Baseline data on like

a. Physicists like {Albert, him, *himself} don’t often make mistakes.
b. I told Albert that physicists like himself don’t have often make mistakes.

How are performatives interpreted? Since Lemmon (1962) and Hedenius (1963), performative clauses have widely been taken to denote truth-values (for recent discussion, see Condoravdi & Lauer 2011 and references therein; for a counter-argument, see Searle 1989). This is inspired, in part, by the fact that performatives look exactly like reportative sentences such as (18), which clearly are run-of-the-mill declaratives (Szabolcsi 1982, and subsequent work). (18) does not perform an act of asking, but rather reports on John’s asking.

Reportative use of ask

John asked who the mother is.

The idea is that reportatives and performatives have the same semantics, with performativity arising from that common semantics when the subject is first person and the tense present. While ‘John [i.e. 3rd person] asked [i.e. past]’ is reportative, I ask is performative. I will not provide a complete analysis of ask here, but rather an informal paraphrase, inspired by Condoravdi & Lauer (2011).

Condoravdi & Lauer analyze a number of performatives including claim, for which they propose, roughly, that \( x \) claims \( p \) is true iff there is a communicative event from \( x \) which commits \( x \) to \( p \) being a publicly manifest belief. If \( x \) is the speaker and the sentence is present tense, as in I claim \( p \), the utterance of the sentence itself constitutes such a communicative event, making the truth-conditions of the sentence self-fulfilling. Because the sentence is trivially true, we do not readily intuit a truth-value. Rather, all we intuit is the performative effect. If the subject or tense is changed, the truth-conditions are non-trivial, and a contingent reportative meaning is intuited: John claimed \( p \) or I claimed \( p \) require that there was some communicative event other than the present utterance by another speaker, or by the same speaker at a different time. In a similar vein, \( x \) asks \( Q \) might be true just in case there is a communicative event from \( x \) which publicly commits \( x \) to soliciting the answer to \( Q \). A matrix question with its performative prefix, I ask \( Q \), is then a self-fulfilling truth, and only performativity is intuited, despite a type t meaning. Although a question
is not intuitively true or false, it formally does have a truth-value. The truth-value is accessible to intuitions when the subject is changed from I to John in the reportative (18).

With a type t performative layer in the LF for a question, there is an appropriate scope site for and at which it can be interpreted as [and]; and can scope at the performative layer. Concretely, I propose that the conjunction in (11-a) has the LF in (19):

(19) LF for (11-a)

\[ \&P [I \_P \_ask [C_P \text{who the mother is}]] [\text{and} [I \_P \_ask [C_P \text{who the father is}]]] \]

Two performatives are conjoined, the first of which performs the act of soliciting an answer to Q1, and the second performs the act of soliciting an answer to Q2. It thus follows that the responder must answer both questions to provide all of the solicited information. Now, question conjunction aligns with the SIH: and scopes not at the Hamblin denoting CP, but at a higher performative layer, with a type t meaning.

2.2 Question disjunction

The existence of question disjunction is controversial. As noted in the introduction, Groenendijk & Stokhof (1989) suggested that questions can be disjoined such that Q1 or Q2 offers the responder a choice between answering Q1 and Q2. The data in (20) are a sampling of potential question disjunctions considered in the literature (from Groenendijk & Stokhof 1989, Szabolcsi 1997).

(20) Potential question disjunctions

a. Where do you live? Or, who did you marry?
b. What did Mary read? Or, what did Judy read?
c. Whom does Mary love? Or, whom does Judy love?

Szabolcsi (1997), however, disputed the intuition. She suggested that or in (20-a)-(20-c) “does not really offer a choice, but, instead, is an idiomatic device that allows one to cancel the first question and replace it with the second” (p. 325). In her view, or in (20-a)-(20-c) is unrelated to logical [or]. It is a separate idiom that could be paraphrased as rather or instead, as in (21).

(21) Paraphrasing idiomatic or

Where do you live? Rather, who did you marry?

She supports the idea that or is not logical disjunction in several ways. First, she notes that the strings in (20) are written as two separate sentences, and must be written that way. Without a prosodic break marking a sentence boundary, she judges these data to be severely degraded.

(22) Intra-sentential disjunction blocked

??Where do you live or who did you marry?

In a similar vein, she observes that the sequences in (20) are fully unacceptable in Hungarian, unless the morpheme inkabb (‘rather; instead’) is added, making the strings more parallel to (21). In addition, she observes that question disjunction is possible in embedded environments in Hungarian, but requires a separate complementizer in each disjunct. In effect, Szabolcsi takes this double-complementizer pattern to show that
questions cannot be directly disjoined. Based on these and similar observations, a number of authors have converged on the conclusion in (23).

(23) **The consensus view**

\[ Q_1 \text{ or } Q_2 \] (with logical disjunction) is never possible as a (matrix) question.

It has recently come to light, however, that the consensus view is not correct. Szabolcsi's idea that \( Q_1 \) or \( Q_2 \) idiomatically cancels \( Q_1 \) and replaces it with \( Q_2 \) predicts that, in answering \( Q_1 \) or \( Q_2 \), the responder must answer \( Q_2 \). Answering \( Q_1 \) should be uncooperative. I refer to this answering pattern as the *cancellation schema*, illustrated in (24), where \( A_1 \) is an answer which completely resolves \( Q_1 \), and \( A_2 \) is an answer which completely resolves \( Q_2 \).

(24) **Cancellation schema: possible answers to \( Q_1 \) or \( Q_2 \):**

- a. \#\( A_1 \)
- b. \( \checkmark A_2 \)

As noted earlier, the cancellation schema contrasts with Groenendijk & Stokhof's (1989) original intuition that the responder can choose to provide an answer to either of the disjoined questions. I refer to this answering pattern as the *choice schema*, illustrated in (25).

(25) **Choice schema: possible answers to \( Q_1 \) or \( Q_2 \):**

- a. \( \checkmark A_1 \)
- b. \( \checkmark A_2 \)

I take the choice schema to be the hallmark of true question disjunction, with controversy about the existence of question disjunction centering on whether or not any examples in fact instantiate the choice schema. As noted in the introductory section, Ciardelli et al. (2015) recently argued that the choice schema *is* attested. They observed (9), where each of the answers in (9-a) (addressing \( Q_1 \)) and (9-b) (addressing \( Q_2 \)) are acceptable and equally successful in resolving the question.

(9) **Instantiating the choice schema**

Where might we rent a car, or who might lend us one?

- a. We might be able to rent a car at Avis.
- b. John might lend us one.

In Section 3, I will introduce further data refining the choice schema. For now, I add just one example. Suppose you work in the office at a school, and a sick child comes in. You must call one of the child's parents. You talk to the child for a minute, and then ask (26). Your question is felicitous (if the child has a mother and a father), and (26-a) and (26-b) are both appropriate answers.

(26) **A further example**

So, what's your mother's name or what's your father's name?

- a. My mother's name is Mary.
- b. My mother's name is John.
To make the data point perfectly minimal to the conjunction example in (11-a), and to streamline the LFs, I will take as the running example in this section a more sterile version of (26):

(27) **Running test example**
Who's the mother or who's the father?
   a. Mary is the mother.
   b. John is the father.

The fact that question disjunction is clearly attested places a new constraint on an appropriate LF for questions, given the SIH: there must be a type t node at which [or] can scope to yield an appropriate meaning for data like (27). So far, our LF for a simple question is (28), and the only type t node we have committed to is the top node, above the performative prefix.

(28) **Recall: LF for (5) so far**
[cp I ask you [cp who is the mother]]

Can disjunction take widest scope? In fact, it cannot, as Krifka (1999) observed that it is impossible to disjoin performative clauses. He illustrates with the example in (29):

(29) **Performatives cannot disjoin**
#I hereby baptize you John or I hereby baptize you Mary.

Although this example seems syntactically well-formed, its meaning is problematic pragmatically. Differing somewhat from Krifka’s own account, I point out that (30) yields implausible ignorance inferences. In general, felicitous use of a disjunction \( p \lor q \) requires that the speaker be ignorant about whether \( p \) is true or false, and similarly for \( q \). Because performative clauses are trivially true, it is implausible that the speaker is ignorant about their truth-value. As expected, the deviance carries over to a disjunction of questions with each introduced by an overt performative:

(30) **Substantiating with questions**
#I’m asking you who the mother is, or I’m asking you who the father is.

Since the LF in (31), with [or] above the performative, would reduce (27) to the deviant example in (30), it cannot be the right analysis. There must be another type t node at which [or] can scope.

(31) **Incorrect LF for (27)**
[&p [cp I ask [cp who the mother is]] [or [cp I ask [cp who the father is]]]]

We could envision different analytical moves. One idea might be to propose a more articulated performative layer with an appropriate scope site for [or]. I will suppress this possibility for now, and explicitly consider and reject it later in the chapter. Rather, I take disjunction to diagnose the presence of a type t node *lower* in the question LF, internal to the Hamblin denoting CP itself. To this end, I will adopt a composition for the Hamblin denotation, which has already been proposed by other authors: the composition is an adaptation of Karttunen (1977), based on von Setchow (1996), Heim (2012) and Fox (2013). I introduced this approach in Chapter 4, and present it here anew:
The interrogative $C$ asserts that two propositions are identical. Its sister is a covert pronoun $PRO$, which moves to adjoin on the clausal spine. $PRO$ leaves a trace, which is interpreted as a variable of type <s,t>. This variable saturates the first argument of $C$. The second argument of $C$ is saturated by the intension of the TP out of which $who$ has moved, through IFA. $Who$ is interpreted ex situ as an existential quantifier binding its trace, (34). $CP_2$ thus provides the equation $p = \text{that } x \text{ is the mother}$, with $x$ existentially bound.

(33) Defining the interrogative complementizer
\[
[C]^w = \lambda p_{st} \cdot \lambda q_{st} . \ p = q
\]

(34) Defining $who$
\[
[\text{who}]^w = \lambda f_{et} \cdot \exists x [f(x)]
\]

Movement of $PRO$ targets a position just above $CP_2$. $PRO$ is not itself interpreted, but triggers Predicate Abstraction, binding the propositional variable $p$. This creates (the characteristic function for) a set of propositions \textit{that }$x$ is the mother\textendash ; the target Hamblin denotation.

(35) Hamblin denotation at $CP_3$
\[
[CP_3]^w = \lambda p_{st} . \exists x [p = \lambda w' . \ [x \text{ is the mother in } w']]
\]

With the Hamblin denotation derived in this compositional fashion, there are multiple scope sites for $[\text{or}]$: the TP, $CP_1$, and $CP_2$ are all of type $t$. Since the disjunction in (27) contains two instances of $who$, one in each disjunct, it must be that $[\text{or}]$ scopes above the landing for $who$ in that example, i.e. $[\text{or}]$ must scope at least as high as $CP_2$. I thus propose the LF in (36) with $[\text{or}]$ at $CP_2$.

(36) LF for (27)
\[
[CP_4 \ldots \text{ask } [CP_3 \ PRO \ \lambda 3 [\text{orp } [CP_2\ a \ who \ \lambda 1 [CP_1 a \ [C \ t_3] [TP \ t_1 \text{ is the mother}]]]
\]

\[\text{or } [CP_2b \ who \ \lambda 2 [CP_1 a \ [C \ t_3] [TP \ t_2 \text{ is the mother}]]]
\]

The trace of $PRO$ is not bound within $CP_2$ and, as such, the two conjuncts have assignment-dependent type $t$ meanings. $[\text{orp}]$ operates on $[CP_{2a}]$ and $[CP_{2b}]$ as in (37).
(37) **Composing with [and]**

a. \[ [\text{or}]([\text{CP}_{2a}]^{w,s})([\text{CP}_{2b}]^{w,s}) \]
b. \[ = 1 \iff [\text{CP}_{2a}]^{w,s} \lor [\text{CP}_{2b}]^{w,s} \]
c. \[ \iff \exists x \ [g(3) = \lambda w . x \text{ is the mother in } w] \lor \exists x \ [g(3) = \lambda w . y \text{ is the father in } w] \]

PRO has moved across-the-board (ATB) above the disjunction to form CP₃, and the meaning at CP₃ is then (38), computed from (37) by abstracting over g(3). This is (the characteristic function for) a set of propositions, in particular a set containing both propositions of the form \( \textit{that } x \text{ is the mother} \) and propositions of the form \( \textit{that } y \text{ is the father} \).

(38) **Prediction for CP₃**

\[ [\text{CP}_3]^w = \lambda p . \exists x \ [p = \lambda w' . x \text{ is the mother in } w'] \]
\[ \lor \exists y \ [p = \lambda w' . y \text{ is the father in } w'] \]

Scoping \( \textit{and} \) above the performative layer resulted in two successive acts of questioning, each posing a separate question. The first conjunct performed the act of questioning ‘\( \textit{Who is the mother?} \)’ and the second conjunct performed the act of questioning ‘\( \textit{Who is the father?} \)’. Scoping \( \textit{or} \) at a lower type \( t \) node has created a single question with a new meaning different from either ‘\( \textit{Who is the mother?} \)’ or ‘\( \textit{Who is the father?} \)’. The Hamblin set for this new question is (38), which is the union of the Hamblin sets for ‘\( \textit{Who is the mother?} \)’ and ‘\( \textit{Who is the father?} \)’. To make this explicit, we could equivalently re-state (38) as:

(39) **Equivalent statement of [CP₃]**

\[ [\text{CP}_3]^w = \lambda p . \ [\text{who is the mother}]^{w}(p) \lor [\text{who is the father}]^{w}(p) \]

The next section will address in detail how the choice answering pattern arises from this Hamblin set. In brief, what is crucial to note for now is that the Hamblin set does contain all possible answers to the disjunction. Recall the illustration of the choice schema from above:

(27) **Recall: choice schema**

Who is the mother or who is the father?

a. Mary is the mother.

b. John is the father.

On the one hand, the proposition expressed by (39-a) — \( \lambda w . \textit{Mary is the mother in } w \) — is in the extension of \([\text{CP}_3] \) because that proposition is in the extension of \([\text{who is the mother}] \). On the other hand, the proposition expressed by (39-b) — \( \lambda w . \textit{John is the father in } w \) — is also in the extension of \([\text{CP}_3] \) because that proposition is in the extension of \([\text{who is the father}] \). The Hamblin set for a question should be the set of its possible answers, and the predicted \([\text{CP}_3] \) precisely fits the bill: the possible answers to (27) are those propositions in either \([\text{who is the mother}] \) or \([\text{who is the father}] \). More generally, we can state the result as the **Union Hypothesis** in (40).

(40) **The Union Hypothesis**

\[ [Q_1 \text{ or } Q_2]^w = \lambda p . \ [Q_1]^{w}(p) \lor [Q_2]^{w}(p) \]
The proposed LF for questions contains a type t node at which [or] can scope to yield an interpretation of $Q_1$ or $Q_2$ as a single question denoting the union of $[Q_1]$ and $[Q_2]$. Given the choice schema, the responder can answer $Q_1$ or $Q_2$ with an answer to $Q_1$ or an answer to $Q_2$ — precisely the answers in $[Q_1]$ or $[Q_2]$ under the Union Hypothesis. The Union Hypothesis closely resembles ideas in Ciardelli et al. (2015), which they develop in Inquisitive Semantics.23

It is important to flag that the presence of a new, lower type t node in the LF does not alleviate the need for the performative layer and corresponding higher type t node. We saw earlier that scoping [or] above the performative would yield a pragmatically deviant meaning. The converse problem arises with and: scoping it at the lower type t node would yield a deviant meaning:

(41) Hypothetical LF with and at CP2
[CP4 ... ask [CP3 PRO $\lambda_3 [\&p [CP2a \text{ who } \lambda_1 [CP1a [C t_3] [TP t_1 \text{ is the mother}]]]$
[and [CP2b \text{ who } \lambda_2 [CP1a [C t_3] [TP t_2 \text{ is the mother}]]]]]

(42) Prediction for CP3
a. $[CP3]^{\pi}$
b. $= \lambda p \cdot \exists x [p = \lambda w'. x \text{ is the mother in } w'] \land \exists y [p = \lambda w'. y \text{ is the father in } w']$
c. $= \lambda p \cdot [\text{who is the mother}]^{\pi}(p) \land [\text{who is the father}]^{\pi}(p)$

The LF in (41) outputs the Hamblin set in (42). This predicted meaning is identical to the one we already entertained in (13) above, where [and] directly operated on the Hamblin sets for the two questions. $[CP3]$ is the intersection of the Hamblin sets for [who is the mother] and [who is the father], and that intersection is empty. The proposed LF for questions, then, can replicate the effect of [and] in a way consistent with the SIH. While not helpful for and, this looks promising for or, which results in union, not intersection.

2.3 Concluding Part I

From the perspective of the SIH, coordination constitutes a valuable tool to probe into the LF for questions. Conjunction diagnoses a covert performative layer and supports the idea that performative clauses denote truth-values. Disjunction sheds light on how the Hamblin denotation comes about compositionally: it argues that there must be a type t node below the Hamblin denoting CP, but above the landing site of wh-words, as in the Karttunen-inspired composition I adopted. The LF for a simple question is (43): [and] scopes at CP4, [or] scopes at CP2, and this cannot be reversed on pain of pragmatic deviance.

(43) LF for basic question
[CP4 I ask you [CP3 PRO $\lambda_3 [CP2 \text{ who } \lambda_1 [CP1 [C t_3] [TP t_1 \text{ is the mother}]]]]$]

I will consider and reject alternative compositional mechanisms in the last part of the chapter. For now, I will continue on to Part II, which probes deeper into the semantic properties of question disjunction. There is still more to be said about exactly how the choice schema arises, but the Union Hypothesis will provide a stalwart pushing off point for this investigation.

---

2In Inquisitive Semantics, questions introduce alternatives and “raise the issue” of which alternative is correct. The idea that questions, in effect, introduce alternatives is thus common to the Hamblin semantic treatment I have adopted and Inquisitive Semantics. As I see it, the substantive innovation of my analysis will come in the second half of the chapter, which connects question disjunction to grammatical theories of mention-some questions.

3A similar idea was also put forward for the analysis of alternative questions in Biezma & Rawlins (2012).
3 Two profiles of question disjunction

To explain how a question is answered, two ingredients are necessary: (i) a Hamblin set, i.e. a set of all propositions which could, in principle, constitute possible answers to the question, and (ii) a theory of which element or elements in the Hamblin constitute appropriate answers in a given circumstance. For disjunction, I have provided the first ingredient: the Hamblin set for $Q_1$ or $Q_2$ is determined in accordance with the Union Hypothesis. Over the next three sections, I develop the second ingredient.

This section focuses on empirical matters in order to figure out exactly what it is we have to explain. I identify two sorts of situations in which question disjunctions may be felicitous. The choice schema is observed in both, but in a subtly different way. The distinction made in this section is purely descriptive and it will not have theoretical status in the final analysis. Indeed, an aim of the analysis is to account for both situation profiles in a unified fashion. Still, en route to the final analysis, it will prove useful to look at the data in a bifurcated way.

3.1 Profile 1: only one question answerable

The first profile of felicitous question disjunction takes the following form. Only one of $Q_1$ and $Q_2$ has a true answer, while the other does not. The questioner knows that that one question has a true answer, but is uncertain which question is answerable. The questioner wants to leave it up to the responder to provide an answer to whichever question is, in fact, answerable. To illustrate, consider (44). Speaker A informs speaker B that the department made a staffing change, but B does not know the nature of the staffing change.

(44) Establishing Profile 1
   a. A: The department made a staffing change this weekend.
   b. B: Really?! Who got hired or who’s leaving?

If there was a single staffing change, exactly one of $Q_1$ and $Q_2$ has a true answer. But, B does not know which one that is and, as such, is not in a position to cancel $Q_1$, as Szabolcsi proposed. Rather, B wishes to give A the option to answer either $Q_1$ or $Q_2$, whichever turns out to be appropriate. If John got hired, A should answer as (45-a) and, if John is leaving, A should answer as (45-b). There is only one correct answer, but that answer could address either $Q_1$ or $Q_2$, depending on what the facts happen to be. This is how the choice schema manifests in Profile 1.

(45) Illustrating the choice schema
   a. $\lor A$: John got hired.
   b. $\lor A'$: John’s leaving.

The example in (46) further instantiates the profile. Suppose that the speaker comes to a church yard and hears bells. They know that bells signify either a wedding or a funeral, but they are uncertain which one is happening. The speaker may ask (46). Again, just one of the two questions has a true answer, and it is up to the interlocutor to answer the right question.\footnote{Thanks to David Pesetsky (p.c.) for offering examples.}

(46) Replicating Profile 1
   Who’s getting married, or who’s funeral is it?
An internet search furnishes additional corroborating data. The three data points that follow are adapted from different websites, and all seem at least likely to have Profile 1. The question in (47-a), for instance, is felicitous if Bob committed one crime, and the speaker is uncertain whether it was a murder, or something else. The intuition is similar for the other two examples.

(47) **Naturally occurring data**

a. Who did Bob kill, or what other crime did he commit?

b. *(Trying to get a product the interlocutor has received.)*
Who did you call, or what else did you do?\(^5\)

c. *(The interlocutor has finally started feeling better in life.)*
Who did you meet or what did you read that helped you?\(^6\)

3.2 **Profile 2: both questions answerable**

The second profile of question disjunction is more familiar from the examples introduced in Section 2.2. *Both \(Q_1\) and \(Q_2\) have a true answer, and knowing the answer to either question is sufficient to achieve some conversational goal. The questioner provides both questions to the interlocutor, who has the option to provide an answer to either question.*

The running example from the last section instantiates Profile 2. Suppose, as before, that you require information about one parent to make a phone call home. If the person has both a mother and a father, then both questions have a true answer, and either an answer to \(Q_1\) or an answer to \(Q_2\) is felicitous, and sufficient to achieve the objective of attaining the required information for the phone call. The interlocutor can answer however they prefer.

(27) **Establishing Class 2**

Who’s the mother or who’s the father?

a. Mary is the mother.

b. John is the father.

The example in (48) is natural in a similar sort of context. Suppose the questioner is a customer service agent at an American Blue Cross call center, and the interlocutor requires assistance. To locate the interlocutor’s file, the questioner needs one of two personal identifiers: either their name, or their social security number. Since Blue Cross customers necessarily have both a name and an SSN, both questions in (47) must have a true answer. The responder could choose to answer with (48-a) or (48-b), and either would be sufficient to achieve the goal of locating the file.

(48) **Replicating Profile 2**

What’s your name or what’s your SSN?

a. My name is John.

b. My SSN is 123-45-6789.

---


A further illustration is (49), which is adapted from a website. Assuming the interlocutor has a preferred superpower and a spirit animal, both questions in (49) have a true answer. The goal in the context is to learn some piece of information about the interlocutor to begin a conversation. Either an answer like (49-a) or one like (49-b) is sufficient to achieve the goal, and the choice between them is left to the interlocutor.

(49) Naturally occurring data
What’s your superpower or what’s your spirit animal?

a. My superpower is flying.
b. My spirit animal is a polar bear.

3.3 Local summary

This section has identified two classes of question disjunction. Both kinds of disjunctions evince the choice schema in the sense that they may, in principle, receive an answer addressing either $Q_1$ or $Q_2$. It is important to emphasize, however, that the choice schema manifests differently in Profile 1 than in Profile 2. In Profile 1, the disjunction could only be answered in one way in a given circumstance. It was up to the responder to figure out the appropriate answer and, depending on the facts, the answer would address either $Q_1$ or $Q_2$. In Profile 2, the disjunction may be addressed in more than one way in the very same circumstance. The responder can address $Q_1$ or they can address $Q_2$, and the choice is one for them to make.

Note that I have not included Ciardelli et al.’s (2015) example (Where might we rent a car, or who might lend us one?) in the above bifurcation. This example would often pattern as Profile 2, assuming there are usually both rental options and borrowing options — but it has a somewhat different character from the other examples in a way I will make precise in Section 5, once more theoretical background is in place. For now, I will stick to the data in the last two subsections: how does the choice schema arise when disjunctions occur in situations with one profile or the other?

4 Question disjunctions are mention-some

Given the Union Hypothesis, $Q_1$ or $Q_2$ is interpreted as a single question, denoting the union of $[Q_1]$ and $[Q_2]$. If disjunctions are just internally complex single questions, then the answering patterns observed with them should fit in with independently observable answering patterns with simple questions. A useful starting point, then, is to ask: how are simple questions answered?

4.1 Mention-all questions

Most questions solicit exhaustive answers. The question in (50-a), for instance, solicits an answer which names all of the people who addressed the committee, as the paraphrase in (50-b) makes clear. This sort of question is commonly referred to as mention-all (‘MA’).

(50) Mention-all question

a. Who addressed the committee?
b. “Tell me everyone who address the comittee.”
We can state the answerhood conditions more precisely based on the Hamblin denotation. I take the Hamblin denotation to be (51), which characterizes a set of propositions that \( X \) addressed the committee for different salient atomic or plural entities \( X \). The set includes the propositions in (51-a)-(51-c). We will see the derivation for this Hamblin downstream in the chapter.

(51) Hamblin set for (50-a)

\[
\begin{align*}
\text{a. } & \lambda \cdot \exists X \ [p = \lambda w' \cdot \forall x <_{AT} X [x \text{ addressed the committee in } w']] \\
\text{b. } & \{ \lambda w \cdot \text{ John addressed in } w, \\
\text{c. } & \lambda w \cdot \text{ Bill addressed in } w, \\
\text{d. } & \lambda w \cdot \text{ John addressed in } w \land \text{ Bill addressed in } w, \ldots \}
\end{align*}
\]

With the Hamblin set now in place, the MA answerhood conditions may be stated in either of two ways. The MA answer to a question with Hamblin set \( Q \) at world \( w \) is either (52-a) or (52-b).

(52) MA answerhood conditions

\[
\begin{align*}
\text{a. } & \text{ The strongest proposition in } Q \text{ that is true at } w. \quad \text{(Dayal 1996)} \\
\text{b. } & \text{ The conjunction of propositions in } Q \text{ that are true at } w. \quad \text{(Karttunen 1977)}
\end{align*}
\]

For the question in (50-a), if just John addressed the committee, the only true proposition in the Hamblin set is (51-b), which is the MA answer. If John and Bill addressed the committee, the true propositions are all of (51-b)-(51-d). Their conjunction is equivalent to the strongest of them, (51-d), which is thus equivalent to the MA answer under either way of stating the answerhood conditions.

4.2 Question disjunctions are not mention-all

Now, let us consider: could question disjunctions be analyzed as MA questions? The answer is this: certain question disjunctions are consistent with MA answerhood conditions, but not all of them. It is in this respect that the bifurcation of data into two profiles is useful. A MA analysis of question disjunction is consistent with Profile 1 situations, but not Profile 2 situations.

To start with Profile 1, a representative data point from Section 3.1 is repeated below. The Union Hypothesis assigns to (44) the Hamblin denotation in (53), which characterizes a set containing propositions of the form that \( x \) got hired, and propositions of the form that \( y \) is leaving.

(44) Recall: Profile 1

Who got hired or who's leaving?

(53) Under the Union Hypothesis

\[
[(44)]^w = \lambda p \cdot \exists x \ [p = \lambda w' . \ x \text{ got hired in } w'] \lor \exists y \ [p = \lambda w' . \ y \text{ is leaving in } w']
\]

Granting, as before, that there was only one staffing change, the Hamblin set will contain exactly one true proposition, and that true proposition will be the predicted MA answer. At a world, \( w_1 \), where John got hired, the true proposition is (54), and MA answerhood conditions correctly that the disjunction should be answered: 'John got hired'.

(54) True proposition in Hamblin set at \( w_1 \)

\[
\lambda w \cdot \text{ John got hired in } w
\]
The choice schema as manifest in Profile 1 data is captured. Although a given question has a unique MA answer at a given world, the identity of that answer varies from world to world. We already noted such variability with the simple question in (50-a). At a world where just John addressed, the MA answer is 'John is address the committee'. At a different world, where John and Bill addressed, the MA answers changes to 'John and Bill addressed the committee'.

Parallel world-by-word variability captures the choice schema in Class 1 data. Compare the MA answer for the disjunction at $w_1$ to its MA answer at another world $w_2$. At $w_1$ above, the MA answer was 'John got hired.' If $w_3$ is such that John is leaving, then the true proposition in the Hamblin set is (55), and that becomes the MA answer: 'John is leaving'.

(55) True proposition in Hamblin set at $w_2$
\[ \lambda w. \text{John is leaving in } w \]

Since the MA answers vary between elements of the Hamblin set across worlds and because the Hamblin set for (44) contains propositions of the form that $x$ got hired and that $y$ is leaving, the MA answer will vary between those two forms, depending on the facts at the world. This is, in fact, how the choice schema manifest in the Profile 1 data point.

Although MA answerhood conditions are viable for a fragment of the data, they do not generalize to Profile 2 examples, as I illustrate with (27). The key point is that Profile 2 data can be answered in different ways at the very same world — and MA answerhood conditions predict a unique answer.

(27) Recall: Profile 2
Who's the mother or who's the father?

The Union Hypothesis assigns to (27) the Hamblin denotation in (56), which characterizes a set containing propositions that $x$ is the mother and propositions that $y$ is the father. If the interlocutor has a mother and a father, the Hamblin set contains two true propositions. At $w_1$, suppose they are are (57-a) and (57-b). As we have seen, both are equally appropriate answers.

(56) Under the Union Hypothesis
\[ \llbracket (27) \rrbracket^w = \lambda p. \exists x [p = \lambda w'. x \text{ is the mother in } w'] \lor \exists y [p = \lambda w'. y \text{ is the father in } w'] \]

(57) True proposition in Hamblin set at $w_1$
\begin{align*}
a. & \lambda w. \text{Mary is the mother in } w \\
b. & \lambda w. \text{John is the father in } w \\
\end{align*}

Now, what would MA answerhood conditions predict? The precise prediction depends on how the MA answerhood conditions are stated. If the complete MA answer is the strongest true proposition in the Hamblin set, (27) would be unanswerable, since there is no strongest true proposition: the two true propositions are logically independent. Alternatively, if the complete MA answer is the conjunction of true answers, the question is answerable, but the MA answer would be (58-a), not (58-b) or (58-c). The responder would have to answer the question disjunction in the way they would answer a question conjunction (cf. (11-a)).

194
Possible MA answering pattern (incorrect)

a. Mary is the mother and John is the father.
b. #Mary is the mother.
c. #John is the father.

Clearly, MA answerhood conditions do not capture the choice schema in Profile 2 data, and a unified analysis of question disjunction cannot analyze them as MA questions. For an alternative, we need to go back to simple questions, and consider how else simple questions can be interpreted.

4.3 Mention-some questions

Not only do MA answerhood conditions not capture question disjunction, but they do not capture all simple questions either. The question in (59) provides an illustration:

Ambiguous simple question
Who can chair the committee?

This question can be interpreted in two ways. First, it can be a MA question, asking the responder to name everyone who is eligible to serve as committee chair. Yet, it can also have another interpretation, where it solicits something weaker: it can ask the responder to name some committee chair, rather than all. Interpreted in the latter way, the question is mention-some (‘MS’).

MA vs. MS paraphrases

a. “Tell me everyone who can chair the committee.”
b. “Tell me someone who can chair the committee.”

The MS interpretation shows a descriptive similarity with the Profile 2 disjunction data. Just as Profile 2 disjunctions have multiple possible answers which the responder can freely choose between in a given circumstance, so too does a MS simple question. If John and Bill are both eligible chairs, then the MA would be (61-a), while both (61-b) and (61-c) are equally cooperative MS answers.

Possible answers to (59)

a. John and Bill can chair. (MA answer)
b. John can chair. (possible MS answer)
c. Bill can chair. (possible MS answer)

In the remainder of the chapter, I translate the descriptive parallel between Profile 2 data and MS questions into an analytical parallel. This is the idea:

Hypothesis
Question disjunctions are simple MS questions.

The idea of linking disjunction to MS was previously pursued in Szabolcsi (1997), which proposed that the disjunction in (63-a) receives a mention-some analysis, matching (63-a) with the paraphrase (63-b).
(63) **Mention-some precedent**

a. Who did Fido or King bite?

b. “Tell me someone who is such that either Fido or King bit him.”

An idea reminiscent of mention-some was also entertained in Groenendijk & Stokhof (1989), based on Bennett (1979) and Belnap (1982). Groenendijk & Stokhof, however, abandoned the approach, since it did not extend to question conjunction. In my view, conjunction should receive a separate analysis, as conjunction of performatives.8

5 **The analysis**

In this section, I build the formal proposal. I show that a MS analysis for question disjunction “follows for free” from the Union Hypothesis, combined with an independent analysis of the MA-MS distinction, due to Fox (2013). Profile 1 and 2 data will be captured together: Profile 1 data are just MS questions that happen to have only one true answer. Section 5.1 introduces the analysis of simple questions. The disjunction data are fit into that analysis in Section 5.2.

5.1 **Analyzing simple questions**

How is the MA-MS distinction analyzed in simple questions? As noted earlier, a full theory of how questions are answered has two ingredients: (a) set of possible answers, and (b) a mechanism to determine which of these possible answers are appropriate at a given world. The mechanism in (b) must capture the MA vs. MS distinction. In principle, this mechanism could involve either of two systems: it could be purely pragmatic, or it could be grammaticalized into the semantics. Both have been considered.

Under the pragmatic view, MA and MS questions are identical in their semantics. The Hamblin set is grammatically determined, but how to choose appropriate answers out of the Hamblin set is entirely up to the pragmatics. The MA and MS answerhood conditions would be two symptoms of a more general answerhood condition, which we might state like this. For a question with Hamblin set Q, p is a complete answer in w if p is described by (64).

(64) **Answerhood condition**

A proposition in Q that is true at w and sufficient to resolve the asker’s decision problem.

Van Rooj (2004) suggested that a speaker asks a question to resolve some dilemma or, in his terminology, “decision problem”. The responder’s job is to answer the question in such a way as to resolve that decision problem (the answerhood condition in (64) is inspired by van Rooj, though formulated differently; see

---

8 Szabolcsi (2016) relates Ciardelli et al.’s data point to what she terms “exemplifications”, illustrated in (i) with a naturally occurring example. In (64), the sentence raises the issue of what the label should state, and gives examples of what it could state. In a similar vein, Ciardelli et al.’s disjunction seems to involve a general question of where we can get a car, and gives examples of how to address that question.

(i) The label must state, for example, the nature of a nutritional or compositional change, or the presence of an allergen.

With reference to an earlier version of this work, Szabolcsi notes that the exemplification intuition is captured under the MS analysis. In addition, she suggests a second way of capturing the intuition, being to analyze exemplifications as conjunction, despite the surface morphology. Ciardelli et al.’s example would paraphrase along the lines of: ‘Where can we get a car. Examples of questions to address this are: where can we rent one, and who might lend us one.’
also Groenendijk & Stokhof 1984). Depending on the particular decision problem, a MA answer may be required, or a MS answer may be sufficient. Determining how to answer a question, then, requires figuring out what decision problem motivated that question. The conversational goals regulate MA vs. MS.

For the question in (59) above, for instance, consider two different contexts. First, suppose I need to fill out a report listing all potential committee chairs to give to my superior who will then decide on a chair. In this context, my dilemma is how to fill out the report, and this can only be resolved with a MA answer. Alternatively, suppose I am at a boring faculty meeting and will be grateful for anyone to take the burdensome task of chairing a time consuming committee. Now, my dilemma is to fill the chair position with anyone who is willing, and I can ask (59) as MS.

Although the pragmatic view seems attractive prima facie, there is more going on. The pragmatic view makes a strong prediction: it predicts that it should be possible to interpret any and all questions as MA or MS, depending on contextual factors. George (2011) showed that this is not so. In particular, he identified certain simple questions which lack a MS reading. The question in (59) has a special property: it contains an existential quantifier, the possibility modal can. George observed that MS readings only arise in simple questions when an existential is present. His original contrast is (65), involving some:

(65) Constraints on mention-some
a. Who are some of your friends? $MA, (*MS$

Fox (2013) provides the example in (66) to replicate the contrast with can. Mirroring the judgments in (65), the first example with can allows MS, just as (59) does. Conversely, the second example without can is necessarily MA. The two examples in Section 4 showed the contrast, too. They are repeated together in (67): whereas (67-a) allows for MS, (67-b), without can, must be MA.

(66) Illustrating with can
a. Where can we get gas? $MA, MS$

(67) Minimal pair with (59)
a. Who can chair the committee? = (59); $MA, MS$

The correlation between the presence of an existential and the availability of MS strongly suggests that the grammar plays a role in regulating the two interpretations: there seems to be some grammatical property that interacts with an existential to derive MS in simple questions. When there is no existential, the grammar cannot derive MS in the same environment.

I will present a recent grammatical analysis of the MA vs. MS distinction in simple questions, due to Fox (2013). By combining Fox’s theory with the Union Hypothesis, I show that the status of Class 1 data as mention-all and Class 2 data as mention-some “follows for free” as a predicted consequence of the system. These results have important consequences: they lead to a full analysis of disjunction rooted in the Union Hypothesis, fit disjunction into the broader interrogative typology on an analytical level, and lend fresh support to Fox’s ideas about how mention-all and mention-some interpretations come about.
5.1.1 Dayal’s ANS

What we need is a grammatical operator that applies to a Hamblin set and whittles it down to the set of appropriate answers at a given world. The operator, ANS — originally proposed by Dayal (1996) to accommodate question embedding — gets close. I already introduced ANS in Chapter 4, and its semantic value is repeated in (68).

(68) Defining ANS (Dayal 1996)
\[
[\text{ANS}]^w = \lambda Q, : \exists p [w \in p \in Q \land \forall p' \in Q [w \in p' \rightarrow p \subseteq p']] 
\]

ANS takes as its argument a set of propositions. It presupposes that the set contains a proposition that is true at the evaluation world and entails all other true propositions. ANS returns that proposition. In effect, it whittles down a Hamblin set to its MA answer, with the MA answer understood as the strongest true proposition in the Hamblin set. If the answerhood conditions are stated as (69), ANS can be seen as adding a grammatical dimension to MA. For a question with Hamblin set Q, p is a complete answer in w if p is described by (69).

(69) Answerhood condition
p is the proposition given by \([\text{ANS}]^w(Q)\).

Fox’s theory will amount to a re-formulation of Dayal’s ANS to accommodate both MA and MS and predict their distribution. To get warmed up, it will be useful to see exactly how Dayal’s own ANS would work in the simple question in (59) and to diagnose exactly where it falls short of deriving MS. There are two possible LFs for (59) and I consider each in turn. This is the place to start:

(70) Possible LF for (59) (each > can)

One point of note has to do with the relative scope of the modal can and the covert distributive operator each, which I assume is present in the LF, following Fox: each takes wide scope above can. This results in a Hamblin set containing propositions that every atom x of X can chair the committee, where X is a singular
or plural entity. Taking John as X yields (71-b), taking Bill as X yields (71-c), and taking the plurality John⊕Bill as X yields a proposition equivalent to the conjunction in (71-d).

(71) **Predicted Hamblin denotation**

a. \[\langle CP_4 \rangle^w = \lambda p . \, \exists X \, [p = \forall x <_{AT} X \, [\Diamond [\text{chair}(x, \text{committee})]]] \]
b. \(\approx \{ [\Diamond [\text{chair}(\text{John}, \text{committee})]], \)
c. \(\Diamond [\text{chair}(\text{Bill}, \text{committee})], \)
d. \(\Diamond [\text{chair}(\text{John}, \text{committee})] \land \Diamond [\text{chair}(\text{Bill}, \text{committee})], \ldots \}

An observation that will become important in a moment is that the Hamblin set is closed under conjunction: for the three propositions shown, the conjunction of (71-b) and (71-c) is (71-d). The same holds for any pair of propositions in the Hamblin set: if \(p\) is in the Hamblin set and \(q\) is in the Hamblin set, \(p \land q\) is in the Hamblin set, as well.

With a Hamblin set in hand, how does ANS integrate? Given its argument structure, ANS must be interpreted just above the node introducing the Hamblin set and below the performative prefix. This means ANS must integrate just above PRO, since abstracting over the trace of PRO, \(t_3\), creates the Hamblin set. Because the Hamblin set is closed under conjunction, if it contains any true proposition, it will necessarily contain a true proposition which entails all other true propositions. The presupposition of ANS will thus be met and ANS will output that unique strongest true proposition, designating it as the appropriate answer to the question.

Concretely, suppose a world, \(w_1\), where John and Bill are both eligible committee chairs. At that world, all three propositions in (71) are true, and (71-d) asymmetrically entails both (71-b) and (71-c). \([\langle \text{ANS} \rangle^{w_1} ([CP_4])]\) outputs the proposition in (71-d) and, to accurately resolve the question, the responder must assert that proposition. They must provide the MA answer: 'John and Bill can chair the committee.' More generally, the proposition ANS picks out at a given world will always express the conjunction of all eligible chairs at that world, leading to a MA answering pattern.

Hence, ANS derives out of one LF for (59) a MA reading, leaving MS so far unaccounted for. Consider a second LF for (59) does not help matters, but it is important to see that a different problem arises than in the first LF. This LF differs from the last one in the relative scope of \(\text{can}\) and \(\text{each}\): now, \(\text{can}\) takes wide scope above each, instead of vice versa.
(72) Alternative LF for (59) (can > each)

\[
\begin{array}{c}
\text{CP}_5 \\
\downarrow \text{ask} \\
\text{ANS} \\
\downarrow \text{CP}_3 \\
\downarrow \text{PRO} \\
\downarrow \lambda 2 \\
\downarrow \lambda 1 \\
\downarrow \text{CP}_1 \\
\text{C} \\
\text{t}_2 \\
\text{TP} \\
\end{array}
\]

The effect of the scope reversal is to change the Hamblin denotation at \(\text{CP}_3\) to a set of propositions of the form \(\text{that it's possible that every atom } x \text{ of } X \text{ chair the committee.}\) This set includes the propositions in (73-b)-(73-d), with \(X\) as John, Bill, and John\&Bill, respectively.

(73) Predicted Hamblin denotation

a. \(\llbracket \text{CP}_3 \rrbracket^w = \lambda p . \exists X [p = \diamond (\forall x <_{AT} X \text{ chair}(x, \text{committee}))]\)

b. \(\approx \{ \diamond \text{chair}(\text{John, committee}) \} ,\)

c. \(\diamond \text{chair}(\text{Bill, committee}) ,\)

d. \(\diamond \text{chair}(\text{John, committee}) \& \text{chair}(\text{Bill, committee}) \)

Comparing (73) with (71) above, it is clear that the scope of the distributive relative to the modal has no effect on what proposition makes it into the Hamblin set when \(X\) is an atomic individual. Taking John as \(X\) yielded (73-b), which is identical to (71-b). By the same token, taking \(X\) as Bill yielded (73-c), which is identical to (71-c). There is, however, a crucial effect when \(X\) is plural: (73-d) is not equivalent to (71-d) and, as such, is not the conjunction of (73-b) and (73-c). As this illustrates, the new Hamblin set, unlike the earlier one, is not closed under conjunction.

The make up of the Hamblin set in (73) places the presupposition of ANS on shakier ground: because the Hamblin set is not closed under conjunction, it is now possible for it to contain a true proposition, but still not contain a true proposition that entails all other true propositions. In a context where there is more than one eligible chair, the new LF will not yield either a MA reading or a MS reading: there will be presupposition failure. At \(w_1\), with John and Bill eligible chairs, the true propositions in the Hamblin set are (73-b) and (73-c). Assuming the committee can have only a single chair, (73-d) is false. Since (73-b) and (73-c) are logically independent, neither true proposition entails the other, and ANS’s presupposition fails. The question is unanswerable.

Overall, Dayal’s ANS yields a mention-all reading when the Hamblin set contains a true answer that entails all other true answers (as with \(\text{each} > \text{can}\)), and otherwise yields presupposition failure (as with \(\text{can}\))
each). Fox's tactic will be to re-formulate ANS so that it converges with Dayal's ANS in its derivation of MA, but repurposes Hamblin sets which yield presupposition failure for Dayal's ANS to derive MS. For Fox, the first LF for (59) will yield MA and the second will yield MS, rather than presupposition failure.

5.1.2 Modifying ANS (Fox 2013)

Fox modifies ANS. To facilitate discussion, I define a notion of an answer being Strongest-True in a particular new sense, which I state as (75). Paraphrasing, a proposition \( p \) in the Hamblin set for a question \( Q \) qualifies as Strongest-True at a world \( w \) just in case (a) \( p \) is true at \( w \), and (b) \( p \) is not entailed by any other proposition in \( Q \) true at \( w \).  

(74) Re-defined ANS
\[
[ANS_2]_w^Q = \lambda Q: \exists p \in Q \land \neg \exists p' [w \in Q \land p' \subseteq p]
\]

(75) Defining 'strongest-true'
For a question \( Q \) at a world \( w \), a proposition \( p \in Q \) is strongest-true iff (a) and (b) hold.

a. \( p(w) \)

b. \( \neg \exists p' [w \in p' \land p' \subseteq p] \)

What is the presupposition? For Dayal, ANS (hence, 'ANS_1') presupposed that the Hamblin set contains a true proposition that entails all other true propositions. Adopting the new terminology, ANS_1 presupposed that the Hamblin set contains a unique Strongest-True proposition. For Fox, ANS (hence, 'ANS_2') carries a weaker presupposition: that the Hamblin set contains some Strongest-True element. If the presupposition of ANS_1 is satisfied, the presupposition of ANS_2 is satisfied as well, but not necessarily vice versa. The second LF for (59), which suffered from presupposition failure under ANS_1 in the previous section, will be presupposition satisfying now, as we will see in detail: it will contain multiple propositions which qualify as Strongest-True. By weakening the presupposition of ANS, Fox reduces the cases which yield presupposition failure and sets the stage to transform previous presupposition failures into mention-some readings.

Now, what does ANS contribute? Whereas ANS_1 returned a proposition (the unique Strongest-True proposition in the Hamblin set at the evaluation world), ANS_2 returns a set of propositions. \([ANS_2](Q)(w)\) returns the set of all propositions in \( Q \) which qualify as Strongest-True in \( w \). If there is a unique Strongest-True proposition, that set will be a singleton; if there are multiple Strongest-True propositions, it will have a higher cardinality. In Fox's analysis, the set output by ANS_2 determines how the question can be answered. For a question with Hamblin set \( Q \), a proposition \( p \) counts as a complete answer in \( w \) if \( p \) is:

(76) Answerhood condition for \( Q \) at \( w \)
A proposition in the set \([ANS_2]^w(Q)\).

With the answerhood conditions defined as (76), the cardinality of the ANS-set determines whether the question has just one complete answer at a given world, or multiple complete answers. If the ANS-set

---

9In fact, this is Fox's "first stab" at re-formulating ANS. This ANS incorrectly predicts mention-some readings with singular which questions. Which professor chaired the committee? carries a uniqueness presupposition that exactly one professor can chair, and does not have a mention-some reading. Accounting for uniqueness with which was the motivation for Dayal's ANS. Fox proposes a further revision to ANS which accounts both for his generalization and uniqueness with singular which. I refer the reader to Fox's work for further details. For our purposes here, we can set this issue aside, since it will not be germane to disjunction.
contains just one element, there is a unique complete answer, which will correspond to the MA answer. On
the other hand, if the ANS-set contains multiple elements, all of them qualify as complete answers, and the
responder can resolve the question with any of them. This optionality will create MS.

(77) Predicted answering pattern
   a. Unique possible answer iff $|[[\text{ANS}_2]]^w(Q)| = 1$.
   b. Optionality between multiple possible answers iff $|[[\text{ANS}_2]]^w(Q)| > 1$.

I will use the terms “MA question” and a “MS question” according to the definitions in (78) and (79). In
effect, I treat a MA question as one which necessary has a unique answer, and a MS question as one which
may have multiple answers.

<table>
<thead>
<tr>
<th>Defining MA question and MS question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MA question:</strong> a question $Q$ is a MA question iff $</td>
</tr>
<tr>
<td><strong>MS question:</strong> a question $Q$ is a MS question iff $</td>
</tr>
</tbody>
</table>

Now, whether a question is MA or MS is entirely predictable from its Hamblin set. If the Hamblin set is
constructed in such a way as to necessarily contain a unique Strongest-True answer at any world compatible
with the presupposition, the question is MA. If the Hamblin set allows for multiple Strongest-True answers
at some world, the question is MS. Hence, this theory makes powerful, grammatically driven predictions
about the distribution of MA and MS.

5.1.3 Illustrating mention-all

Let us return to example (59) and see concretely how ANS$_2$ derives both MA and MS readings. An LF with
each above can was responsible for MA with ANS$_1$, and is likewise responsible for MA with ANS$_2$. The LF
from the previous subsection updates as (80), with the only change being that ANS$_1$ is swapped for ANS$_2$.
The Hamblin set to which ANS applies is the same as before, closed under conjunction.
(80) Updated first LF for (59) (each > can)

(71) Recall: corresponding Hamblin denotation

a. \([\text{CP}3]^w = \lambda p . \exists X [p = \forall x <_{AT} X [\Diamond [\text{chair}(x, \text{committee})]]]

b. \(\approx \{\Diamond [\text{chair}(\text{John, committee})],\)

c. \(\Diamond [\text{chair}(\text{Bill, committee})],\)

d. \(\Diamond [\text{chair}(\text{John, committee})] \land \Diamond [\text{chair}(\text{Bill, committee})], \ldots}\)

Consider the now familiar world, \(w_1\), at which John and Bill are eligible chairs. At this world, (71-b), (71-c), and (71-d) are all true, and only (71-d) qualifies as Strongest-True: (71-d) is not entailed by (71-b) or (71-c), but (71-b) and (71-c) are both entailed by (71-d). Hence, there is a unique Strongest-True proposition. As we saw in the previous section, ANS\(_1\) would have its presupposition satisfied and return (71-d), which the responder should thus provide to resolve the question. ANS\(_2\) reaches a similar end through minimally different means. Its weaker presupposition is also satisfied, and it returns the singleton set containing (71-d):

(81) Computing the ANS-set

a. \([\text{ANS}_2]^{w_1}(\text{[CP}_5]^{w_1})\)

b. \(= \{\Diamond [\text{chair}(\text{John, committee})] \land \Diamond [\text{chair}(\text{Bill, committee})]\}\)

The responder’s task is to provide some proposition in the ANS-set. Because the unique Strongest-True proposition is the only proposition in the ANS-set, the responder must provide that. ANS\(_2\), like ANS\(_1\), thus predicts the MA answer: ‘John and Bill can chair the committee.’. This discussion has offered a particular illustration of the key general point that a singleton ANS-set directly translates to a MA interpretation. Because the Hamblin set in (71) is closed under conjunction, it will necessarily contain a unique Strongest-True proposition at any world where the presupposition of the question is satisfied, and the ANS-set will always be a singleton. With each > can, (59) is a MA question by virtue of its Hamblin set.
5.1.4 Illustrating mention-some

Now, what about the MS reading? To derive MS, we need a different Hamblin set, and one with just the right make up follows from the second LF, where can takes wide scope over each. The LF is updated in (82) to replace ANS$_1$ with ANS$_2$, and the associated Hamblin set, crucially not closed under conjunction, is repeated below the structure.

(82) Alternative LF for (59) (can > each)

At $w_1$, there are now just two true propositions in the Hamblin set: (73-b) and (73-c). Since (73-b) is true and not entailed by (73-c), (73-b) qualifies as Strongest-True. In a similar vein, since (73-c) is true and not entailed by (73-b), (73-c) counts as Strongest-True, as well. This is a key result: there are two true propositions in the Hamblin set and both count as Strongest-True. With ANS$_1$, this state of affairs led to presupposition failure, due to the absence of a unique Strongest-True proposition. No such problem arises with ANS$_2$, however: since there is some Strongest-True proposition, the existential presupposition of ANS$_2$ is met. The ANS-set, computed in (83), now contains multiple members, which leads to MS.

(83) Computing the ANS-set

Now, the responder has a choice to make: they can provide either of the two propositions in (83-b). They can answer the question with 'John can chair the committee.' or with 'Bill can chair the committee.' and
both are complete answers — the MS answering pattern. In sum, the LF which led to presupposition failure with ANS1 is repurposed for a MS reading with ANS2. The Hamblin set is not closed under conjunction, so can contain multiple Strongest-True propositions, deriving MS.

5.1.5 Constraints on mention-some

As noted at the outset of this section, a strong argument for a grammatical theory of the MA vs. MS comes from the fact that not all questions have MS readings. In particular, MS is bled when an existential quantifier is absent. This came out, for instance, in the data point in (50-a):

(50-a) Recall: unambiguous simple question
Who addressed the committee?

a. “Tell me everyone who chaired the committee.”

b. **“Tell me someone who chaired the committee.”

ANS2 predicts the judgment. Since there is no modal for each to scopally interact with, there is just one available LF for this sentence, (84), which results in the Hamblin set in (85). This set is closed under conjunction and, thus, will furnish a unique Strongest-True proposition at any world compatible with the presupposition, deriving a MA reading.

(84) LF for (84)
... [CP4 ANS2 [CP3 PRO λ2 [who λ1 [[C t2] [IP [t1 each] [chair the committee]]]]]]

(85) Corresponding Hamblin denotation
a. \([CP_3]^w = λp . \exists X [p = \forall x <_{AT} X [chair(x, committee)]]\]

b. \(≈ \{ λw . John addressed in w,\)
c. \(λw . Bill addressed in w,\)
d. \(λw . John addressed in w \land Bill addressed in w, \ldots\)

For a concrete illustration, suppose that John and Bill address the committee at w1, but no one else did. All of (85-b), (85-c), and (85-d) are true at w1, and (85-d) is the only one that qualifies as Strongest-True, since (85-d) asymmetrically entails (85-b) and (85-c). ANS2 thus returns the singleton set in (86) and, in turn, the responder must answer the question by supplying (85-d): ‘John and Bill each addressed the committee.’ is the answer on order — the MA answer.

(86) Computing the ANS-set
a. \([ANS_2]^w = ([CP_3]^w)\]

b. \(= \{ λw . John addressed in w \land Bill addressed in w\]

With no other LF to consider, there is no way to coax a MS reading out of (50-a). Hence, ANS2 explains why the availability of MS correlates with the presence of a modal in the LF: the modal is required to yield an LF whose Hamblin set has the right make-up to possibly contain multiple Strongest-True propositions.
5.1.6 Local summary

In this subsection, I have introduced Fox's (2013) theory of the MA vs. MS distinction, developed on the basis of simple questions. The key idea is that whether a question is interpreted MA or MS is predictable from its Hamblin set: a MA question has a Hamblin set which will necessarily deliver a singleton ANS-set, while a MS question has a Hamblin set which can deliver an ANS-set with multiple elements. The stage is set for a return to disjunction.

5.2 Analyzing the disjunctions

I will now make the push at caching out the main hypothesis that question disjunctions are MS questions. Combined with the Union Hypothesis, the system with \( \text{ANS}_2 \) predicts that the disjunction data are MS questions, and the choice schema is immediately predicted. I present Profile 2 data first, then Profile 1, and finally I return to Ciardelli et al.'s (2015) data point.

5.2.1 Profile 2 disjunctions

Let us re-consider the familiar example in (27), the LF for which is now (87). \( \text{ANS}_2 \), having ATB-moved above the conjunction, operates on the Hamblin denotation in (88).

(27) **Recall: Profile 2 prototype**
Who's the mother or who's the father?

(87) **LF for (27)**
\[ \text{CP5} \ldots \text{ask} \text{CP4 ANS2 CP3PRO}\lambda 3 [\&P \text{CP2a who} \lambda 1 [\text{CP1a } \text{C t3} ] [\text{FP t1 is the mother}]] \]
\[ \text{and} \text{CP2b who} \lambda 2 [\text{CP1a } \text{C t3} ] [\text{FP t2 is the mother}] \]

(88) **Hamblin denotation**
\[ \lambda p. \exists x [p = \lambda w. \text{x is the mother in w}] \lor \exists y [p = \lambda w. \text{y is the father in w}] \]

Given the Union Hypothesis, the Hamblin set for (27) is constructed in such a way as to allow for a multi-membered ANS-set and, thus, is formally a MS question. To illustrate, suppose at \( w_1 \) that Mary is the mother and John is the father. Then, there are two true propositions in the Hamblin set, one of the form \( \text{that x is the mother} \) and one of the form \( \text{that y is the father} \), as in (89). Because those two true propositions are logically independent of one another, both qualify as Strongest-True. The existential presupposition of \( \text{ANS}_2 \) is met, and the predicted ANS-set is the doubleton in (90).

(89) **True propositions at \( w_1 \)**
\[ a. \lambda w. \text{Mary is the mother in w} \]
\[ b. \lambda w. \text{John is the father in w} \]

(90) **Computing ANS-set at \( w_2 \)**
\[ a. [\text{ANS}_2]^{w_2}([\text{who is the mother or who is the father}]^{w_2}) \]
\[ b. = \{\lambda w. \text{Mary is the mother in w, } \lambda w. \text{John is the father in w}\} \]

The choice schema follows straightaway as a predicted consequence. Any proposition in the ANS set constitutes a complete answer to the question. With the ANS-set in (90), therefore, the responder has a
choice to make. They can answer the question with ‘Mary is the mother’ or, equally well, they can answer with ‘John is the father’ — capturing the choice schema.

The result is very general. Suppose any schematic disjunction $Q_1$ or $Q_2$ where answers to $Q_1$ are logically independent of answers to $Q_2$. Because $[Q_1 \text{ or } Q_2]$ contains all the answers to $Q_1$ and all the answers to $Q_2$, it will contain multiple Strongest-True propositions at any world where $Q_1$ has a true answer and $Q_2$ has a true answer (as in the Profile 2 data). In the system with $\text{ANS}_2$, $Q_1$ or $Q_2$ is thus a MS question, answerable with any Strongest-True proposition $[Q_1 \text{ or } Q_2]$ contains.

5.2.2 Profile 1 disjunctions

Without any modifications, the system perfectly captures intuitions with the Profile 1 data, as well. Let us return to the disjunction in (91), discussed in Section 3.1 as Profile 1:

(44) Recall: Profile 1 prototype
Who got hired or who’s leaving.

(91) LF for (44)
$[CP_5 \ldots \text{ask} [CP_4 \text{ANS}_2 [CP_3 \text{PRO } \lambda 3 [\& P [CP_{2a} \text{who } \lambda 1 [CP_{1a} [C t_3 [TP t_1 \text{ got hired}]]]
[\text{and } [CP_{2b} \text{who } \lambda 2 [CP_{1a} [C t_3 [TP t_2 \text{ is leaving}]]]]]]]

(92) Hamblin denotation
$\lambda p . \exists x [p = \lambda w . x \text{ got hired in } w] \lor \exists y [p = \lambda w . y \text{ is leaving in } w]

In the context in which I presented (44), it was granted that there was just a single staffing change. In that context, (44) will have a unique answer. If the world is $w_1$, at which the staffing change was John being hired, the Hamblin set will contain the one true proposition in (93):

(93) True proposition in Hamblin set at $w_1$
$\lambda w . \text{John got hired in } w$

Being the only true proposition in the Hamblin set, qualifies as Strongest-True, and $\text{ANS}_2$ returns the singleton set in (94). The complete answer to the question is ‘John got hired.’.

(94) Computing ANS-set at $w_1$
a. $\llbracket \text{ANS}_2 \rrbracket^w (\llbracket \text{who got hired or who’s leaving} \rrbracket^w)$
b. $= \{ \lambda w . \text{John got hired in } w \}$

The choice schema as it manifests in Profile 1 is captured through world-by-world variability in the identity of the true answer, much in the same way as sketched in Section 4.2. If the world is $w_2$, at which John is leaving, the unique true proposition in the Hamblin set would be (95), rather than (93), and accordingly the ANS-set would be (96). The complete answer is, then, ‘John is leaving.’.

(95) True proposition in Hamblin set at $w_2$
$\lambda w . \text{John is leaving in } w$
Computing ANS-set at \( w_2 \)

a. \([\text{ANS}_2]''\{[\text{who got hired or who's leaving}]''\}

b. \( = \{\lambda w . \text{John is leaving in } w\} \)

Contextually granting that there was one staffing change means there is a unique Strongest-True proposition in the Hamblin set and, depending on what the facts happen to be, that will either be an answer \textit{that }x\textit{ got hired} or an answer \textit{that }y\textit{ is leaving} — capturing the choice schema. Profile 1 data and Profile 2 data are captured in a unified system.

With respect to its Hamblin set, (44), like (27), has the formal property of a MS question: the Hamblin set is constructed in such a way that it could contain multiple Strongest-True propositions. At a world where John got hired and Bill is leaving, for instance, the union in (92) would contain two Strongest-True propositions — \( \lambda w . \text{John got hired in } w, \lambda w . \text{John is leaving in } w \) — giving rise to a multi-membered ANS-set. In context, however, it is granted that the Hamblin set will contain only one Strongest-True proposition and, thus, (44) behaves more like a MA question. It has a singleton ANS-set and thus a unique complete answer. Perhaps the best way of conceptualizing (44) is as a MS question in a situation where there happens to be only one true answer.

5.2.3 Disjoining MS questions

At this point, I want to return to Ciardelli et al.'s (2015) example. I noted in Section 4 that this example has a somewhat different profile from the other data. I did not explain the difference there, but the background is now in place for an explanation.

(9) Adapted from Ciardelli et al. (2015)
Where might we rent a car, or who might led us one?

In the other disjunction data, the two disjuncts, if asked in isolation, are individually MA questions. In (27), for instance, \([\text{who is the mother}]\) necessarily contains a unique Strongest-True proposition, as does \([\text{who is the father}]\). After all, neither question contains an existential modal. We saw that disjoining these two MA questions manufactures a MS question. In Ciardelli et al.'s example, the individual disjuncts are each MS questions. I now show how that example fits in the system.

The effect of disjoining two MS questions is to manufacture a new MS question. Each of the disjuncts in (9) itself contains an existential modal and is most naturally interpreted as a MS question. The relevant Hamblin set for the question in the left disjunct is made up of propositions of the form \textit{it’s possible that we rent a car at every atom of }X,\textit{ for different entities }X.\textit{ Sample elements from this Hamblin set are:}

(97) \([\text{where might we rent a car}]''\)

a. \( \approx \{\Diamond [\text{we rent a car at Avis}]\}

b. \( \Diamond [\text{we rent a car at Hertz}]\)

c. \( \Diamond [\text{we rent a car at Avis } \land \text{ we rent a car at Hertz}]\}

Let us assume a world, \( w_1 \), at which Avis and Hertz are potential renting locations. Moreover, let us assume that we will certainly not rent at multiple places. Then, (97-a) and (97-b) are true, while (97-c) is false. Since (97-a) and (97-b) are not in an entailment relation, both qualify as Strongest-True. If the question
were asked in isolation, then, it would be a mention-some question with (97-a) and (97-b) as possible answers. The situation is parallel for the question in the second disjunct:

\begin{align*}
(98) & \quad \mathbf{[who \ might \ lend \ us \ a \ car]^w} \\
& \quad \text{a. } \approx \{\Diamond \ [\text{John lends us a car}] \\
& \quad \text{b. } \Diamond \ [\text{Fred lends us a car}] \\
& \quad \text{c. } \Diamond \ [\text{John lends us a car} \land \text{Fred lends us a car}] \}
\end{align*}

The Hamblin set is made up of propositions of the form it's possible that every atom of X lends us a car, for different entities X. If John and Fred are potential lenders and we will certainly not borrow from multiple people, the propositions in (98-a) and (98-b) are true, while (98-c) is false. (98-a) and (98-b) are again not in an entailment relation, and both qualify as Strongest-True, making the question mention-some. According to the Union Hypothesis, the Hamblin set for the disjunction is:

\begin{align*}
(99) & \quad \mathbf{[where \ might \ we \ rent \ a \ car, \ or \ who \ might \ lend \ us \ one]^w} \\
& \quad = \lambda p. \ [\text{where might we rent a car}]^w(p) \land [\text{who might lend us a car}]^w(p)
\end{align*}

Because the union characterized in (99) contains all elements of (97) and (98) together, there are four true propositions in the Hamblin set for the disjunction: (97-a), (97-b), (98-a), and (98-b). None of these propositions entails any other, so they all qualify as Strongest-True, and any one of them is an appropriate mention-some answer to the disjunction — a prediction clearly accordant with intuitions. When two MS questions, Q₁ and Q₂, are disjoined, the output is a new MS question answerable with any answer to Q₁ or any answer to Q₂.

5.2.4 Local summary

I showed how Fox’s (2013) system for analyzing simple MA and MS questions immediately captures question disjunctions as MS questions, when combined with the Union Hypothesis. The choice schema is predicted (as manifest in both Profile 1 and Profile 2 situations), and questions disjunctions are fit into the general typology of question types. Whether the individual disjuncts are themselves MA or MS questions, the disjunction is an internally complex MS question.

6 Taking into account the pragmatics

Although the distribution of MA and MS readings is grammatically conditioned, there are also still pragmatic constraints on when it is appropriate to utilize a question LF with a MS reading. In this section, I flag the pragmatic properties of simple MS questions, and demonstrate that disjunctions share the pragmatic signature of MS. For this discussion, I focus on the Profile 2 cases, which clearly function as MS in context. I show that taking into account the pragmatics can help explain: (i) further data which initially appear not to fit into the proposal; and (ii) why the question disjunctions considered in the literature prior to Ciardelli et al. (2015) were unacceptable.
6.1 The pragmatics of mention-some

As discussed earlier, Van Rooij (2003) suggests that speakers ask a question to resolve a “decision problem”. In his system, a MS question is acceptable only when a MS answer has equal utility to a MA answer to resolve the decision problem. While drawing inspiration from van Rooij, I re-formulate the generalization as (100) to fit better with a grammatical view of MS.

(100) Pragmatic condition on mention-some

A speaker should ask a MS question if the answer to a MS question would be of sufficient utility to resolve their decision problem.

First, consider the simple question in (59). I noted earlier that a MS reading would be natural in the context of a faculty meeting, where the speaker is looking for anyone willing to chair an unpleasant committee. In that context, the decision problem is to find some arbitrary committee chair. Knowing a single person who can serve as chair is sufficient to resolve the dilemma, and a MS answer is as useful as a MA answer. The MS LF would thus be appropriate.

(59) Simple question with mention-some reading

Who can chair the committee?

Problem: finding an arbitrary chair for a committee.

Pragmatic condition met: one chair is all that’s needed.

Extending to disjunctions, consider first Ciardelli et al.’s (2015) example. This datum clearly has the pragmatic signature of MS. The dilemma is how we can acquire a car, and knowing one way to acquire a car is sufficient to resolve this dilemma. All of the other Profile 2 data I introduce in Section 3.2 similarly have the pragmatic signature of MS, as sketched below.

(101) Ciardelli et al.’s (2015) example

Where might we rent a car, or who might have one we could borrow?

Problem: dilemma of where to find a car.

Pragmatic condition: one place to find a car is all that’s needed.

(102) The other data

a. Who’s the mother or who’s the father?

Problem: dilemma how to make a call to one parent.

Pragmatic condition: one parent’s information is sufficient.

b. What’s your name or what’s your SSN?

Problem: dilemma of how to identify the interlocutor.

Pragmatic condition: one identifier is sufficient.

c. What’s your superpower or what’s your spirit animal?

Problem: dilemma of how to begin a conversation.

Pragmatic condition: one piece of information is sufficient.
6.2 The pragmatics helps explain further data

At this point, I want to consider a puzzling new data point, not introduced in Section 3: example (103), which involves a disjunction of a constituent question and a polar question.

(103) **A new example**
Who can take care of the kids, or do you want to take them with us?

Consider first the polar question in isolation, the syntax for which is given in (105). I assume that polar questions involve a special interrogative complementizer, defined as (105), which yields the Hamblin denotation below, characterizing a set of propositions containing two members, *that we will take the kids* and its negation *that we will not take the kids*.

(104) **LF for polar question**

```
CP4
  ...
    ask
      CP3
        ANS2
          CP2
            PRO
              λ1
                CP1
                  C_{pol} t1
                    TP
                      you want to take the kids with us
```

(105) **Defining interrogative C**

\[
[C]^w = \lambda p_{st} \cdot \lambda q_{st} \cdot p = q \lor p = \neg q
\]

(106) **Hamblin denotation at CP3**

\[
[[CP_2]]^w = \lambda p \cdot \lambda w^* \cdot \forall w^* \in W(you)(w^*) \left [ \text{we take the kids in } w^* \right ] \\
\lor p = \forall w^* \in W(you)(w^*) \left [ \text{we do not take the kids in } w^* \right ]^{10}
\]

Given a treatment of polar questions, it is clear that the assumed syntax for questions attributes a well-formed structure to (103), parallel to the structures it assigns to disjunctions where both disjuncts are constituent questions. I take the LF for (103) to be (107), which predicts the Hamblin denotation in (108) for CP3. Accordant with the Union Hypothesis, the effect of disjunction is to union the Hamblin sets of the two disjoined questions. The set in (108) contains propositions *that X can take care of the kids* (i.e. elements of [[who can take care of the kids]]), and the propositions *that you want to take the kids with us* and *that you don't want to take the kids with us* (i.e. elements of [[do you want to take them with us]]).

---

10 *Want* is a neg-raising predicate in English, and I show the output of neg-raising.
Now, suppose a world, \( w_1 \), at which John can take care of the kids, but the interlocutor in fact wants to take them along. At \( w_1 \), the Hamblin set in (108) contains two true propositions: (109-a) and (109-b). Since these propositions are independent of one another, both qualify as Strongest-True, and the ANS-set is the doubleton in (110). The disjunction is a MS question, and both propositions in the ANS-set are predicted to be complete answers.

**True proposition in Hamblin set at \( w_1 \)**

a. \( \lambda w \cdot \text{John can take care of the kids in } w \)
b. \( \lambda w \cdot \forall w' \in W(you)(w) \text{ [we take the kids in } w'] \)

**Computing the ANS-set**

a. \( [\text{ANS}_2]^{w_1}([\text{CP}_3]^{w_1}) \)
b. \( = \{ \lambda w \cdot \text{John can take care of the kids in } w \} \)

Empirically, the prediction does not seem right. If the responder wants to take the kids, they must answer with (111-a) and it is infelicitous to answer with (111-b). Hence, the puzzle: in the scenario considered, (103) is predicted to have multiple MS answers, since the ANS-set contains multiple propositions — but in actual fact there is only one appropriate answer.
(111) **Actual answering pattern**

Who can take care of the kids, or do you want to take them with us?

a. I want to take them with us.

b. #John can take care of the kids.

Does (103) undermine the overall analysis? I think not. I suggest that (103) is a MS question with the ANS-set in (110), and that the more restricted answering pattern is a consequence of the pragmatics. In asking (103), the decision problem the speaker is attempting to resolve is the dilemma of what to do with the kids. (111-a) and (111-b) are both acceptable answers to (103), as both equally well resolve the decision problem. However, they resolve it in different ways — and the questioner will act based on how the responder’s answer resolved the decision problem. If the responder provides (111-b), the questioner will send the kids to John, and if the responder provides (111-a), the questioner will bring the kids. If the responder’s goal is to bring about the effect that they bring the kids, the responder must, then, answer (111-a). The example in (103) is understood by taking into consideration the pragmatics.

### 6.3 Why disjoining questions isn’t fully productive

This chapter has argued that disjoined questions are attested and analytically predicted. Yet, as discussed at the outset, a longstanding consensus in the literature held that disjoined questions do not exist. While the data presented in Section 3 conclusively demonstrate the existence of disjoined questions, it remains to be explained: why are disjoined questions not fully productive? In this regard, I agree with a suggestion in Ciardelli et al. (2015) that the answer lies in the pragmatics. Two of the degraded examples discussed in the literature are repeated in (112). These examples appear to be Profile 2, assuming that Mary and Judy both read something in (112-a), and that the interlocutor is both married and lived somewhere in (112-b).

(112) **Recall: unacceptable question disjunctions**

- a. What did Mary read or what did Judy read?
- b. Who did you marry or where did you live?

The problem is that these examples do not satisfy the pragmatic constraints on MS. As Ciardelli et al. (2013) state, “it is difficult to see what kind of motivation (or what kind of decision problem, to follow van Rooij 2003) a speaker could have that would lead her to raise or even consider the issue expressed by [these disjunctions].” Re-casting slightly in the present analysis, it is difficult to see what decision problem could be equally well addressed by an answer to either one of the two disjuncts in (112-a) or (112-b). A prediction of the account is that these examples should improve with enough contextual help.

Illustrating first with (112-a), suppose that you want to pick a new book to read. You know that Mary and Judy have excellent literary taste, and you would be happy to read anything they’ve read. Talking to Mary, you might felicitously say: “You and Judy have the best taste, and I want to read what one of you is reading. So, what you have been reading, or what’s Judy been reading?” This is a variant of (112-a).

The situation is similar with (112-b). Suppose that you work as an archivist, and someone comes to you looking for help in locating records on their great-grandfather. Records are organized in two ways: based on spousal relations, and by place of residence. Either piece of information is equally effective as the other to search the database. After explaining the database to your client, it seems felicitous to ask a variant of (112-b): ‘Who did he marry or where did he live?’
6.4 Local summary

In addition to having the defining semantic property of MS questions, question disjunctions behave pragmatically like MS questions: they are felicitous when a MS answer is sufficient to resolve some decision problem. The pragmatics reconciled (103) with the analysis developed in the preceding section and, following Ciardelli et al. (2015), explained why disjunction is not fully productive.

7 Alternative analyses

This chapter has pursued an analysis of question coordination with two key features: (i) question conjunction has and scope at a high type t node, above a covert performative layer, while (ii) the CP embedded by the performative verb has a Hamblin denotation computed in such a way that there is a lower type t node at which or can scope. I now consider alternative analyses to (i)-(ii) and argue they are insufficient.

7.1 Alternatives for conjunction

To handle conjunction without a covert performative, the overall analysis must change in some fundamental way. As I see it, there are two options, and both lead to pathological results with disjunction. I outline each in turn. Note that to illustrate both theories, I assume world arguments, rather than parameters.

7.1.1 Partition semantics for questions

The first potential alternative comes from Groenendijk & Stokhof ('G&S'; 1989). While I have assumed a Hamblin denotation for the interrogative CP throughout the dissertation, G&S have a completely different interrogative semantics, specifically that of G&S (1984), by which questions partition the space of possible worlds. To introduce the system, consider a simple question:

(113) A G&S simple question
   a. Who got hired?
   b. \([ ([113-a] = \lambda w . \lambda w' . \forall x [\text{be-hired}(x, w) \leftrightarrow \text{be-hired}(x, w')] ] \]

If we assume that there are just two salient people, John and Mary, (113-a) carves the space of possible worlds up into four non-overlapping subsets of worlds: (i) those worlds at which no one got hired, (ii) those at which only John got hired, (iii) those at which only Mary got hired, and (iv) those at which both got hired. By providing an answer to the question, the responder indicates in which cell the actual world is located. Formally, the questions denotes an equivalence relation between worlds, as in (113-b). Two worlds w and w' stand in the relation in (113-b) just in case the people who got hired at w are exactly the same as the people who got hired at w'. All the worlds within any one of cells (i)-(iv) stand in the relation with one another, while no two worlds in different cells stand in the relation. For a question to be well-formed for G&S, it is crucial that the question carve up the space of possible worlds into non-overlapping cells. They require that questions partition logical space, and mutual exclusivity of cells is one property of a partition. I state this property as (114), paraphrased: if w and w' are in one cell and w is not in the same cell as w", w' must not be in a cell with w" either.
Partition Condition
For a question Q to be well-formed, the relation R that Q denotes must be such that for any worlds w, w', and w", if R(w,w') & ¬R(w,w") → ¬R(w',w").

G&S can readily handle question conjunction without recourse to a performative, provided that and can type-lift to operate on relations between worlds. With and defined as (115), (116-a) denotes the equivalence relation in (116-b): two worlds w and w' stand in the relation just in case the same people got hired in w and w' and the same people are leaving in w and w'.

Type-lifted variant of and
a. \[ \text{and}_6 = \lambda R_{s, st} . \lambda R'_{s, st} . \lambda w . \lambda w' . \left[ \text{and}_6((R(w)(w'))(R'(w)(w'])) \right] \]
b. \[ = \lambda R_{s, st} . \lambda R'_{s, st} . \lambda w . \lambda w' . R(w)(w') \land R'(w)(w') \]

Illustrative question conjunction
a. Who got hired and who's leaving?
b. \[ \left[ (116-a) \right] = \lambda w . \lambda w' . \forall x \left[ \text{be-hired}(x, w) \Leftrightarrow \text{be-hired}(x, w') \right] \land \forall y \left[ \text{leave}(y, w) \Leftrightarrow \text{leave}(y, w') \right] \]

The relation in (116-b) induces a partition with 16 cells, graphically represented in (117). Cell 1 contains exactly those worlds where no one was hired and no one is leaving. Cell 2 contains exactly those worlds where no one was hired and only John is leaving. And so forth. No world can be in multiple cells, so the Partition Condition is satisfied. To answer the question, the responder must localize the actual world in a particular cell and, as such, they must indicate everyone who was hired and everyone was fired — the information that the question intuitively demands.

Partition induced by (116-a)

<table>
<thead>
<tr>
<th></th>
<th>no one hired</th>
<th>only John hired</th>
<th>only Mary hired</th>
<th>both hired</th>
</tr>
</thead>
<tbody>
<tr>
<td>no one leaving</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>only John leaving</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>only Mary leaving</td>
<td>3</td>
<td>7</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>both leaving</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

Despite this success for conjunction, disjunction does not receive a viable analysis. G&S themselves observed that disjoining equivalence relations does not yield a well-formed question, since the output does not satisfy the Partition Condition. Consider:

Type-lifted variant of or
a. \[ \text{or}_6 = \lambda R_{s, st} . \lambda R'_{s, st} . \lambda w . \lambda w' . \left[ \text{or}_6((R(w)(w'))(R'(w)(w'))) \right] \]
b. \[ = \lambda R_{s, st} . \lambda R'_{s, st} . \lambda w . \lambda w' . R(w)(w') \lor R'(w)(w') \]

Illustrative question disjunction
a. Who got hired or who's leaving?
b. \[ \left[ (119-a) \right] = \lambda w . \lambda w' . \forall x \left[ \text{be-hired}(x, w) \Leftrightarrow \text{be-hired}(x, w') \right] \lor \forall y \left[ \text{leave}(y, w) \Leftrightarrow \text{leave}(y, w') \right] \]
Two worlds \( w' \) and \( w'' \) stand in the relation in (119) just in case either the same people got hired in \( w' \) and \( w'' \), or the same people are leaving in \( w' \) and \( w'' \). To make clear that this relation does not conform to the Partition Condition, note the three worlds in (120):

(120) **Not a partition**

a. \( w_1 \): no one got hired, only John is leaving.
b. \( w_2 \): no one got hired, only Mary is leaving.
c. \( w_3 \): only John got hired, only Mary is leaving.

\( w_1 \) is in a cell with \( w_2 \) (because the same people got hired at \( w_1 \) and \( w_2 \)), but not \( w_3 \) (because different people got hired and are leaving at \( w_1 \) and \( w_3 \)). \( w_2 \), however, is in a cell with \( w_3 \) (because the same people are leaving in \( w_2 \) and \( w_3 \)). Hence, cells overlap: \( w_2 \) is in both a cell with \( w_1 \) and a cell with \( w_3 \), which are in different cells. Despite this problem with the Partition Condition, G&S (1989) were, as far as I know, the first to suggest empirically that question disjunction is possible. To accommodate disjunction within a partition-based theory, G&S proposed that questions are type-lifted to a generalized quantifier prior to being disjoined. I illustrate first with an embedded question (following exposition in Szabolcsi 1997, 2016):

(121) **Embedded question disjunction**

Sue found out \( [or\_p \text{ who got hired or who's leaving}] \).

The two disjuncts denote relations which, when applied to the actual world, output the propositions in (122-a) (*that the people who got hired are the ones who actually got hired*) and (123-a) (*that the people who are leaving are the ones who are actually leaving*), respectively. These propositions are Montague lifted, as in (122-b) and (123-b).

(122) **Montague-lifting the first disjunct**

a. \( \lambda w'. \forall x [\text{be-hired}(x, w_0) \leftrightarrow \text{be-hired}(x, w')] \)
b. \( \lambda P_{stt} . P(\lambda w'. \forall x [\text{be-hired}(x, w_0) \leftrightarrow \text{be-hired}(x, w')]) \)

(123) **Montague-lifting the second disjunct**

a. \( \lambda w''. \forall y [\text{leave}(y, w_0) \leftrightarrow \text{leave}(y, w'')] \)
b. \( \lambda P_{stt} . P(\lambda w''. \forall y [\text{leave}(y, w_0) \leftrightarrow \text{leave}(y, w'')]) \)

The resultant meanings are disjoined with a different higher-type meaning for *or*. The output, as in (125), is a set of sets of propositions. A set of propositions is an element of the set just in case it contains either the proposition in (122-a) or the proposition in (123-a) (or both).

(124) **Type-lifted variant of *or***

a. \( [or_7] = \lambda Q_{stt, t} . \lambda Q'_{stt, t} . \lambda P_{stt} . [(Q(P))(Q'(P))] \)
b. \( = \lambda Q_{stt, t} . \lambda Q'_{stt, t} . \lambda P_{stt} . Q(P) \lor Q'(P) \)

(125) **Output of disjunction**

\( [orP] = \lambda P_{stt} . P(\lambda w'. \forall x [\text{live}(you, x, w_0) \leftrightarrow \text{live}(you, x, w')] \)
\( \lor P(\lambda w''. \forall x [\text{marry}(you, y, w_0) \leftrightarrow \text{marry}(you, y, w'')]) \)
The meaning in (125) is well suited to compose in an embedded environment such as (121). The full LF for (121) could be (126), where the disjunction QRs, leaving a trace of proposition type, as in (126). Found out denotes a relation between an individual and a proposition. Predicate Abstraction in (126) creates the set of propositions that Sue found out, which is the argument of (125). The sentence in (121) is correctly predicted to be true just in case either Sue found out that the people who came in fact came, or Sue found out the people who are leaving are in fact leaving.

(126) **Full LF for (121)**

\[
[[\text{orp who got hired or who's leaving}] \lambda l [\text{Sue found out }t_l]]
\]

The effect of type-lifting is to require the question disjunction to be the argument of some predicate of propositions. This is possible with embedding, but does not extend to a matrix environment, since there is no appropriate predicate. Szabolcsi (1997) first noticed this issue: "[matrix interrogatives] are genuine questions. Thus it seems natural to interpret them in a way that directly links them to possible answers, which is what the unlifted interpretations do; and it is not natural to interpret them as lifted questions" (p. 324).

Given G&S's heavy reliance on type-lifting, their approach is clearly incompatible with the SIH. The SIH would thus lead us to expect that their analysis is not the correct one — and, indeed, it has proven insufficient on its own terms: it offers no viable treatment of matrix question disjunction. Their analysis accords with the consensus view that question disjunction is unattested, but not with the observation of genuine question disjunction in Ciardelli et al. (2015) and the present chapter. I take it that the Hamblin denotation for the interrogative CP is not negotiable.

### 7.1.2 Pointwise Functional Application

The only other way I can think of to accommodate conjunction without a performative is to change how the Hamblin denotation is derived compositionally and combine that with a particular entry for and. Rather than adopting a Karttunen-inspired composition, let us consider the composition that Hamblin himself proposed. Hamblin assumed that every node in the structure denotes a set (not the characteristic function for a set). In the simple declarative in (127-a), the DP John denotes the singleton set in (127-a) and the VP got hired denotes the singleton set in (127-b). Each denotes the singleton set containing their familiar denotation.

(127) **Simple declarative**

John got hired.

a. \[[\text{John}] = \{\text{John}\}\]

b. \[[\text{got hired}] = \{\lambda x . \lambda w . x \text{ got hired in } w\}\]

Hamblin proposes a composition principle similar to Functional Application, but re-defined for two sets to compose. The principle is **Pointwise Functional Application.** A set \[[\beta]\] containing elements of type \(\sigma\) composes with a set \[[\gamma]\] containing elements of type \(<\sigma, \tau>\) to output a new set \[[\alpha]\] containing elements of type \(\tau\). The elements of \[[\alpha]\] are those which would result from composing every element of \[[\gamma]\] with every element of \[[\beta]\], each in turn, via regular Functional Application.
(128) **Pointwise Functional Application**

If $\alpha$ is a branching node with daughters $\beta$ and $\gamma$, and $[\beta]^s \subseteq D_\sigma$ and $[\gamma]^s \subseteq D_{<\sigma,\tau>}$, then $[\alpha]^s = \{a \in D_\tau : \exists b \exists c \ [b \in [\beta]^s \land c \in [\gamma]^s \land a = c(b)]\}$

In (127), $[\text{John}]$ and $[\text{got hired}]$ each contains just one element, so $[\text{got hired}](\text{[John]})$ also contains just one element — the proposition *that John got hired*.

(129) **Interpreting (127)**

a. $[(127)] = [\text{got hired}](\text{[John]})$

b. $= \{[\lambda x \cdot \lambda w \cdot x \text{ got hired in } w](\text{John})\}$

c. $= \{\lambda w \cdot \text{John got hired in } w\}$

In questions, Pointwise FA is put to more substantive work. Hamblin proposed that the *wh*-word *who* introduces a set of all salient individuals. The ingredients of the simple question are interpreted as in (130-a) and (130-b). Because $[\text{who}]$ contains multiple members, so does $[\text{got hired}](\text{[who]})$. Pointwise FA yields a set containing the proposition that results from applying $\lambda x \cdot \lambda w \cdot x \text{ got hired in } w$ to John, and the proposition that results from applying $\lambda x \cdot \lambda w \cdot x \text{ got hired in } w$ to Mary, and so forth, as in (131). In this way, the alternatives introduced by *who* propagate up the tree to derive a set of propositions of the form *that*)

(130) **Simple question**

Who got hired?

a. $[\text{who}] = \{\text{John, Mary, } \ldots \}$

b. $[\text{got hired}] = \{\lambda x \cdot \lambda w \cdot x \text{ got hired in } w\}$

(131) **Interpreting (130)**

a. $[\text{got hired}](\text{[who]})$

b. $= \{[\lambda x \cdot \lambda w \cdot x \text{ got hired in } w](\text{John}), [\lambda x \cdot \lambda w \cdot x \text{ got hired in } w](\text{Mary}), \ldots\}$

c. $= \{\lambda w \cdot \text{John got hired in } w, \lambda w \cdot \text{Mary got hired in } w, \ldots\}$

This system offers a new way to analyze question conjunction, provided that *and* is assigned the entry in (132): *and* denotes the singleton set containing a propositional conjunction operator. If this is the sole meaning for *and*, this is broadly consistent with the SIH, since this *and* would necessarily take propositional scope. I refer to it as *andhs* (i.e. *and* in Hamblin semantics).

(132) **Hypothetical entry for and**

$[\text{andhs}] = \{\lambda p_{st} \cdot \lambda q_{st} \cdot \lambda w \cdot p(w) \land q(w)\}$

The two conjuncts in (133) would each denote sets of propositions, and $[\text{andhs}]$ could compose with those sets of propositions via Pointwise Functional Application to create a new set of propositions, (134), where each element of the new set is a conjunction.
(133) **Question conjunction**
Who got hired and who's leaving?
   a. \([\text{who got hired}] = \{\lambda w . \text{John got hired in } w, \ldots\}\)
   b. \([\text{who's leaving}] = \{\lambda w . \text{John is leaving in } w, \ldots\}\)

(134) **Output of Pointwise FA**
\([\text{and}_{hs}]([\text{who got hired}])([\text{who's leaving}])\)
   a. \(\{\lambda w . \text{John got hired in } w \land \text{John is leaving in } w, \ldots\}\)
   b. \(\lambda w . \text{John got hired in } w \land \text{Mary is leaving in } w, \ldots\)\]
   c. \(\lambda w . \text{Mary got hired in } w \land \text{John is leaving in } w, \ldots\)\]
   d. \(\lambda w . \text{Mary got hired in } w \land \text{Mary is leaving in } w, \ldots\)\]

Since a question conjunction does solicit a conjunctive answer, the set of propositions in (134) contains the possible answers to the conjunction in (133). Concretely, at a world, \(w_1\), where John got hired and Mary is leaving, \(\text{ANS}_2\) applies to (134) to return the singleton set containing (134-b). The responder must, then, answer the question with 'John got hired and Mary is leaving.' — which is exactly the intuitively required answer at \(w_1\).

(135) **Computing ANS-set**
   a. \([\text{ANS}_2]([\text{who got hired and who's leaving}]) (w_1)\)
   b. \(\{\lambda w . \text{Mary got hired in } w \land \text{Mary is leaving in } w\}\)

In Section 2, I proposed that (133) is a conjunction of two performative clauses: the first asks the question 'Who got hired?' and the second asks the question 'Who's leaving?'. Since both questions are asked, the responder must answer each, i.e. they must supply a conjunctive answer that \(x\) got hired and \(y\) is leaving. Pointwise composition with \([\text{and}_{hs}]\) has derived a single question, with the Hamblin set containing conjunctions of the form that \(x\) got hired and \(y\) is leaving. In this way, the same answering pattern is predicted without a covert performative.

Once again, the rub is disjunction. Assuming that \(\text{and}\) and \(\text{or}\) have parallel lexical entries, \([\text{and}_{hs}]\) leads us to expect the entry for \(\text{or}\) in (136): the singleton set containing a meaning for \(\text{or}\) which operates on propositions. This \(\text{or}\) has pathological consequences.

(136) **Hypothetical entry for or**
\([\text{or}_{hs}] = \{\lambda p_{st} . \lambda q_{st} . \lambda w . p(w) \lor q(w)\}\)

Just as \([\text{and}_{hs}]\) pointwise composed with two Hamblin sets in a question conjunction, \([\text{or}_{hs}]\) should be able to do the same in a question disjunction. For the question in (137), this would result in a set of propositions, each of which is a disjunction.

(137) **Question disjunction**
Who got hired or who's leaving?
   a. \([\text{who got hired}] = \{\lambda w . \text{John got hired in } w, \ldots\}\)
   b. \([\text{who's leaving}] = \{\lambda w . \text{John is leaving in } w, \ldots\}\)
Output of Pointwise FA

\[ \text{[or}_h] \text{([who got hired])([who's leaving])} \]

- a. \( \lambda w. \text{John got hired in } w \lor \text{John is leaving in } w \)
- b. \( \lambda w. \text{John got hired in } w \lor \text{Mary is leaving in } w \)
- c. \( \lambda w. \text{Mary got hired in } w \lor \text{John is leaving in } w \)
- d. \( \lambda w. \text{Mary got hired in } w \lor \text{Mary is leaving in } w \}

By the choice schema, the responder should provide either an answer of the form that x got hired or an answer of the form that y is leaving. Yet, the set in (139) does not contain propositions of those forms: it contains proposition of the form that x got hired or y is leaving, building the expectation that the responder should answer with a disjunction. To illustrate the prediction in one of its most curious forms, I assume that John and Mary are the only two salient individuals, and single out the following elements of (138):

Consider three elements of the Hamblin set

- a. \( \lambda w. \text{John got hired in } w \lor \text{Mary is leaving in } w \)
- b. \( \lambda w. \text{John got hired in } w \lor \text{John is leaving in } w \)
- c. \( \lambda w. \text{John got hired in } w \lor \text{John and Mary are leaving in } w \)

At a world, \( w_2 \), where John got hired and no one is leaving, these three propositions are exactly those in (138) that are true. Since (139-c) asymmetrically (139-a) and (139-b), only (139-c) qualifies as Strongest-True, and \( \text{ANS}_2 \) would return the singleton set containing (139-c). As such, the prediction is that the answer to the question should be 'John got hired or John and Mary are leaving.' — which clearly is not right. At \( w_2 \), the answer is intuitively just ‘John got hired.’

Computing ANS-set

- a. \( \text{[ANS}_2\text{]}\text{([who got hired or who's leaving])}(w_2) \)
- b. \( \{\lambda w. \text{John got hired in } w \lor \text{John and Mary are leaving in } w\} \)

Assuming that conjunction and disjunction have minimally different lexical entries, this problem with disjunction is sufficient grounds to reject the approach to conjunction based on Pointwise Functional Application with \[ \text{[and}_h] \]. Note that the syntax and semantics for questions I adopted in Section 2 not only generates the correct denotation for a disjunction, but, moreover, it cannot generate the pathological meaning in (139). Let us reflect on the characteristic function for (139), noting the scope of the existential quantifiers and the disjunction relative to the identity operator: the existentials scope above the identity operator, while the disjunction scopes below.

Characteristic function for (139)

\[ \lambda p. \exists x \exists y [p = \lambda w. x \text{ got hired in } w \lor y \text{ is leaving in } w] \]

In the syntax-semantics mapping I presented, the two existential quantifiers come from two wh-words, and the ‘=’ operator comes from the interrogative complementizer. As such, to derive the function in (141), there would have to be two wh-words above the complementizer, while the disjunction scoped lower, inside the TP. This would implicate a structure like (142) — which is not syntactically viable, since movement of the two wh-words out of their respective conjuncts violates the Coordinate Structure Constraint. The ungrammaticality of the corresponding surface string (*‘Who who got hired or is leaving?’) corroborates.
Overall, I attempted to do away with the covert performative in two ways — by changing the overall semantics for questions, and by changing the internal composition of the Hamblin denotation to invoke Pointwise Functional Application.\textsuperscript{11} Both attempts proved unsuccessful and, as far as I can see, there are no other options to try in this service.

7.2 Alternative for disjunction

According to the Union Hypothesis, a question disjunction is interpreted as a single question. This was achieved by scoping or at a type t node below the Hamblin denoting CP. In this section, I consider an alternative analysis, which modifies the performative layer to create an appropriate higher scope site for or, as well as and. Illustrating here with the Profile 2 prototype in (27), I noted earlier in the chapter that or cannot scope above ask in the LF in (19) without creating failed ignorance inferences. Performative clauses express trivial truths; yet, the disjunction requires that the speaker be ignorant about the truth of the two performative disjuncts.

(27) Recall: Profile 2 prototype
Who’s the mother or who’s the father?

\textsuperscript{11}Note that Haida & Repp (2013) proposed an analysis with Pointwise Functional Application, with a different assumption about the question meaning. They took a question to denote the set of its true answers (Karttunen 1977), rather than the set of its possible answers (Hamblin 1973). While their approach could again handle conjunction, it still falls short of handling disjunction. In fact, Haida & Repp explicitly accepted the earlier intuition that matrix question disjunction is unavailable, and developed their proposal so as to block it. If questions denote a set of true propositions and disjunction is pointwise, the denotation for $Q_1$ or $Q_2$ is a set of disjunctions, each constructed in such a way that both disjuncts are necessarily true. Since a disjunction is in general infelicitous when both disjuncts are known to be true, Haida Repp take it that a matrix question disjunction is not answerable, and blocked on that basis. Given the existence of matrix question disjunction, Haida Repp’s analysis cannot be maintained.
Recall: unviable LF

\[\text{[[CP I hereby ask you who the mother is] [or [CP I hereby ask you who the father is]]]}\]

It is, however, possible that the performative layer does not simply comprise *ask*, but rather has a more articulated structure with multiple scope sites. For instance, the covert material could be *I hereby ask you to tell me*. The sentence in (27) could, then, have the same structure as (143), where disjunction scopes between the two levels of the performative layer (*ask* \(\lor\) *tell*). (143) is a single performative speech act and is fully acceptable, just like (27).

Potential decomposed performative

I hereby ask you [PRO to tell me who the mother is] or [PRO to tell me who the father is].

This idea gains some initial plausibility from the fact that a decomposition along these lines has been independently proposed in Sauerland (2009) and Sauerland & Yatsushiro (2014, S&Y). They discuss questions such as (144), with the operator *again*. Their focus is on counterpart data in German and Japanese, but the English example is sufficient to make the point here.

Question with *again*

What is your name again?

They observe that *again* does not necessarily presuppose that the speaker has asked the question before, nor does it necessarily presuppose that the interlocutor has told the questioner their name before. (144) allows a “remind-me” reading, paraphrased: “I used to know your name (having learned it one way or another), but now I forget. Could you remind me what your name is?” (adapted from S&Y p. 2). To derive the remind-me reading, S&Y propose that *again* scopes at an intermediate level of a complex performative layer, informally paraphrased as (145) (see §3 in S&Y for formal details). The presupposition *again* introduces in (145) is simply that the interlocutor’s name was previously known to the questioner.

S&Y’s decomposed performative

I hereby ask you to make it again known to me what your name is.

If disjunction can scope at the same level as where *again* scopes in (145), (27) could be captured, though the paraphrase is slightly different from (143):

S&Y-inspired analysis of (27)

I hereby ask you to make it [known to me ANS who the mother is] or [known to me ANS who the father is].

Nonetheless, I am skeptical of this approach to question disjunction for two reasons. First, only a small set of operators seems able to scope at intermediate levels of the performative layer. The additive focus particle *too*, for instance, cannot take intermediate scope, despite being ostensibly very similar to *again*: (147) is not licensed by someone other than the interlocutor having announced their name, or by the speaker previously knowing the birth date of the interlocutor.

Replacing *again* with *to*

#What is your name too?
As far as Y&S can tell, it seems only particles like *again* that trigger a repetitive presupposition can take intermediate scope in English and German. In Japanese, even those particles can't: intermediate scope seems to be restricted to a special suffix *kke* whose unique function is to trigger remind-me presuppositions. Thus, while I cannot demonstrate that disjunction is unable to take intermediate scope, it seems unlikely, given how the vast majority of other operators behave.

One way to more concretely dissociate my proposal from a performative theory is through presupposition. To do this, it will be useful to consider the Profile 1 prototype, assuming as before that there was just one staffing change:

(44) **Recall: Profile 1 prototype**

Who got hired or who's leaving?

In my proposal, (44) is a single question, and its LF contains a single ANS$_2$ operator, as in (148-a). ANS$_2$ introduces the presupposition that the union of the Hamblin sets for the two disjuncts contains some Strongest-True element. In other words, (44) should presuppose that someone got hired or someone's leaving, as in (148-b). If there was any staffing change, that presupposition is met.

(148) **Proposal: predicted presupposition**

a. [... ANS$_2$ [who got hired or who's leaving]]

b. Someone got hired or someone's leaving.

In a performative analysis, there are two questions, and each should come with a separate ANS operator, as in (149-a). ANS in the left conjunct triggers the presupposition that someone got hired, and ANS in the right conjunct triggers the presupposition that someone is leaving. In general, presuppositions project out of disjuncts, (perhaps) unless they are entailed by the negation of the other disjunct. The first disjunct in (149-a) (that you make it known who got hired) does not entail the presupposition of the second disjunct (that someone is leaving), or vice versa. Accordingly, both presuppositions should project, and (149-a), should overall presuppose (149-b).

(149) **Alternative: predicted presupposition**

a. [I hereby ask you to make it [known to me ANS who got hired]
   [or [known to me ANS who's leaving]]]

b. Someone got hired and someone's leaving.

Unless one of the presuppositions is suspended (or 'locally accommodated'; Heim 1983), the Profile 1 context should result in presupposition failure. If there was just a single staffing change, the conjunction in (149-b) is false. To test empirically, I will use *tell* in place of *I hereby ask you to make it known to me*:

(150) **Testing for presupposition failure**

The department finally made a staffing change!

a. Really! ??Tell me who got hired or tell me who's leaving.

b. Really! Who got hired or who got fired?

Consistent with the prediction, there does seem to be a squeamishness felt in (150-a), since it, unlike (150-b), most naturally requires that there was both a hiring and a departure. Although I believe further work is
needed to test the performative analysis sketched in this section more fully, I conclude that the proposal is more successful, as far as I can tell from the available evidence.

7.3 Local summary

In this section, I have presented alternative possible analyses of question conjunction and disjunction, and concluded that the original analysis achieves the greatest success: and scopes above a covert performative verb, while or scopes lower, internal to the interrogative CP itself.

8 Conclusion

This chapter began with the puzzle of question coordination, which is a prima facie problem for the hypothesis that and is always interpreted as [and, since questions do not obviously have type t meanings. I proposed an LF for questions which includes a performative prefix above which and scopes. The question CP itself has a Hamblin denotation, which is derived in such a way that there is a lower type t node internal to the CP, at which or scopes.

The proposal for or set up the second part of the chapter, where I provided a full analysis of question disjunction. The LF for disjunction predicted that Q1 or Q2 should be interpreted as a single question, with its Hamblin set the union of the Hamblin sets for Q1 and Q2. By combining that result with Fox’s (2013) theory of mention-some questions, I proposed that Q1 or Q2 is interpreted as mention-some. The proposal accounted for intuitions about how disjunctions are answered for a range of data, and fit disjunction into the broader interrogative typology.

9 A program for further inquiry

This is the last chapter of the thesis addressing conjunction. I have developed, and empirically defended, analyses of three constructions — apparent object DP conjunction, conjunctive pivots in pseudo-clefts, and questions — compatible with the SIH. Of course, this barely scratches the surface of cases where and appears to type-lift. Testing for covert syntax in other environments is a broad research program. In the following, I offer speculative remarks about what covert syntax might look like in some additional cases, as a stimulus for further inquiry. I will only discuss English data, but there is are further questions as to what strategies different languages use to create the necessary underlying syntax to host and how the different structures are mapped to phonology.

9.1 On verb conjunction

The covert syntax seen in this chapter already reconciles certain further data with the SIH. Consider, in particular, the intransitive verb conjunction in (151), repeated from Chapter 1:

(151) Intransitive V conjunction
    John laughed and danced

While this example might appear to involve [and2] directly operating on the <e,t> predicates laughed and danced, it could just as easily be parsed as vP conjunction, as in (152). Given this syntax, (151) has a parallel
structure to the one I have proposed to be underlying for apparent object DP conjunction. Due to the covert traces associated with the vP Internal Subject Hypothesis, the conjuncts are of type t and and is interpreted as [and] — accordant with the SIH.

(152) Parse as vP conjunction

\[ [TP \text{John}_1 \ [\&P \ [vP \text{t}_1 \text{ laughed}] \ [\&P \ [vP \text{t}_1 \text{ danced}] ]] \]

It might be possible to analyze apparent transitive verb conjunction as vP conjunction, as well. Let us consider the example in (153), and suppose the underlying vP conjunction structure in (154).

(153) Transitive V conjunction

John hugged and petted the dog.

(154) Parse as vP conjunction

\[ [TP \text{John}_1 \ [\&P \ [vP \text{t}_1 \text{ hugged the dog}] \ [\&P \ [vP \text{t}_1 \text{ petted the dog}] ]] \]

Observe that the object, the dog, is shared between the two conjuncts, and occurs linearly at their right edge. There is a known mechanism — Right Node Raising (‘RNR’) — by which rightmost shared material may be pronounced once, at the end of the sentence. Example (155) illustrates:

(155) Right Node Raising

John hugged and Mary petted the dog.

There is disagreement about the syntax of RNR: whether it involves ellipsis (e.g. Wexler & Culicover 1980, Kayne 1994, Hartmann 2000, Ha 2008), rightward movement (e.g. Ross 1967, Sabbagh 2007), multi-dominance (e.g. McCawley 1982, Bachrach & Katzir 2007, 2009), or some combination of different mechanisms (Barros & Vicente 2011, Hirsch & Wagner 2015). Regardless, whatever mechanism is available in (155) should also be able to apply in (153). The final structure for (153), assuming rightward movement, would be (156).

(156) Structure after RNR

\[ [TP [TP \text{John}_1 \ [\&P \ [vP \text{t}_1 \text{ hugged the dog}] \ [\&P \ [vP \text{t}_1 \text{ petted the dog}] ]] \ [DP \text{the dog}_2] ] \]

Given that vP conjunction and RNR are independently attested, it might be surprising if a parse like (156) were not at least available. I leave testing for this syntax to the future.

9.2 On subject DP conjunction

The arguments in Chapter 3 that conjunction of object DPs is best analyzed with CR, raises the question: what are the range of CR mechanisms that create the appearance of DP conjunction across its distribution? In this subsection, I speculate on the analysis of subject conjunction, the most obvious next case to consider. Consider the test example:

(157) Subject DP conjunction

Every student and every professor came.
This example could again involve an RNR derivation. The underlying syntax would be the TP conjunction in (158-a), with a verbal projection containing *came* then undergoing RNR to result in a single pronounced occurrence of *came* at the left edge.

(158) **Possible derivation**
   a. \[&p \; [TP \; every \; student \; [vP \; came]] \; [and \; [TP \; every \; student \; [vP \; came]]] \]
   b. \[[[&p \; [TP \; every \; student \; t1] \; [and \; [TP \; every \; student \; t1]]] \; [vP \; came1]] \]

Example (159) provides independent evidence that RNR is available for verbal projections, again making the suggested derivation epiphenomenal. In (159), *came* is associated with both conjuncts, but appears once, at the end of the sentence.

(159) **Baseline RNR**
    John said that Bill and Mary said that Sue came.

A natural way of testing for an RNR derivation for (157) would involve examining subject-verb agreement. In English, RNR generally allows for either singular or plural agreement (Grosz 2015). That pattern seems to be matched with apparent conjunction of certain subject quantifiers, notably universals. Speakers I consulted accept both *is* and *are* in (160).

(160) **Agreement with every and every**
    Every student and every professor *is/are* coming to the party.

Existential quantifiers, however, seem to require plural agreement. The minimally different example in (161) is accepted with *are*, while *is* is sharply ungrammatical.

(161) **Agreement with a and a**
    A student and a professor *are/*is coming to the party.

I am uncertain at present what to make of these agreement facts. These data — as well as agreement with apparent conjunction of subject quantifiers more generally — merit study.

### 9.3 On conjunction within DPs

At this point, I want to comment briefly on how conjunction within DPs fits with the SIH. In a sentence like (162), *and* appears to scope within the DP:

(162) **Conjunction within DPs**
    John saw the linguist and philosopher at the party.

This sentence has at least an available reading which presupposes that there was a unique person at the party who was both in linguistics and philosophy (i.e. someone like Bob Stalnaker), and asserts that John saw that person. This reading would not derive from a vP-level CR structure. Setting aside questions about how exactly ellipsis would proceed, an underlying vP structure would necessarily contain two object DPs, attributing to (162) the syntax in (163)
Hypothetical CR structure

\[ TP \text{John} [\text{[} vP \text{t saw the linguist] [and [} vP \text{t saw the philosopher]}]] \]

One way to distinguish (162) from (163) is via presupposition. The presupposition of (162) would be met if Stalnaker were the only linguist-philosopher at the party, even if the party was well stocked with phonologists and metaphysicists. In the same scenario, the presupposition of (163) would fail, since (163) presupposes that there is a unique linguist (from the first DP) and a unique philosopher (from the second). It must be possible to interpret and below the definite in (162).

Given the most obvious syntax of the DP, scoping and below the definite necessitates a higher-type and, in particular [and2], the variant which composes with <e,t> predicates. The structure would be as in (164), with and directly conjoining the NPs linguist and philosopher.

Most obvious syntax for DP

\[ DP \text{the} [\text{[NP linguist] [and [NP philosopher]]]}] \]

Pursuing the SIH leads to a more different analysis with additional syntax to create a type t node. One first hypothesis about the nature of this additional syntax might be to replace (164) with (165):

Syntax compatible with the SIH

\[ DP \text{the} [\text{NP3 PRO}1 [\&P [\text{NP1 t linguist} [\text{and [NP2 t philosopher]}]]]}] \]

Here, a covert pronoun is introduced into the structure. It starts off as a predicate-internal argument within the NPs linguist and philosopher, and moves across-the-board to a position just above the conjunction. The trace left with movement saturates the argument of linguist in the left conjunct and, likewise, saturates the argument of philosopher in the right conjunct to create type t meanings for the NPs. And scopes at this type t level, and the traces are then abstracted back over, once the binder index is reached just above and. PRO itself is not interpreted. Its sole role is to move in order to create the traces and trigger abstraction. This structure leads to the right result:

Interpreting the structure

a. \[ [\&P]^{w,s} = 1 \text{ iff } g(1) \text{ is a linguist in } w \land g(1) \text{ is a philosopher in } w \]
b. \[ [\text{NP3}]^w = \lambda x . x \text{ is a linguist in } w \land x \text{ is a philosopher in } w \]
c. \[ [\text{DP}]^w = \lambda x [x \text{ is linguist in } w \land x \text{ is a philosopher in } w] \text{ (if defined)} \]

The “PRO-strategy” in (165) was previously adopted in Heim & Kratzer (1998) to create scope positions within PPs. There are, however, complications. If the PRO-strategy is available, over-generation problems arise of a similar sort to what we saw with type-ambiguity. Consider again the apparent conjunction of DPs in (167), repeated from Chapter 3. I argued that scope freezing in this example is most easily understood if it lacks a parse where the quantifiers are directly conjoined. Yet, if the PRO-strategy is available, a maid and a cook should have as a possible structure (167).

Recall: scope freezing

Some company hired a maid and a cook. \( (\text{some} > \text{and}, *\text{and} > \text{some}) \)

Possible structure with PRO strategy

\[ DP3 \text{PRO} \lambda 1 [\&P [DP1 t1 [DP1 a maid] [and [DP2 t1 [DP2 a cook]]]]] \]
In (168), PRO starts off within the DPs, which are directly conjoined. Much as in (165), PRO ATB-moves out of the two DPs to adjoin above the conjunction. This leaves traces internal to the DPs. If these traces could be interpreted as variables of type &lt;e,t&gt;, rather than type e, they are of the right type to saturate the &lt;e,t&gt; argument of a maid and a cook, creating truth-values. &lt;[and]&gt; could then operate on those truth-values. Abstracting back over the traces just above where and integrates would re-create a quantifier type. As sketched in (169), this yields exactly the same result as if a maid and a cook were straightaway coordinated with &lt;[and3]&gt;. The PRO-strategy re-creates in the syntax what type-ambiguity does in the semantics.

(169) Interpreting the structure
   a. [DP &lt;p [NP the linguist] [and [NP the philosopher]]]]
   b. [DP the &lt;[&lt;f,x,y;[g]]&gt; &lt;x &lt;[g(y) &amp; x]] [exists y [cook(y) &amp; x]] [exists y [cook(y) &amp; x]]]

Now, we may be able to envision a reason to allow the LF in (165), but disallow the one in (168). Two possible options: (i) traces can be of type e, but not of type &lt;e,t&gt;, or (ii) DPs, unlike NPs, do not have a syntactic argument position available for PRO to fill. Neither is obvious: (i) wh-movement in (170) seems to leave an &lt;e,t&gt; trace, and (ii) possessives, as in (171), seem to require an argument in spec-DP, assuming that 's is a D head.

(170) Apparent &lt;e,t&gt; trace
   a. John is happy.
   b. How &lt;[f,t]t;1?&gt; is John &lt;[f,t]t;1?

(171) Apparent argument in spec-DP
   John's dog

There may be ways to make (i) or (ii) go through (see e.g. Poole 2017, who argues that, despite appearances, &lt;e,t&gt; traces are not attested; also Chierchia 1984, Landman 2006). Nonetheless, I want to suggest that we do away with the PRO-strategy in the first place. Looking back at the NP co-ordination in (162), the derivation might actually be (172), rather than (165).

(172) Alternative syntax for (165) compatible with the SIH
   a. [DP &lt;[&lt;f,x,y;[g]]&gt; &lt;x &lt;[g(y) &amp; x]] [exists y [maid(y) &amp; x]] [exists y [maid(y) &amp; x]]]
   b. [DP the &lt;[&lt;f,x,y;[g]]&gt; &lt;x &lt;[g(y) &amp; x]] [exists y [maid(y) &amp; x]] [exists y [maid(y) &amp; x]]]

In (172), the overt determiner does not originate in D, but as an argument within each NP. D is then a derived position, which results from movement of the determiner from its NP-internal site to the D head. In (172), the moves across-the-board out of the two conjoined NPs. This movement leaves the necessary traces to create type t meanings for the NPs and triggers abstraction above the conjunction to return to an &lt;e,t&gt; meaning in D, now in D, can operate on. If the PRO-strategy does not exist, but rather what's going on is movement of an interpreted operator, the right division is made between NP conjunction and DP conjunction. Whereas determiners can originate as arguments within NPs, it is not at all obvious what would originate as an argument in D outside of possessives.

It is important to note that there is one place in the earlier discussion which makes use of the PRO strategy: the analysis of questions I presented in Chapter 4 and pursued at length in this chapter. Recall that PRO originates as the argument of the interrogative C, and moves onto the clausal spine, leaving a proposition type trace:
In this case, it is easy to dispense with PRO, since there is an interpreted operator which could be generated as the argument of C: namely, ANS. That is, the structure in (173) could be modified to (174) without any change in the predicted meaning:

(174) Recall: LF for simple question

The intent in this LF is that ANS itself originates as the argument of C and moves onto the clausal spine, leaving a proposition type trace in its base position. The viability of this idea rests on ANS always being present wherever a question occurs. One environment where this may not work is the pseudo-cleft environment in Chapter 4. There, I proposed that Exh$_{ep}$ composes with the entire Hamblin set for the pre-copular question. For that to be so, ANS must be absent, since it whittles down the Hamblin set to the subset of Strongest-True answers. One possibility is that Exh itself originates as the argument of C in the pre-copular question in the pseudo-cleft, rather than ANS, but I leave evaluating the plausibility of that to future consideration.
9.4 Collective predication

A further important issue to address is the analysis of collective predication, as in (175-a), and NP co-ordinations such as (175-b), issues introduced earlier, at the beginning of Chapter 3.

(175) Conjunction involving sums
   a. John and Mary met.
   b. Every man and woman got married today.

And in (175-a) seems to form the mereological sum of John and Mary, with met than predicat ed of that sum. In the hydra case, the sentence is interpreted with the universal quantifier ranging over man-woman sums. As noted in Chapter 3, it is not in principle a problem for the SIH for and to be ambiguous between multiple unrelated basic meanings. The intent of the SIH is that logical and does not have type-lifted variants.

Nonetheless, it does not seem accidental that collective predication and hydras involve the same lexical item as logical conjunction and this is a common property across many languages. As such, we are led to ask whether sum formation and really exists, or whether it can be re-analyzed with logical and. If the latter, the SIH then leads us to ask whether it can be analyzed with the basic meaning, [and]. I am aware of two lines of research which analyze sum formation conjunction with logical and: one due to Winter (2001) and pursued further in Champollion (2015), the other due to Schein (2015). The former as it currently stands involves a type-lifted variant of logical and, while the latter explicitly sets as its goal invoking just [and]. I sketch each in turn, and leave further investigation to the future.12

In Winter’s system, there is a covert operator which “existentially raises” John to a quantifier meaning, modeled directly as a set of sets, as in (176-a): John is raised to the set of sets which contain John as an element. Likewise, Mary is existentially raised to the quantifier meaning in (176-b): the set of sets which contain Mary as an element. Conjunction intersects those two sets to create the set of sets which contain both John and Mary, as in (177).

(176) Existentially raising the conjuncts
   a. [John] = \{X : John ∈ X\}
   b. [Mary] = \{X : Mary ∈ X\}

(177) Conjoining John and Mary
   a. [John and Mary] = \{X : John ∈ X\} ∩ \{X : Mary ∈ X\}
   b. = \{X : John ∈ X ∧ Mary ∈ X\}

The set in (177-b) includes a set containing just John and Mary, as well as many sets which contain them in addition to other entities. A minimization operator applies to return the minimal member, i.e. \{John, Mary\}. Winter defines met so that it composes with a set of sets.

Champollion extends the analysis to NP co-ordinations like (175-b). Man is existentially raised to a quantifier meaning: the set of sets containing a man and, likewise, woman is existentially raised to the set of sets containing a woman. These quantifiers are intersected to yield the set of sets containing a man and a woman. This includes sets containing just a man and a woman, as well as larger sets containing a man

12 Note that we can profitably ignore possible worlds altogether in presenting this theory.
and a woman in addition to other entities. A minimization operator again applies to return just the set of man-woman pairs, and every would operate on that set.¹³

Moving between sets and their characteristic functions (as Champollion does), and in the Winter-Champollion analysis would be interpreted as [and₃], since it operates on quantifiers. Fox (2013) proposes a revision to the system, modifying just its semantic types. In Fox’s revision, a different type-lifted variant of and is invoked. The derivation for (175-a) would proceed from an LF like this:

(178)  LF for (175-a)

\[
[[\text{DP} \equiv [\text{MIN } \& \text{P} [\text{Op John} ] [\text{and} [\text{Op Mary}]]]]] \text{ met}
\]

Whereas Winter and Champollion model plural individuals as sets, Fox models them as sums proper. A covert operator thus transforms John not into a quantifier, but into an <e,t> predicate. That predicate operates on entities (atomic or sum), and characterizes the set of entities which include John as an atom (i.e. John, John+Mary, John+Mary+Sue, ...), as in (179-a). Likewise, Mary is shifted to an <e,t> predicate characterizing the set of entities containing Mary as an atom, as in (179-b). Those two predicates are conjoined, as in (180), and that conjunction requires [and₂].

(179)  Interpreting the conjuncts

a. \[ [\text{Op John}] = \lambda X . \text{John} < X \]
b. \[ [\text{Op Mary}] = \lambda X . \text{Mary} < X \]

(180)  Interpreting the &P

a. \[ [\& \text{P}] = [\text{and₂}([\text{Op John}])([\text{Op Mary}])] \]
b. \[ = \lambda X . \text{John} < X \land \text{Mary} < X \]

A minimization operator applies to return a new <e,t> predicate characterizing the singleton set containing the minimal element in the extension of (180), i.e. John+Mary. That predicate forms the restrictor of a covert existential quantifier, and the sentence comes out true iff John+Mary is in the extension of met.

Although these analyses involve type-lifted variants of and, they set the stage to analyze conjunction apparently creating sums with [and] itself, since they invoke logical and, rather than a separate operator. In order for [and] to compose, a type t node is needed internal to the NP, much as was the case for the NP co-ordinations discussed in the previous subsection. In that spirit, the LF in (178) could cede to a derivation along the lines of:

(181)  Possible LF for (175-a)

a. \[ [[\text{DP} \equiv \text{MIN } \& \text{P} [\exists [\text{Op John} ] [\text{and} [\exists [\text{Op Mary}]]]]]] \text{ met}] \]
b. \[ [[\text{DP} \equiv \lambda X [\text{MIN } \& \text{P} [\text{t₁ [Op John] } [\text{and} [\text{t₁ [Op Mary]]]]]]] \text{ met}] \]

The covert existential originates just above the silent operator that shifts John to predicate type in the left conjunct, and at a parallel position in the right conjunct. From there, it moves to the D head position, leaving traces saturating the <e,t> predicates, creating the type t meanings:

¹³Champollion illustrates with an example involving a covert existential (Ten men and women got married today).
Interpreting the conjuncts

a. \[ [t_1 \text{ Op John}]^g = 1 \text{ iff John } < g(1) \]
b. \[ [t_1 \text{ Op Mary}]^g = 1 \text{ iff Mary } < g(1) \]

The resultant type t meanings can then be conjoined with \[ [\text{and}] \]: the conjunction is true of an entity, \( g(1) \), just in case \( g(1) \) is an entity with John as an atom and is an entity with Mary as an atom.

Interpreting the \&P

a. \[ [\&P]^g = [\text{and}][[t_1 \text{ Op John}]^g][[t_1 \text{ Op Mary}]^g] \]
b. \[ = 1 \text{ iff John } < g(1) \land \text{ Mary } < g(1) \]

Since the landing site of the determiner is above the MIN operator, \( g(1) \) is not abstracted over until after MIN integrates. MIN must, therefore, be redefined to operate on type t meanings. Perhaps it is an Exh operator, of the sort discussed in Chapter 4, so that it asserts that \( g(1) \) is an entity with John as an atom, and Mary as an atom — and nothing else. I leave working out the details of the derivation to a future occasion, but this does seem a promising direction to pursue.

The second line of analysis for collective predication comes from Schein (2015), which I noted at the outset of the thesis is an important antecedent for this project. He pursues the hypothesis that and is always interpreted as and , with a particular focus on collective predication. To achieve this, he adopts a quite different view of the syntax-semantics mapping than I have, by proposing a novel theory of translation into a semantics with plural event pronouns (what he calls Eventish). It would take us too far afield to present the system in full detail, but a rough sketch of his analysis of a sentence like (175-a) is given in (184):

Schein-style analysis of (175-a)

\[ \exists e \exists e' \text{ [John is a participant in } e \land \text{ Mary is a participant in } e' \land e + e' \text{ is a meeting]} \]

In prose, (184) says that there is an event \( e \) in which John is a participant, and an event \( e' \) in which Mary is a participant, and the plurality of \( e \) and \( e' \) is a meeting event. The overt and visible in (175-a) corresponds to the leftmost \( \land \) in (184). Comparison of the Winter-Champollion approach with Schein’s is a matter of importance, but one I must postpone.

9.5 Local summary

The aim in this section was to lay out a template for further inquiry by making programmatic suggestions about what covert syntax might look like in a range of different environments to allow and to be interpreted as [and]. I did not intend to argue that the ideas here are correct. My hope is that they will lay a fruitful foundation for future work. Now, it is time to leave conjunction and move on to focus operators.
Chapter 6
Focus operators

1 Introduction

The preceding chapters have focused on coordination. In this chapter, I shift attention to another class of expression with a broad surface distribution: focus operators. I will restrict attention to *only* and, much as I did with *and* in Chapter 3, I will zero in on just one profile of data: cases where the object DP is focused or properly contains the focused constituent. In those data, overt *only* may precede the vP, or it may precede the object DP. The data in (1) illustrate. In those examples, *one* is focused, as signaled by prosody: *one* is produced with greatest prosodic prominence.\(^1\)

(1) **Pre-vP vs. pre-DP only**
   
a. John *only* learned \(\text{ONE}_F\) language.
   
b. John learned *only* \(\text{ONE}_F\) language.

How does *only* compose in both environments? Given that the vP has a sentential meaning under current syntactic assumptions, *only* in (1-a) may be interpreted as the sentential operator [only]. The case which seems to require type-lifting is (1-b), where *only* seems to be interpreted as [only\(_3\)], the variant of *only* able to compose with quantifiers. Is [only\(_3\)] really invoked in (1-b)?

The SIH predicts that (1-b) must involve a richer syntax than meets the eye so that pre-DP *only* operate on a sentential meaning, despite appearances. In this chapter, I argue that the prediction of the SIH is correct. I provide empirical evidence that pre-DP *only* actually composes with the vP, and develop a syntax which achieves this, along with a theory of syntax-phonology mapping to explain how the underlying syntax comes to be obscured in the surface string.

The rest of the introductory section takes care of preliminaries. I spell out the full analysis of (1-a) based on the basic meaning [only], and show how (1-b) would be analyzed with type-ambiguity, setting the stage for the remainder of the chapter. Note that this chapter is written in the second version of intensional semantics from Chapter 2, with world arguments instead of a world parameter (for discussion, see Chapter 2, Section 3).

1.1 The basic meaning of *only*

The basic meaning of *only* is very similar to the meaning of Exh, which we discussed in detail in the last chapter. I proposed a revision to the traditional meaning for Exh in order to account for data with adverbs such as *possibly*, and I provided data with overt *only* showing that it must be revised in

\(^1\)Small capitals indicate the word containing the most strongly stressed syllable.
kind. Since adverb data will not make an appearance in this chapter, we can revert to a more classical formulation here. For this chapter, I adopt the definition of only in (2), from Chapter 2. Applied to a prejacent proposition p, ]only[ presupposes that p is true and asserts that non-weaker alternatives to p in ALT are false.\(^2\)

(2) **Basic meaning for only**

\[
[\text{only}]^{ALT} = \lambda p . \lambda w . p(w) . \forall p' \in ALT [p'(w) \rightarrow p \subseteq p']
\]

The source of ALT was a primary concern of the last chapter, where it was crucial to take ALT to be an argument of Exh and only. In this chapter, we will be more concerned with how the prejacent is assembled, and I relegate ALT to parameter status to streamline the LFs.

As discussed in Chapter 2, Rooth (1985) assumed a basic meaning for only along the lines of ]only[, but proposed that this basic meaning is never actually attested. ]only[ was a phantom operator which gave rise to attested higher-type meanings through type-shifting. This is a natural conclusion to reach based on the traditional assumption that just TPs have sentential meanings. Under that assumption, ]only[ would only be attested if only adjoined to the TP — and only cannot in general adjoin that high. Modifying the earlier examples in (1) by moving only to sentence-initial position outputs a sharply unacceptable result: \(^3\)

(3) **Constraints on pre-TP only**

Only John learned ONEF language.

Whereas Rooth suggested that only is never interpreted as ]only[, I will pursue the hypothesis that ]only[ is its sole meaning. The most straightforward case to re-analyze with ]only[ is the example in (1-a), where only is pre-verbal. With the classical syntax in (4), only must type-lift to operate on a property \((\lambda x . \lambda w . x \text{ learned } \geq 1 \text{ language in } w)\).

(4) **Classical LF for (1-a)**

\[
[\text{TP John} [\text{only} [\text{VP learned one language}]]]
\]

The situation changes, however, with the VP Internal Subject Hypothesis, as discussed in Chapter 3 with and. The syntax in (4) cedes to (5-a), where the subject John originates inside the verbal projection and raises to spec-TP. With QR of one language, the full LF is (5-b).

(5) **Updated LF for (1-a)**

a. \[
[\text{TP John}_1 [\text{only} [\text{VP t1 learned one language}]]]
\]

b. \[
[\text{TP John}_1 [\text{only} [\text{VP2 one language}_2 [\text{VP1 t1 learned t2}]]]]
\]

\(^2\)There are a number of disagreements about how to state [only]. Whereas Horn (1969) takes p(w) to be presupposed, Horn (1996) adopts a weaker presupposition that some proposition in ALT be true, though not necessarily p (see also Wagner 2006, von Fintel & Iatridou 2007). Even if p(w) is part of the entry, p(w) may be presupposed, asserted (e.g. Kuroda 1969, Lakoff 1970, Taglicht 1974), or implicated (Karttunen & Peters 1979). These issues are not relevant here.

\(^3\)Initial only is tolerated when the subject is focused (Only JOHN_{EF} learned Spanish) and, in that case, there is the possibility that only forms a constituent with the subject.
*Only* attaches to vP$_2$, which has a sentential meaning. In particular it expresses the proposition in (6), which constitutes an appropriate prejacent for [only]. [only] thus triggers the presupposition that John learned at least one language.

(6) **The prejacent**

[[vP] = λw . John learned ≥1 language in w]

To calculate the assertion, a value for ALT is required. ALT is a set of propositions whose make-up depends on the focus structure in the sentence. In general, ALT contains those propositions which are just like the prejacent, but with the focused element in the prejacent replaced by alternatives to it. I take ALT in (1-a) to be the set of propositions in (7), informally propositions of the form *that John learned* $n$ *languages* for different numbers $n$.

(7) **The ALT parameter**

a. ALT ≈ {λw . John learned ≥1 language in w,
b. λw . John learned ≥2 languages in w,
c. λw . John learned ≥3 languages in w, ...}

Having reverted to the original definition for [only], rather than the definition from Chapter 5, [only] negates any alternatives not entailed by the prejacent. The alternative in (8-a) is entailed (in fact, it is equivalent to the prejacent), while the other alternatives are logically stronger. These alternatives are negated, entailing that John did not learn two or more languages. Collapsing presupposition and assertion, the sentence conveys that John learned at least one language, but not two or more languages, i.e. he learned exactly one language — the correct meaning.

1.2 **Defining [only]$_3$**

Now, we turn to the meaning whose existence is under contention in this chapter: [only]$_3$, which seems to be required when *only* precedes the DP, rather than the vP. As defined in (8), [only]$_3$ operates on a quantifier intension (type <est, st>) to output a new meaning of the same type. [only]$_3$ is a function from quantifiers to quantifiers.

(8) **Defining [only]$_3$**

a. \[\text{[only]$_3$}^{ALT} = \lambda F_{\text{est, st}} . \lambda f_{\text{est}} . \text{[only]}(F(f))\]

b. \[= \lambda F_{\text{est, st}} . \lambda f_{\text{est}} . \lambda w : F(f)(w) . \forall p' \in \text{ALT} \left[ p'(w) \rightarrow F(f) \subseteq p' \right]\]

How, then, is example (1-b) with pre-DP *only* built syntactically and interpreted? The base structure would be the obvious one where *only* adjoins directly to the DP that follows it, as in (9).

(9) **LF for (1-a)**

[[TP John$_1$ [vP t$_1$ learned [only [DP one language]]]]]
The internal composition of one language is not of special interest, but the final meaning is (10). That could be derived with one itself a quantificational determiner, or one could be a cardinality predicate which composes with language inside the restrictor of a covert existential quantifier. 4

(10) **Interpreting one language**

\[ [DP] = \lambda f_{ext} \cdot \lambda w . \exists x [\#(x) = 1 \land \text{language}(x)(w) \land f(x)(w)] \]

Since \([DP]\) is a quantifier, \([\text{only}_3]\) can compose with it to create the new quantifier in (11), which wraps up the contributions of both only and one language into a single meaning. Applied to a property \(f\) at a world \(w\), the quantifier introduces the presupposition that \(f\) contains at least one language in its extension at \(w\).

(11) **[only one language]**

a. \(= [\text{only}_3]^{ALT}(\text{one language})\)

b. \(= \lambda f_{ext} . [\text{only}]^{ALT} ([\text{one language}](f))\)

c. \(= \lambda f_{ext} . \lambda w : [\text{one language}](f)(w) . \forall p' \in ALT [p'(w) \rightarrow [\text{one language}](f) \subseteq p']\)

To compute the assertion, a value for ALT is required. I assume that ALT is a set of propositions, regardless of where only attaches. 5 In this case, ALT is the same set of propositions as in the last section: propositions that John learned \(n\) languages for different \(n\). As such, \([\text{only}_3]^{ALT}(f)(w)\) asserts that \(x\) does not contain two or more languages in its extension at \(w\). Collapsing presupposition and assertion, \([\text{only}_3]^{ALT}(f)(w)\) conveys that \(f\) is true of exactly one language at \(w\).

Now, like any object quantifier, only one language must QR, as in (12). QR derives the property of being something John learned, and that property saturates the quantifier [only one language], as in (13). The overall meaning is, then, the correct one: the proposition true just in case the property of being something John learned contains exactly one language, i.e. the proposition that John learned exactly one language. To streamline the composition in (13), I immediately fill in John as the value for the subject trace.

(12) **Full LF for (1-b)**

\[ [TP \text{ John}_1 [[\text{only}_3 \text{ [DP one language]}] \ [\text{tp}_1 \text{ learned } \text{t}_2]]] \]

(13) **The overall meaning**

a. \([\text{only one language}](\lambda x . \lambda w . \text{John learned } x \text{ in } w)\)

b. \(= \lambda w : \text{John learned } \geq 1 \text{ language in } w\)

\[ \forall p' \in C [p'(w) \rightarrow [\lambda w . \text{John learned } \geq 1 \text{ language in } w] \subseteq p']\]

4 Note that \# is a function which takes an atomic or plural individual and returns the number of its atomic members.

5 Rooth (1992) introduces focus alternatives in such a way that they must be the same semantic type as the sister of the focus operator. In that case, ALT with \([\text{only}_3]\) would be a set of quantifier meanings, rather than a set of propositions. Rooth (1985) fixes the value for ALT through a covert operator R (for ‘restrictor’), defined in such a way that it yields a set of propositions, regardless of the sister of only. Since I will reject \([\text{only}_3]\) and my arguments are independent of how the value for ALT comes about, I will gloss over this issue and simply assume that ALT is an appropriate set of propositions.
Although the correct meaning is derived in this instance, the SIH leads us to expect that this analysis is not the correct one, and that *only* constructions involve some covert syntax which makes the appearance of *[only]* deceptive.

### 1.3 Preview of main claims

In Chapter 3, I pursued the hypothesis that apparent object DP conjunction involves underlying vP conjunction. In this chapter, I pursue a parallel hypothesis for *only*: when it precedes an object DP, the interpreted operator is still *[only]*, scoping at the vP. The idea is flagged:

$$\text{(14) The central hypothesis}$$

For both pre-DP and pre-vP *only*, *[only]* scopes at the vP.

The argument proceeds in three steps. First, I provide empirical evidence that pre-DP *only* does not directly compose with the DP. To do this, I adapt diagnostics for hidden syntax from Chapter 3, in particular the tests involving split scope (Chapter 3, Section 5) and VP-ellipsis licensing (Chapter 3, Section 4). Both diagnostics converge on a key conclusion that *only* and the DP may make their semantic contributions at distinct syntactic heights. *Only* may be interpreted at one place in the structure, while the DP is interpreted lower. The pattern is unexpected if *only* DP is a single quantifier — but is a prediction of the hypothesis.

The second part of the chapter provides a detailed analysis caching out the central hypothesis. The proposal must explain how *only* can appear at different positions in the linear string, but still have its semantic contribution localize uniformly at the *[only]*. To achieve the desideratum — phonological variability, but semantic uniformity — I propose a richer underlying syntax for *only*, and take the phonology and the semantics to interact with that syntax in different ways. In a nutshell, the syntactic claim is that *only* constructions involve *two* underlying heads, one at the vP, and the other with the object DP containing the focus. The two heads Agree for an *[ONLY]* feature:

$$\text{(15) The analysis in a nutshell}$$

$$\begin{align*}
[TP John_t [ONLY_t [iONLY()]] [vP t_1 learned [F [uONLY(+)] [DP oneF language]]]] \\
\Rightarrow \text{Phonology: only can realize the [ONLY] feature on ONLY or F.} \\
\Rightarrow \text{Semantics: the interpreted head is ONLY, bearing [iONLY]; F is inert}
\end{align*}$$

The mapping from syntax to phonology is variable such that overt *only* may optionally realize one or the other head. The syntax-semantics mapping, by contrast, is invariant: the vP-level head is always interpreted as *[only]*, while the DP-level head is semantically inert. *Only* constructions exhibit surface variability, but their underlying syntax and semantics is constant.

Under the proposal, the phonology and semantics crucially come apart with pre-DP *only*. Pre-DP *only* realizes the DP-level head, which is semantically inert. The vP-level head is interpreted, but phonologically abstract. In effect, pre-DP *only* will amount to dummy Agreement morphology with
the vP-level head. The situation is reminiscent of analyses of negative indefinites (Penka 2011) and negative concord (Zeijlstra 2008, 2011), which analyze DP-level negative morphology as Agreement with an abstract interpreted negation higher in the structure.

The third and final part of the chapter discusses previous work on different languages, based on which my own proposal is built. Barbiers (2014) and Bayer (1999, 2016) observe that 'only' in colloquial Dutch can involve two overt morphemes simultaneously, which they explicitly take to reveal two underlying heads. I will discuss related data from Vietnamese in Höle (2013) and Erlewine (to appear). Moreover, Horváth (1997) argues that a feature like [ONLY] must exist, based on focus fronting in Hungarian (also Cable 2007, 2010). Lee (2004) is the closest semantic precedent: she argues that the Korean morpheme corresponding to pre-DP only should be analyzed as dummy Agreement morphology, with the meaning at a higher operator. By building on these works, the proposal fits English only naturally into the cross-linguistic picture.

1.4 Roadmap for the chapter

The chapter follows the logic just laid out. I present the key scope and ellipsis data in Sections 2 and 3, respectively. I present the proposal in Section 4, and then show how it helps account for additional scope data in Section 5. The cross-linguistic arguments are presented in Section 6. Finally, I compare the proposal with alternatives in Section 7, and conclude in Section 8, where I lay out a template for continued investigation of focus operators, based on the chapter.

2 Test 1: Split scope

Chapter 3 introduced a series of diagnostics to show that and in apparent object DP conjunction can take scope at the vP. This chapter extends two of those diagnostics to pre-DP only. The first is “split scope”. Recall the split scope signature in apparent object DP conjunction: and can scope at one height, while the apparent DP conjuncts scope lower. This section establishes a similar signature with pre-DP only: only can scope at one height, while the following DP scopes lower. If only directly operates on the DP, there is no viable derivation for split scope. Only must make its semantic contribution in a different way, separate from the DP. At this point, the goal is not to present an analysis of the data, but rather to lay out the problem.

2.1 The split scope signature

To get a feel for how only can take scope above or below other operators, let us start with example (16), where only DP occurs in an embedded clause introduced by require. Because the DP following only is Spanish, which is inert for scope, this datum isolates the scope behavior of only.

(16) Scope of only
    John is required to learn only Spanish.

Footnote 6: For related observations of split scope, see von Fintel & Iatridou (2007). Wagner (2006) observes split scope with pre-vP only, with the data attributed to Daniel Büring and Yael Sharvit. We will discuss the pre-vP only data later on (see Sections 4, 5, and 7), including Wagner’s treatment of it (see Section 7).
Taglicht (1984) observed that (16) is ambiguous. On one reading, the requirement is that John learn Spanish and no other language. On the other reading, it is required that John learn Spanish, and there is no further requirement: he need not learn other languages — but can if he wants. If John learns Spanish and French, he has violated the former requirement, but satisfied the latter. The two readings are flagged in (17). The contrast between them is one of scope: in (17-a), require scopes above only and, in (17-b), the scope order is reversed.

(17) Scope ambiguity in (16)

- a. “The requirement is that John learn Spanish and nothing else.” (require > only)
- b. “The only requirement is that John learn Spanish.” (only > require)

The sentence in (17) made it possible to detect the scope of only — but not the scope of the following DP. To test for split scope, we must replace Spanish with a quantifier which itself scopally interacts with require. The DP one language fits the bill. Its scope interaction with require is illustrated in (18), first without only, to get a feel for the readings at stake.

(18) Scope of one language

John is required to learn one language.

On the reading in (19-a), the existential contributed by one language takes scope above the universal modal, yielding a specific de re reading of one language: there is some particular language such that John learns that language at every world compatible with what is required, i.e. at every ‘R-world’. It might be, for instance, that John learns French at every R-world. The de re reading can be brought out by following (18) with the continuation in (19-b).

(19) Available: one > require

- a. \( \exists x \ [\text{language(x)}(w_0) \ & \ \forall w' \in R(w_0) \ [\text{learn(x)}(\text{John})(w')]] \)
  “There is one particular language that John is required to learn.”
- b. ... namely, French.

A second option is for the existential to take narrow scope below the universal modal. One language may then be interpreted de dicto. The de dicto reading is given in (20-a): at every world compatible with what is required, there is a language that John learns. Now, the language might be different at different R-worlds. He might learn French at one R-world, and German at the next. The continuation in (20-b) biases the reading.

(20) Available: require > one

- a. \( \forall w' \in R(w_0) \ [\exists x \ [\text{language(x)}(w') \ & \ \text{learn(x)}(\text{John})(w')]] \)
  “The requirement of John is that he learn any one language.”
- b. ... he can learn any one he chooses.
With this in hand, we are ready to combine only and one language to test for split scope. The key matter of interest: when only precedes one language, can only and one language scope at different heights relative to require? The test example in (21) mixes (16) and (18) together.

(21) **Scope of only + one language**

John is required to learn only one language.

Clearly, split scope is possible: (21) allows a reading where only scopes above require, while one language is interpreted non-specifically, below require. Under the split scope reading, the only requirement is that John learn one language, but he is allowed to learn more (diagnosing only > require); moreover, he can learn any language he chooses (diagnosing require > one language).

(22) **Paraphrase: only > require > one language**

“The only requirement of John is that he learn any one language.”

Note that the split scope reading is crucially distinguishable from alternative readings where only and one language scope together. If only and one language both scope above require, then the sentence would paraphrase as (23). On this reading, unlike (22), one language is de re, so John must learn some particular language. Alternatively, if only and one language both scope below require, the reading would be (24), again distinct from (22): while (22) allows for John learning multiple languages, (23) limits John to a single language, due to the narrow scope of only.

(23) **Paraphrase: only, one language > require**

“There’s only one particular language that John is required to learn.”

(24) **Paraphrase: require > only, one language**

“The requirement of John is that he learn one language and no more.”

The readings in (23) and (24) are, of course, available — but the crucial fact is that the reading in (22) is available too. Split scope is illustrated more dramatically with the new example in (25-a). The only pragmatically plausible reading of the sentence is (25-b): it is required that you publish three papers, though you may publish more (diagnosing only > have), and you can publish any three papers you want (diagnosing a de dicto reading for three papers; have > three).

(25) **Split scope prototype**

a. To be considered for tenure, you have to publish only three papers.

b. “The only requirement is that you publish any three papers.”

In the given scenario, only clearly scopes above have. If only took narrow scope, (25-a) would say that you must publish exactly three papers to go up for tenure. This would be an odd requirement, since presumably someone with more than three papers should also be eligible for tenure.

(26) **Implausible: require > only > three**

“You must publish exactly three papers.”
Moreover, because *publish* is a creation verb, the only sensible reading of *three papers* is the de dicto one. If *three papers* took wide scope above *require*, (25-a) would say that there are three particular papers that you must publish to be considered for tenure — which is nonsense. The sole reasonable interpretation of (25-a) has split scope.

(27) **Implausible: only, three > require**
    “There are only three particular papers which you must publish.”

Having established that *only* exhibits a split scope signature, stated as (28), we turn to the analytical matter: can the type-ambiguity analysis of pre-DP *only* derive split scope?

(28) **Split scope signature**
    Pre-DP *only* scopes above an operator that its DP associate scopes below.

### 2.2 Insufficiency of [only3]

If *only* operates directly on the DP as [only3], split scope does not derive. To build up, let us consider how scope ambiguity *would* derive in Taglicht’s original example:

(29) **Taglicht ambiguity**
    John is required to learn only Spanish.

For [only3] to compose with *Spanish*, *Spanish* must be interpreted as a quantifier, albeit a scopally inert one. This would be achieved by Montague lifting *Spanish* from type e to type <est, st>. *Only Spanish* is then interpreted itself as the quantifier in (30).

(30) **[only Spanish]**
    a.  = [only3]ALT([Spanish])
    b.  = λf est . [only]ALT (λw . f(Spanish)(w))

Taglicht’s ambiguity then becomes a pedestrian quantifier scope ambiguity. Like any quantifier, *only Spanish* QRs, and the scope of *only* is determined by the QR site of *only Spanish*. If *only Spanish* QRs within the embedded clause, as in (31), *only* scope below *require*. Conversely, if *only Spanish* QRs into the matrix clause, as in (32), *only* takes wide scope.

(31) **Low QR: require > only**
    \[TP \text{John} \lambda 1 [vP t1 \text{is required}] [TP PRO1 \lambda 2 [vP [only Spanish] \lambda 3 [vP t2 \text{learn t3}]])\]

(32) **High QR: only > require**
    \[TP \text{John} \lambda 1 [vP [only Spanish] \lambda 3 [vP t1 \text{is required}] [TP PRO1 \lambda 2 [vP t2 \text{learn t3}]])\]

Split scope, however, is problematic. If [only3] creates a single quantifier *only DP* which QRs as a constituent, *only* and the DP should necessarily scope at the same height. Concretely, in (121),
only three papers would have the quantifier meaning in (33), with the contributions of only and three papers wrapped up into that single quantifier.

(25-a) **Split scope prototype**  
You have to publish only three papers.

(33) **[only three papers]**  
a. = [only_3]^{ALT} ([three papers])  
b. = \lambda f_{est} \cdot [only]^{ALT} (\lambda w . \exists X [\#(X) \geq 3 \land \text{paper}(X)(w) \land f(X)(w)])

Mirroring what we saw with Taglicht's example, there are two QR options for only three papers: either the quantifier QRs below have, or it QRs above have. QR below have results in only and three papers both taking narrow scope, while QR above have results in them both taking wide scope.

(34) **Low QR: have > only > three papers**  
\[ TP \text{ you } \lambda 1 \ [vP \ t1 \text{ have } [TP \text{ PRO1 } \lambda 2 \ [vP \text{ [only three papers] } \lambda 3 \ [vP \ t2 \text{ write } t3]]]] \]

(35) **High QR: have > three papers > only**  
\[ TP \text{ you } \lambda 1 \ [vP \text{ [only three papers] } \lambda 3 \ [vP \ t1 \text{ have } [TP \text{ PRO1 } \lambda 2 \ [vP \ t2 \text{ write } t3]]]] \]

To accommodate split scope, a different analysis is needed, one which allows only to make its semantic contribution remote from where the DP three papers makes its contribution.

2.3 **Type-lifting the quantifier**?  
As we saw with and in Chapter 3, it is in principle possible to rework the semantics to derive split scope while still having only compose with the DP. Three papers would be Montague lifted from type <est,st> to type <<<est,st>,st>,st>.

(36) **Lifting the quantifier**  
a. [three papers] = \lambda f_{est} \cdot \lambda w . \exists X [(X) \geq 3 \land \text{paper}(X)(w) \land f(X)(w)]  
b. [three papers\textsuperscript{\uparrow}] = \lambda F^{<<<est,st>,st>,st> . \lambda w . F([three papers])}(w)

To compose with [three papers\textsuperscript{2}], only itself would be lifted to a higher type than [only\textsuperscript{3}]. The relevant denotation, which I call [only\textsuperscript{5}], is given in (37). The output of composing [only\textsuperscript{5}] with [three papers\textsuperscript{\uparrow}] is the new meaning in (38), again of type <<<est,st>,st>,st>.

(37) **Higher-type only**  
[only\textsuperscript{5}]^{ALT} = \lambda G^{<<<est,st>,st>,st>} . \lambda F^{<<<est,st>,st>} . [only]^{ALT} (G(F))

(38) **Higher-type only three papers**  
a. [only three papers\textsuperscript{\uparrow}]  
b. = [only\textsuperscript{5}]^{ALT} ([three papers\textsuperscript{\uparrow}])  
c. = \lambda F^{<<<est,st>,st> . [only]^{ALT} (F([three papers]))}
Given these semantic ingredients, the detailed LF for split scope is (39). Only three papers QRs in two steps. First, it move to the edge of the embedded vP, below have, and then it moves to the edge of the matrix vP, above have. The first step of QR leaves a trace (t₁) of type e, able to compose with publish. The second step leaves a trace (t₂) of quantifier type, i.e. type <est, st>. Abstracting over t₂ creates a property of quantifiers, as in (40), which can saturate the argument of [only three papers]. The resultant meaning derived in (40) is the split scope reading.

(39) LF for split scope

\[
\text{TP w } \left[ \text{vP3 [only three papers]} \right] \left[ \text{vP2 } \lambda_2 \ [\text{vP you}_3 \text{ have} \ [\text{TP } t_2 \ [\text{vP1 PRO}_3 \text{ publish } t_1]]]] \right]
\]

(40) Argument of only three papers

\[
[\text{vP}_2] = \lambda f_{est, st} . \Box [f(\lambda x . \lambda w . \text{you publish } x \text{ in } w)]
\]

(41) The overall meaning

a. \([\text{only three papers}']([\text{vP}_2])\)

b. \(= [\text{only three papers}'](\lambda f_{est, st} . \Box [f(\lambda x . \lambda w . \text{you publish } x \text{ in } w)])\)

c. \(= [\text{only}] (\text{ALT}) (\Box [\text{three papers}](\lambda x . \lambda w . \text{you publish } x \text{ in } w))\)

How exactly has split scope derived? The quantifier variable \(f\) in \([\text{vP}_2]\) is contributed by \(t_2\). When the lifted meaning for only three papers applies to \([\text{vP}_2]\), the basic quantifier \([\text{three papers}]\) ends up being filled in for \(f\). In this way, \([\text{three papers}]\) is interpreted as though it were at the site of \(t_2\), below have. Moreover, saturating \([\text{vP}_2]\) with \([\text{three papers}]\) creates a proposition, and only three papers is defined such that \([\text{only}]\) applies to that proposition, giving only widest scope. Although only three papers is syntactically above have as a unit, the semantics splits scope: only is interpreted above have, while three papers is interpreted as though it were at a lower trace position.

To re-iterate, this derivation provides a mechanism for the syntactic scope of a quantifier to mismatch its semantic scope: three papers is syntactically above have, but interpreted below have. This is the problem. As discussed in Chapter 3, Condition C effects show that syntactic and semantic scope cannot be disentangled. An example from Fox (1999) is modified to include only:

(42) Counter-evidence from Condition C

Only one new theory by Quine\(_1\) seems to him\(_{1, \sqrt{2}}\) to be needed.

The DP only one new theory by Quine originates in the embedded clause, and raises to the matrix subject position. Let us ignore the scope of only, and focus on the interpretation of one new theory by Quine. The only sensible reading has one new theory by Quine interpreted de dicto, below need. The pattern of co-reference shows that this semantic result requires syntactically reconstructing the DP into the embedded clause: with the DP syntactically low, him c-commands Quine,\(^8\) and Condition C explains why him and Quine cannot be co-referent. If there were a mechanism for the DP to remain syntactically high, but still be interpreted low, a de dicto reading could derive without syntactically

\(^7\)To streamline the structure, I show the subjects of the matrix and embedded clauses reconstructed into the vPs.

\(^8\)I assume that there is a mechanism for the complement of a preposition to c-command out of the PP: with the PP [to him], [him] has the c-command domain of the PP.
reconstructing the DP, and no Condition C effect would obtain. The derivation just sketched for split scope constitutes exactly such a mechanism and, if available, could apply in (42) as well. As such, the Condition C effect in (42) rules out the derivation.

2.4 Concluding split scope

Overall, I conclude that there is no viable way to derive split scope with type-ambiguity. If only is interpreted as \([\text{only}_3]\), only three papers is a run-of-the-mill <est,st> quantifier, and only and three papers necessarily scope at the same height. A higher-type interpretation which could in principle derive split scope is disproven by Condition C effects in related data. Given the tight correlation between syntactic and semantic scope, what is needed to derive split scope with pre-DP only is a mechanism for only and the DP to be at different syntactic positions in the LF: only must be syntactically high at the same time that the DP is syntactically low. Achieving this spread requires a parse where, despite appearances, pre-DP only does not compose with the DP at all. Before giving my solution, I present more data corroborating a syntactic scope split in a different way.

3 Test 2: VP ellipsis

In this section, I establish a fact about VP ellipsis licensing which, as far as I know, has gone unnoticed in the literature: that a VP containing an object DP with only preceding it can license ellipsis of another VP that does not contain only. Schematically:

\[
\text{(43) Key ellipsis profile} \\
V \text{ only } D \text{ can serve as antecedent for } V-DP.
\]

The example in (44) provides a concrete illustration. The first sentence in (44-a) is the familiar split scope prototype and its VP appears to be publish only three papers. The VP in the sentence that follows in (44-b) is elided, and that VP (Δ) is interpreted without any sign of only: Δ is ‘publish (at least) three papers’, as paraphrased in (44-c).

\[
\text{(44) Instantiating the key profile} \\
a. \text{ To be considered for tenure, you have to publish only three papers.} \\
b. \text{... and now you're telling me that you can't } \Delta? \\
c. \text{‘... and now you're telling me that you can’t publish at least three papers?’}
\]

For ellipsis licensing to proceed in the key schema, it must be possible for only to be interpreted outside the antecedent VP, at the same time that the DP remains inside the antecedent VP. That is, at LF, V only DP (publish only three papers) must be seen as V DP (publish three papers). This corroborates a syntactic scope split — and, again, is not predicted if only operates on the DP.
3.1 Insufficiency of type-ambiguity

When ellipsis data were previously discussed in Chapters 3 and 4, I adopted the licensing condition in (45), following Takhashi & Fox (2005). I maintain this condition here. Note that 'PD' in (45) stands for 'Parallelism Domain', and 'AC' stands for 'Antecedent Constituent'.

(45) Parallelism Condition
   a. A VP (VP_e) can elide if VP_e is reflexively dominated by a constituent PD, and the linguistic context provides an antecedent AC for PD which is semantically identical to PD, modulo focused marked constituents.
   b. PD is semantically identical to AC modulo focus if there is a focus alternative to PD, PD_{Alt}, such that for every assignment function g, \[ PD_{Alt} = AC \].

To determine what the AC must look like, we need to first pick a PD. To this end, the full LF for (44-b) is given below. The VP that elides is publish three papers. At LF, three papers QRs to the edge of the vP. In principle, any node reflexively dominating the VP is a possible PD. The best candidate is TP_1. Three papers occurs inside TP_1, and the subject trace in the vP is bound. As such, TP_1 has the property meaning in (46-b), informally published (at least) three papers. I will take TP_1 to be the PD, and set aside the other options.

(46) The PD in (44-b)
   a. \[ TP_2 you [TP_1 \lambda 1 [\_p_3 [three papers] [\_p_2 \lambda 2 [\_p_1 t_1 [vP published t_2]]]]]] \]
   b. \[ PD = \lambda x . \lambda w . x published \geq 3 papers in w \]

Now, what must the AC look like? Since the PD does not contain any focused element, the only focus alternative to the PD is the PD itself. As such, an appropriate AC must be semantically equivalent to [PD], i.e. it must denote the property published (at least) three papers.

(47) Desideratum for the AC
   \[ AC = \lambda x . \lambda w . x published \geq 3 papers in w \]

For (44-a) to furnish an AC with that meaning, its LF must incorporate a constituent published three papers, which contains the DP three papers, but does not contain only. If only operates directly on three papers to form a single quantifier, there is no such constituent. Only three papers QRs as a unit, and any possible AC either contains both only and three papers, or neither. Concretely, only three papers QRs to the edge of the embedded vP in (48). The underlined constituent is a candidate for an AC in this structure, since it contains three papers. However, it is not actually a viable AC, since it also contains only. Its semantic value is thus (48), informally publish exactly three papers, which is distinct from [PD].

(48) Possible LF for (44-a)
   \[ TP you \lambda 1 [\_p_1 t_1 have [TP PRO_1 \lambda 2 [\_p only three papers] \lambda 3 [\_p t_2 write t_3]]]] \]
(49) **Best candidate AC in (48)**

a. \([\text{AC}]= \lambda x . ([\text{only}])_{\text{ALT}} ((\text{three papers}) (\lambda y . \lambda w . x \text{ published } y \text{ in } w))\)

b. \(= \lambda x . ([\text{only}])_{\text{ALT}} (\lambda w . x \text{ published } \geq 3 \text{ papers in } w)\)

c. \(= \lambda x . \lambda w . x \text{ published exactly } 3 \text{ papers in } w\)

d. \(\neq [\text{PD}]\)

In (50), *only one language* QRs to the edge of the matrix vP instead. If we take the underlined constituent as the AC, the converse problem obtains to what we saw in (48). This AC does not contain *only*, but it does not contain *three papers* either. Indeed, the trace left by QR is not bound within the AC, resulting in the assignment-dependent denotation in (50), clearly different from [PD]. To include *three papers* in the AC, the AC would have to be the constituent just above \(\lambda 3\), but that does not help matters, since *only* and *have* would then also be included.

(50) **Alternative LF for (44-a)**

\([TP \text{ you } \lambda 1 [\text{wp } [\text{only three papers } \lambda 3 [\text{wp t1 have } [TP \text{ PRO1 } \lambda 2 [\text{wp t2 write t3}]]]]]\]

(51) **Best candidate AC in (50)**

a. \([\text{AC}]^8 = \lambda x . x \text{ published } g(2)\)

b. \(\neq [\text{PD}]\)

To sum up, the AC in (44-a) needs to include *three papers*, but not *only*, and there is no way to achieve this if *only* directly operates on *three papers*. There are only two options: either *only three papers* QRs inside the AC (as in (49)), or it QRs outside the AC (as in (51)). If the former, the AC includes both *only* and *three papers* and, if the latter, it includes neither.

### 3.2 Concluding VP-ellipsis

To capture the ellipsis data, *only* must be interpreted outside the vP at the same time that *three papers* is interpreted inside the vP. As in Test 1, we reach the conclusion that a different analysis is needed to generate a syntactic scope split. Overall, Tests 1 and 2 have converged on a general puzzle: scope and ellipsis data arise which are not expected if pre-DP *only* composes with the DP. For both profiles of data, *only* must be able to make its contribution at one place in the structure, while the DP makes its contribution elsewhere. A new analysis is needed to accommodate the splits.

### 4 Proposal

Scope splits are readily accommodated if the semantic contribution associated with pre-DP *only* is actually made at the vP, with the interpreted operator \([\text{only}]\). Before addressing the syntax in detail, we can establish the reasonableness of the hypothesis by noting that pre-vP *only* displays parallel scope and ellipsis behavior to what we saw with pre-DP *only* in Tests 1 and 2 — and, moreover, that such behavior is naturally expected with pre-vP *only* on the analysis in Section 1.1. Example (52) is a minimal variant of the split scope prototype from Test 1, now with pre-vP *only*. *Only* precedes the matrix vP and scopes above *have*. *Three papers* is de dicto below *have*. 

247
Pre-vP only: replicates split scope

You only have to publish three papers.

The obvious syntax is (53), where only attaches to the vP, and operates on the vP as [[only]]. With only operating on the vP, its scope is entirely independent of that of the DP. Three papers is free to QR on its own to a scope site in the embedded clause, as shown, resulting in the correct scope split: only contributes above have, while three papers stays below.

Obvious structure for (52)

\[\text{TP you } \lambda 1 [\text{vP only } [\text{vP } t_1 \text{ have } [\text{TP PRO}_1 \lambda 2 [\text{three papers}]] [\text{vP } t_2 \text{ publish } t_3]]]]\]

The ellipsis fact from Test 2 replicates in kind in (54). The elided VP in (54-b) is interpreted as ‘publish (at least) three papers’ with its antecedent coming from (54-a). The sentence in (54-a) is just a repetition of the scope example with pre-vP only in (52).

Pre-vP only: replicates ellipsis

a. To be considered for tenure, you only have to publish three papers.

b. And now you’re telling me that you can’t Α? (Α = ‘publish (at least) three papers’)

The LF for (54-a) in (53) above has within it an appropriate AC which contains three papers, but not only. That AC is the underlined constituent in (53). The subject trace in the vP is bound by ‘λ2’, and three papers occurs within that constituent, binding the object trace. The result is the property denotation publish (at least) three papers — identical to Α in (54-b).

In effect, what Tests 1 and 2 show, then, is that pre-DP only displays properties that are expected if only operates on the vP as [[only]]. My strategy to resolve the puzzles, therefore, is to unify the semantics of only across the two positions: whether it precedes the object DP, or precedes the vP, its semantic contribution is made at the vP. The analysis must explain how only can vary in overt position without its semantics varying.

One possibility is that only always attaches at the vP, but other constituents can move around it to derive different word orders. Let us see how the scope example in (52) with pre-vP only could turn into its minimal counterpart with pre-DP only:

Possible derivation for pre-DP only

a. \[\text{TP you } \lambda 1 [\text{vP only } [\text{vP } t_1 \text{ have } [\text{TP PRO}_1 \lambda 2 [\text{three papers}]] [\text{vP } t_2 \text{ publish } t_3]]]]\]

b. \[\text{TP you } \lambda 1 [\text{vP only } [\text{vP } t_1 \text{ have } [\text{TP PRO}_1 \lambda 2 [\text{three papers}]] [\text{vP } t_2 \text{ publish } t_3]]]]\]

c. \[\text{TP you } \lambda 1 [\text{vP only } [\text{vP } t_1 \text{ have } [\text{TP PRO}_1 \lambda 2 [\text{three papers}]] [\text{vP } t_2 \text{ publish } t_3]]] [\text{three papers}]]\]

Ignoring the QR step shown in the earlier LF, the base structure for (52) is (55-a). Perhaps at PF, the object DP three papers could evacuate the matrix vP, as in (55-b), which could then undergo leftward remnant movement to a position above only, as in (55-c). The resultant word order has only flanked by the object DP to its right, and the rest of the vP to its left. If the optional movement steps do not take place, pre-vP only derives; if they do take place, pre-DP only derives. Assuming movement, when it happens, is invisible to the semantics, both word orders would share the same LF, being (53) above. A
derivation for only with leftward remnant movement was proposed in Kayne (1998) (see also Johnson 2009 for a remnant movement derivation of gapping). While this derivation is possible, I believe that the alternative proposal sketched in the introduction can garner more independent evidence cross-linguistically, and I pursue that alternative.

4.1 Syntax: bipartite

The key starting point of the analysis is a claim about covert syntax. To build up, let us set aside the complicated split scope cases for the time being, and revert to the very rudimentary examples presented at the outset of the chapter. The data from (1) are repeated in (56), with the paradigm extended to show that pre-vP and pre-DP only cannot co-occur.

(56) **Extended paradigm from (1)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>John <strong>only</strong> learned one language.</td>
</tr>
<tr>
<td>b.</td>
<td>John learned <strong>only</strong> one language.</td>
</tr>
<tr>
<td>c.</td>
<td>*John <strong>only</strong> learned <strong>only</strong> one language.</td>
</tr>
</tbody>
</table>

Despite the ungrammaticality of (56-c), I propose that pre-vP only and pre-DP only are instantiations of separate heads, and that both of these heads are always present in the structure. What we see as plain only is really a front for a bipartite underlying syntax. One projection, call it XP, occurs on the clausal spine. In all of the examples we will see in this chapter, X takes a vP as its complement. X corresponds to the perceived scope position of only.

(57) **Positioning of X**

X takes a propositional node as its complement, in the cases at hand vP.

At an intuitive level, the role of the second projection, YP, is to signal the associate of only. I take it that Y must c-command the associate, and occur "sufficiently local" to it. I will not provide a precise definition of "sufficiently local" here, but rather operationalize locality as in (58), for cases where the focus is contained in the DP complement to a verb.\(^9\)

(58) **Positioning of Y**

In the cases at hand, Y takes the object DP as its complement.

Putting these pieces together, the shared structure for (1-a) and (1-b) is (59). X takes the vP as its complement, and Y takes the object DP containing the focus as its complement.

(59) **The bipartite syntax**

\[
[TP \text{John}_1 [X [vP \text{t}_1 \text{learned} [Y [DP \text{one}_F \text{language}]]]]]]
\]

\(^9\)The idea that focus heads prefer to occur local to the focus has been discussed in different contexts. For a "closeness" constraint on sentential focus heads, see Jacobs (1983, 1986), Büring & Hartmann (2001), Erlewine (to appear).
An Agree relation forges a link between the two heads. I take it that the grammar includes features particular to specific focus operators, so that *only* is associated with an [ONLY] feature (following Horvath 2000, 2005). For concreteness, I assume the distribution of features in (60).

(60) **Feature specification**

\[ TP \text{John}_1 [X_{\text{ONLY}()}] [vP t_1 \text{learned} Y_{\text{ONLY}(+)}] [DP \text{one}F \text{language}]]

Pesetsky & Torrego (2007) propose that a feature being valued or unvalued is orthogonal to it being interpretable or uninterpretable. In (60), the higher head bears an interpretable, but unvalued [ONLY] probe, while the lower head bears an uninterpretable, but valued [ONLY] feature. ONLY probes downward and finds F in its c-command domain, and enters into an Agree relation with F.\(^{10}\) The bipartite syntax is output to the semantic and phonological modules. The crucial idea has to do with what these interfacing modules do with the bipartite structure, and relies on a claim that the two modules do quite different things. Let us consider the phonology first, and then the semantics.

### 4.2 Mapping to phonology: variable

I propose that variability in the surface position of *only* reflects variability in how one and the same structure is mapped to an overt string. The main idea is this: that overt *only* can realize either head in the bi-partite structure, with these two options alternating in free variation. *Only* can realize X, which places it in pre-vP position. Or, it can realize Y, which places it pre-DP. To formalize this, I follow the idea of Distributed Morphology (Halle & Marantz 1993) that phonological information is inserted into syntactic terminals through a PF process of Vocabulary Insertion (‘VI’). As it relates to *only*, the morphological component of English looks like this:

(61) **Morphological hypothesis (English)**

a. **Vocabulary Items**

\[ \alpha_{\text{ONLY}} \rightarrow \text{only} \]

b. **Constraints on realization**

If two \[ \alpha_{\text{ONLY}} \] features are in a probe-goal relation, at least one is realized.

If two \[ \alpha_{\text{ONLY}} \] features are in a probe-goal relation, at most one is realized.

English morphology has a Vocabulary Item, (61-a), which maps any [ONLY] feature — interpretable or uninterpretable — to *only*. The application of VI is, however, constrained such that *only* should be inserted exactly once between two heads in an Agree relation. This latter condition follows from the conjunction of the two constraints in (61). Of particular note is (61-b), which blocks the unattested pattern in (56-c) (*John only learned only one language*). As we will see later on, this constraint is an idiosyncrasy of certain languages, and is not operative universally.

\(^{10}\)Note that the Agree relation could be modeled differently without affecting the overall analysis: F could be the probe and ONLY the goal, with F probing upwards in the structure, rather than downwards. Upwards Agree has been adopted in Zeijlstra (2004, 2012) and Bjorkman & Zeijlstra (2014), but has been criticized in Preminger (2013) on the basis of \(\phi\)-Agreement. I remain agnostic as to which direction of Agree is correct for the phenomenon at hand, and adopt the above distribution of features purely for illustration.
4.3 Mapping to semantics: constant

Since the syntax of only is fixed across cases, the semantics should also be fixed: the same structure will be compositionally interpreted the same way, independent of its phonological realization. The question is: how does the semantics interpret the structure? The answer comes from our earlier discussion showing that only, regardless of position, exhibits the properties that would be expected if it were interpreted as [only], scoping at the vP. Based on this finding, I propose that the locus of interpretation is the higher head, X. The lower head, Y, I will take to be semantically inert. X occurs at the vP, and is interpreted as [only]. Based on this doling out of semantic labor, I re-christen the two heads with more intuitive labels. X, bearing responsibility for the meaning, will now be called ONLY. Y will be called F, indicating that it is some dummy functional element.

(62) The bipartite syntax (re-labelled)

\[ [TP \text{ John}_1 [\text{ONLY} \text{[ONLY]}] [vP \text{ t}_1 \text{ learned} [FP F \text{[ONLY(+)]} [DP \text{ one}_F \text{ language}]]]]] \]

We will see ample evidence for the correctness of a bipartite syntax, but this syntax does raise a question: if ONLY is sufficient for interpretation, why would natural language include F in the structure? What purpose does F serve? First, let me point out that this question is not very unique to F. Case morphology, for instance, is in even a more puzzling situation than F. It has been proposed that case morphology is housed in a K head, and that head not only is not itself interpreted, but it is not specifically associated with any interpreted head. The existence of case has been given functional explanations. One approach, for instance, proposes that case is useful for dissimilation to help keep multiple nominals in the same domain distinct from one another (Baker 2015, Richards 2010, Yuan 2016). Functional explanations have also been posited for NPI morphology (see Barker 2017 for an idea based on scope disambiguation). Turning back to F, we might imagine explanations in a similar spirit for F. Most obviously, since F is quite local to the focused constituent, its functional advantage, when overt, may be to help disambiguate the associate of ONLY.

It is also possible that F might have meaning. Cable (2007, 2010) proposed a similar bi-partite structure for questions, and took the lower head to introduce a choice function variable operating on a set of focus alternatives. F could do that too, and [only] could be minimally re-defined to bind a choice function variable. Though, the deeper explanatory question would then shift to why language would make use of choice functions when the original definition of [only] without choice functions achieves the same end. At the conclusion of the chapter, I will raise another direction to pursue to perhaps imbue F with some meaning. Still, the concrete claim I will pursue here is that ONLY is the locus of interpretation, and F is inert. I leave the question of why F exists as an open mystery.

4.4 Explaining Tests 1 and 2

The analysis is now in place: the syntax is bipartite, only can realize either ONLY or F, while the interpreted operator is always [only] at ONLY. Before moving back to the scope and ellipsis and puzzles, I want to re-iterate how pre-vP only and pre-DP only are alike and different in this view. In the case of pre-vP only, the semantics and phonology coincide as both occur at ONLY. In the case of pre-DP only, the phonology is moved to F, while the semantics remains at ONLY. Overt pre-DP only
amounts, in effect, to dummy Agreement morphology signaling the presence of abstract ONLY in the structure. From this treatment of pre-DP *only*, the scope and ellipsis data follow. Let us begin with Taglicht’s example, exhibiting ambiguity just in the scope of *only*.

(16) **Taglicht’s ambiguity**

John is required to learn only Spanish.  

\( (\text{require} > \text{only}, \text{only} > \text{require}) \)

Under the type-ambiguity theory, *only Spanish* was a quantifier, and the ambiguity in (62) reflected alternative heights of QR for that quantifier. The proposal re-envisions the ambiguity. The two readings derive from different attachment sites for ONLY:

(63) \( \text{LF1: require} > \text{only} \)

\[ \text{TP you}_{1} [\text{V}_{P} \text{t}_{1} \text{are required}] [\text{TP PRO}_{1} \lambda_{2} [\text{ONLY} [\text{V}_{P} \text{t}_{2} \text{learn}] [\text{FP F Spanish}]])] \]

(64) \( \text{LF2: only} > \text{require} \)

\[ \text{TP you}_{1} [\text{ONLY} [\text{V}_{P} \text{t}_{1} \text{are required}] [\text{TP PRO}_{1} \lambda_{2} [\text{V}_{P} \text{t}_{2} \text{learn}] [\text{FP F Spanish}]])] \]

ONLY takes the *embedded* \( V_{P} \) as its complement in LF1, and the *matrix* \( V_{P} \) as its complement in LF2. \( F \) attaches local to the focused DP *Spanish* in both cases. Because \( F \) is semantically inert, the perceived scope of *only* tracks the location of ONLY: LF1 yields the *require > only* reading, and LF2 the *only > require* reading. With regard to phonology, if overt *only* realizes \( F \), *only* is placed in a position linearly preceding *Spanish* in both structures, deriving Taglicht’s surface string. Since abstract ONLY determines scope and is independent of the DP, the stage is set for split scope:

(121) **Split scope prototype**

You have to publish only three papers.

Pre-DP *only* is again just dummy morphology at \( F \), which in this case takes *three papers* as its complement. The first ingredient for split scope is for *only* to take scope above *have*, which is achieved by having abstract ONLY integrate in the matrix clause, as in (65-a). To derive the second ingredient for split scope — that *three papers* scopes below *have* — the FP independently QRs to a scope position in the embedded clause, as in (65-b).

(65) **Deriving split scope**

a. \( \text{Ingredient 1: only} > \text{have} \)

\[ [\text{TP you} \lambda_{1} [\text{ONLY} [\text{V}_{P} \text{t}_{1} \text{have}] [\text{TP PRO}_{1} \lambda_{2} [\text{V}_{P} \text{t}_{2} \text{publish}] [\text{FP F three papers})]])]] \]

b. \( \text{Ingredient 2: have} > \text{three papers} \)

\[ [\text{TP you} \lambda_{1} [\text{ONLY} [\text{V}_{P} \text{t}_{1} \text{have}] [\text{PRO}_{1} \lambda_{2} [\text{FP F three papers}] \lambda_{3} [\text{V}_{P} \text{t}_{2} \text{publish}])])]] \]

Having achieved a syntactic scope split — ‘*only*’ is interpreted in the matrix clause, and *three papers* in the embedded clause — the ellipsis fact from Test 2 is captured, as well.
Ellipsis licensing

a. To be considered for tenure, you have to publish only three papers.

b. And now you’re telling me that you can’t Δ?

('publish ≥3 papers')

The first sentence in (66), identical to the split scope prototype, has the LF just seen in (65-b). The underlined constituent in (65-b) is an appropriate Antecedent Constituent to license ellipsis of Δ. That constituent does not contain the interpreted head ONLY, but does contain the FP three papers. Because F is inert, the FP is semantically identical to just the DP three papers and, accordingly, that constituent has the denotation in (67-a) ('publish (at least) three papers'). Since the AC and the Parallelism Domain are semantically identical, the Parallelism Condition is satisfied.

The AC and the PD

a. $\lambda x . \lambda w . x \text{ published } \geq 3 \text{ papers in w}$

b. $\lambda x . \lambda w . x \text{ published } \geq 3 \text{ papers in w}$

I noted at the outset of the chapter that pre-vP only shows the same scope and ellipsis patterns as pre-DP only. In the present proposal, the sentence in (68-a) would have the same structure as its counterpart with pre-DP only — the structure in (65-b) above — and the facts in (67) would thus derive in the same way as what we saw with pre-DP only.

Scope and ellipsis with pre-vP only

a. To be considered for tenure, you only have to publish three papers.

b. And now you’re telling me that you can’t Δ?

The alternation between pre-vP only and pre-DP only is just a matter of surface phonology, resulting from optionality in the morphological component responsible for translating the bipartite syntax into an overt string: only can realize F (resulting in pre-DP only), or ONLY (resulting in pre-vP only).

4.5 Local summary

So far, I have presented two puzzles with pre-DP only, and resolved both puzzles by localizing the meaning of 'only' uniformly in a sentential operator at the vP. Because there is never an interpreted component of only that operates on the DP, syntactic scope splits between the interpreted operator and the DP are readily derived — and this is what was needed to resolve the puzzles, split scope in Test 1, and ellipsis licensing in Test 2.

Concretely, I have proposed that only constructions involve two heads: ONLY at the vP and F local to the focused constituent, the object DP in the data we have seen. ONLY and F Agree for [ONLY] features. The phonology and semantics divide labor differently between heads: overt only can realize either head, while the semantics uniformly interprets ONLY as [only]. Pre-DP only is dummy Agreement morphology at F, signaling the presence of abstract interpreted ONLY.

While the particulars are obviously different from what I said about and in Chapter 3, the profile of analysis is exactly the same. There is an unexpectedly rich underlying syntax that is variably mapped to phonology, creating the illusion of a Free Operator. In fact, interpretation is constant, with
the FO scoping at the vP. With an underlying vP-level CR syntax in Chapter 3, and was interpreted as [and]. With the bipartite syntax here, only is interpreted as [only].

4.6 Excursus: negative indefinites

Before moving on, I want to take an excursus into the analysis of DP-level negation in negative indefinites and negative concord to bring out a similarity with how I have analyzed DP-level only. Much as pre-DP only appears to be part of a single quantificational DP, a negative indefinite like no boy seems to have negation as part of a single quantifier. The most obvious hypothesis is that negative existential components are both wrapped up into a single quantificational determiner that is stored in the lexicon and phonologized as no. This hypothesis is captured in the lexical entry in (69):

\[(69)\text{ Apparent denotation for } no\]
\[\text{[no]} = \lambda f, st. \lambda g, st. \lambda w. \neg \exists x [(f(x)(w) \land g(x)(w))]\]

At least in certain languages, however, there is evidence that something more is going on with negative indefinites. That evidence comes from scope splits, just like the one we saw with only. Just as pre-DP only could take scope above an operator, while the DP took scope below that operator, the negative component of a negative indefinite can take higher scope than the existential component. The relevant data have been discussed most extensively in German, where keine is the negative indefinite form that would be most obviously glossed as no. Split scope with kein(e) is illustrated in (70) (Bech 1955, Jacobs e.g. 1982; this example is due to Jacobs).

\[(70)\text{ Negative indefinite kein(e)}\]
\[\ldots \text{ weil du keine Jacke anziehen brauchst} \]
\[\ldots \text{ because you no jacket wear need} \]
\[\ldots \text{ because you do not need to wear a jacket.' (} \neg \text{ need} > \exists\]}

The critical observation relates to the scope of the \(-\) and \(\exists\) apparently contributed by kein(e) relative to brauchst ('need'). The most natural reading of (70), reflected in the gloss, has \(-\) scope above the universal modal while \(\exists\) scopes below. But, this is surprising: if \(-\) and \(\exists\) are both wrapped up in kein(e), they should have to scope at the same height. The quantifier keine Jacke would QR. If keine Jacke QRed above brauchst, both \(-\) and \(\exists\) would scope above. Alternatively, if keine Jacke QRed below brauchst, both would scope below.

One compelling line of solution for split scope with negative indefinites takes \(-\) out of the lexical entry for kein(e) and, indeed, out of the DP altogether. Rather, negation is re-localized in a separate abstract operator. Penka (2011), in particular, derives split scope from the structure in (71). Abstract NEG is interpreted as sentential negation, and takes scope above brauchst. The quantificational determiner is just a regular indefinite, as in (71), and can scope independently below brauchst.
The remaining explanandum is the morphology of *keine*. This form occurs only when negation is part of the interpretation, even though, at least on Penka’s analysis, it is not itself negative. In effect, Penka treats *keine* as a special indefinite determiner that surfaces in the presence of the NEG head, which she achieves formally with an Agree relation, as sketched.

Zeijlstra (2004) pursues a similar form of analysis for negative concord. Illustrating with Italian, the form *nessuno* appears to encode the meaning of a negative indefinite in (73), but when *nessuno* occurs below an overt negation, its negative contribution is no longer detectable, as in (74) (Zeijlstra 2004:39). Example (74) is interpreted nobody called (\(\neg \exists\)), not somebody called (\(\neg \neg \exists\)).

(73) **Negative indefinite (Italian)**
Nessuno ha telefonato.

N-body has called
‘Nobody called.’

(74) **Overt negation can co-occur (Italian)**
Non ha telefonato nessuno.

NEG has called n-body
‘Nobody called.’

Like *kein(e)* in Penka’s analysis, *nessuno* in Zeijlstra’s does not itself encode negation, but rather is a regular indefinite whose distribution is restricted to the environment of a higher negation. In (74), negation is encoded in *non* and, in (73), it is abstract, as in Penka’s datum. Zeijlstra again posits an Agree relation between the negation and the indefinite. He explicitly adopts an Upwards Agree model where a [uNEG] probe on the indefinite looks upwards in the structure to find the NEG head, specified [INEG], as its goal.
(75) **Structure for (73)**

\[ TP \text{ NEG} [\text{NEG}] [TP \exists_1 \text{uNEG} \text{ ha telefonato}] \]

Overall, Penka's and Zeijlstra's ideas about negative indefinites mirror the situation with pre-DP *only* under my proposal. With negative indefinites, the visible negative form occurs within the quantificational DP, but is just dummy Agreement morphology, tracking the presence of higher abstract NEG. The locus of interpreted negation is NEG. In a similar way, pre-DP *only* is just dummy Agreement morphology at F, tracking the presence of a interpreted ONLY. In this way, the proposal brings ONLY together with independent ideas about another Free Operator, which I have not been able to treat in detail in the thesis: negation.

4.7 **Next steps**

In the remainder of this chapter, I have two principle objectives. First, I will look at pre-vP *only* in greater depth, and show that the proposal accounts for a range of further scope facts with pre-vP *only*, including certain facts which remain mysterious under the type-ambiguity theory. Then, I will turn to cross-linguistic matters. As noted earlier, there are different syntactic mechanisms which could result in *only* appearing pre-DP, but being interpreted at the vP, such as movement. Reviewing previous work on different languages, I will collect together a range of arguments that *only* constructions cross-linguistically do involve a bipartite syntax of the sort I have proposed. From that perspective, the proposal can be seen as pursuing a null hypothesis that *only* in English involves that same syntax as parallel constructions cross-linguistically. The proposal has fit the semantic desideratum in with that syntax, again with cross-linguistic precedent to be discussed.

5 **Scope of pre-vP *only***

The proposal has, in important respects, unified pre-vP *only* and pre-DP *only*. Both involve the same bipartite structure and, in both cases, the interpreted operator scopes at the vP. My aim in this section is to look more closely at pre-vP *only*, and to consider whether the proposal can account for known interpretive differences between pre-vP *only* and pre-DP *only*. I will show that it can and, moreover, that a closer look at the data with pre-vP *only* reveals subtle properties that are better understood under my proposal than with type-ambiguity.

The discussion will center on data with this profile: *only* precedes an embedded vP, with the embedded clause introduced by a predicate with which *only* is non-commutative. In classic examples, again due to Taglicht (1984), *only* obligatorily takes narrow scope. Crnič (2014), however, showed that this is not always the case. His data suggest a surprising generalization where the size of the focus in the embedded clause affects the scope of *only* relative to the embedding predicate. In the classic examples where scope is frozen, *only* associates with an object DP. In Crnič's example, where scope is free, *only* associates with the entire vP. I show that frozen scope in the classic examples, and the interaction of scope with the size of the focus follow from my proposal.
5.1 Where scope is frozen

Earlier, we saw the example in (76), due to Taglicht, where only occurs just prior to the object DP, and scope ambiguity arises. ‘Only’ can take surface scope below require, or inverse scope above.

(76) **Pre-DP only: free scope**
John is required to learn only Spanish.

I proposed that, when the DP Spanish is focused, F attaches local to that DP. Pre-DP only, then, realizes F, which leaves the interpreted head, ONLY, abstract. Since ONLY is abstract, it can attach in different positions without affecting the surface string. The two structures for (76) are repeated below. Both map to the same surface string if F is realized, and each derives a different reading.

(77) **Structure 1 for (76): require > only**

\[
[TP \text{ John}_1 [vP \text{ t}_1 \text{ is required } [TP \text{ PRO}_1 \lambda \text{2 } [\text{vP t}_2 \text{ learn } [FP \text{ F Spanish}]]]]]
\]

(78) **Structure 2 for (76): only > require**

\[
[TP \text{ John}_1 [\text{ONLY } [vP \text{ t}_1 \text{ is required } [TP \text{ PRO}_1 \lambda \text{2 } [vP t_2 \text{ learn } [FP \text{ F Spanish}]]]]]]
\]

Now, let us consider what happens when the example is minimally modified so that only precedes the embedded vP, rather than the object DP. The relevant example is (79) and Taglicht observed that this example is unambiguous. Moving only to pre-vP position results in disappearance of the inverse scope reading, while the surface scope reading survives. This leads to the main puzzle for this subsection: why is scope frozen when overt only precedes the embedded vP in (79)?

(79) **Pre-vP only: frozen scope**
John is required to only learn Spanish.

In fact, the proposal accounts for the contrast between pre-DP only and pre-vP only. Under the proposal, overt only must realize one of two heads: F or ONLY. I proposed that the placement of F is constrained by the focus structure of the sentence such that F must attach directly to the object DP when that DP is focused. In (79), just like in (76), the focused constituent is the DP Spanish and, accordingly, F should attach just above that DP. Pre-vP only, then, cannot be a realization of F, but rather must be a realization of ONLY. It follows, in turn, that the position of pre-vP only disambiguates the attachment site of ONLY and thus disambiguates scope. The surface structure for (79) must be Structure 1 above, repeated in (80):

(80) **Structure for (76) (= Structure 1): require > only**

\[
[TP \text{ John}_1 [vP \text{ t}_1 \text{ is required } [TP \text{ PRO}_1 \lambda \text{2 } [\text{vP t}_2 \text{ learn } [FP \text{ F Spanish}]]]]]
\]
prior to all of the material in the matrix vP, required to only learn Spanish. Hence, only Structure 1 accords with constraints on the placement of F and is compatible with the surface string.

Now, what is predicted for scope? Structure 1 directly derives the uniquely attested surface scope reading. Deriving inverse scope would require some operation to apply to Structure 1 at LF to covertly turn Structure 1 into Structure 2. Then, Structure 1 would be pronounced, by something mimicking Structure 2 would be what the semantics interprets. As far as I can see, the covert operation would have to move ONLY itself from the embedded clause into the matrix clause above require. Moreover, in order to shift scope, that movement would have to not leave a trace. If there were a trace just above the embedded vP, it would be of type <st,st> to compose with the vP and [[only]] would semantically reconstruct back down to the trace position.

(81) Moving ONLY?

\[ \text{Moving ONLY?} \]

I assume that covert movement of ONLY is blocked for one or more of three reasons. First, being a head, ONLY may not be syntactically eligible to move on its own at all, especially to an adjunct position. Second, it may be that movement must leave a trace, rendering the movement chain in (81) ill-formed. Or, third, this derivation may be blocked for economy reasons: it is uneconomical to merge ONLY in the embedded clause in the narrow syntax and then covertly move it to the matrix clause, when it could merge directly in the matrix clause and avoid the movement step, as in Structure 2. Hence, Structure 1 is the only viable structure for (79) and, assuming that ONLY is interpreted where it is pronounced, that structure derives just the attested reading.

Overall, when the object DP is focused, pre-DP only realizes F, while pre-vP only must realize ONLY, assuming that F must attach at the DP. When F is overt, ONLY is abstract and there is thus ambiguity in where ONLY attaches, yielding multiple scope readings. When ONLY is overt, the position of only in the linear string signals the hierarchical attachment site of ONLY and thus disambiguates scope, assuming that ONLY is interpreted where it is pronounced. Scope is free when F is overt (in (76)) and frozen when ONLY is overt (in (79)). Phonological variability interacts with the semantics in just the right way to predict exactly the observed scope facts.

Because the assumption about the placement of F has now played a crucial role in deriving an empirical result, I want to expand on that. First, to appreciate the importance of the locality condition, note that (82) would have been a possible parse for (79) if F could attach as high as the edge of the embedded vP. In this structure, F attaches at that site and overt pre-vP only could, then, be a realization of F. This would leave ONLY abstract and free to take matrix scope, much as it was in the pre-DP only example, where only clearly did realize F.

(82) Blocked structure for (79)

\[ \text{Blocked structure for (79)} \]

To motivate rejecting this structure, we should ask: is there independent evidence that F cannot attach as high as the vP when the DP is focused? Some independent support for this conclusion comes from a domain which at first might look unrelated: namely, questions. As noted briefly earlier, Cable (2007) proposed an analysis of questions with a bipartite structure similar to the one I have pursued for focus.
The question in (83-a) has the underlying syntax in (83-b): there is an interrogative complementizer at the C head position, and another head Q more local to the wh-word. C is the primary locus of interpretation, and two heads Agree for Q features. The QP is what gets attracted to spec-CP. Wh-movement is re-envisioned as QP movement.

(83) **Cable’s syntax for questions**

a. Which language did John learn?

b. \[ \text{[CP}_C \text{[Q}_{(EPP)}] \text{[TP John}_1 \text{[vP}_t \text{learned [QP Q}_{(+)\text{] which language}]}} \]\]

An analogy may be drawn between ONLY and the interrogative C, F and Q, and the wh-word and the focus. Indeed, Cable extended his bipartite syntax from questions to certain focus fronting constructions, as we will see further later, and I have proposed a similar bipartite syntax in general. If we assume this analogy between heads, then we may gain insight into constraints on where F can attach by examining constraints on where Q can attach. For the case at hand, my assumptions about the placement of F precisely parallel the Q paradigm in (84):

(84) **Distribution of Q**

a. Which language did John learn.

b. *Learn which language did John.

The DP *which language* is analogous to the focused DP. That DP may move in (84-a), showing that Q may attach just above that DP. On the other hand, the vP cannot move in (84-b), showing that Q cannot attach that high. It would take us too far afield to probe further into the relationship between F and Q here, but I will discuss some further data in the final section of the chapter. For now, I take the question paradigm to constitute at least suggestive evidence for the locality constraints on F — and, with those constraints in place, frozen scope with pre-vP only is captured.

5.2 Where scope is free

Although this chapter is almost exclusively focused on examples where the object DP is focused, in this subsection, I will take a brief foray into examples where the size of the focus is larger. In particular, I will consider data where an entire embedded vP is under focus. By doing so, it is possible to test a further, more fine-grained prediction about scope which arises from the proposal. The basic idea is this. If the placement of F is constrained by locality, changing the size of the focus should change the placement of F. Whereas F must attach local to a focused DP, if the entire embedded vP is focused, then F should attach higher, to that vP. Pre-vP only then could realize F, and abstract ONLY then could take scope above the embedding predicate, yielding scope freedom. Changing the focus to be the vP should enable wide scope of only.

Evidence supportive of precisely this generalization was, in fact, noted in Crnić (2014). The most obvious way to force scope onto the entire verbal material is to take an intransitive vP, where all there is a verb and nothing more. This is achieved in Crnić’s (85). The associate of ‘only’ is show up. The focused constituent could be the verb itself, the VP (show up), or the vP (t show up). The choice between these is not relevant for the point, but I will assume the vP is focused for concreteness.
Changing the focus

This class requires you to only [show up]$_F$.

The proposal predicts this sentence to be ambiguous between the structures in (86-a) and (86-b). In each, F attaches just above the focused vP. If only realizes F, both structures map to the surface string in (85). Abstract ONLY attaches at the embedded vP in (86-a), and at the matrix vP in (86-b).

Available structure for (85)

a. $[TP \text{this class}_1 [vP \text{t}_1 \text{requires } [TP \text{you}_2 \text{[ONLY } [F [vP \text{t}_2 \text{show up}]]]]]]$

b. $[TP \text{this class}_1 [\text{ONLY } [vP \text{t}_1 \text{requires } [TP \text{you}_2 \text{[F [vP \text{t}_2 \text{show up}]]]]]]$

Crnič observed that (85) is, in fact, ambiguous, exactly in the way predicted. The two readings are paraphrased in (87). First, there is a pragmatically odd surface scope reading where (85) says that this class has a requirement that you do nothing but show up. The alternative, more pragmatically sensible reading has inverse scope. The sentence most naturally conveys that this class requires you to show up, but does not require anything more than that. In other words, the class does not require you to be a loafer, but allows you to be a loafer.

Available readings for (85)

a. “The requirement is that you show up and do nothing else.” (require > only)

b. “The only requirement is that you show up.” (only > require)

Overall, then, the proposal makes very fine-grained correct predictions about scope with pre-vP only. It predicts frozen scope when the DP is focused, and pre-DP only must realize ONLY, but free scope when the vP is focused, and pre-vP only may realize ONLY. The proposal predicts a surprising and apparently correct generalization that the location of the focus below pre-vP only affects the scope of ‘only’ relative to the higher embedding verb.\(^\text{11}\)

\(^\text{11}\)Note that Crnič further observed that pre-vP only is able to take scope above an embedding verb in examples such as (i). In this case, the focus is inside the object DP. The difference from (79) is that the DP contains a relative clause exhibiting Antecedent Contained Ellipsis (ACE) (cf. Nakanishi 2012 on even).

(i) To win the championship, we are required to only beat ONE team that our opponents are.

To allow for ACE, one team that our opponents are must move above required (and be interpreted de re). At the same time, only must be interpreted above require as well. The sentence paraphrases:

(ii) “For only one team that our opponents are required to beat are we required to beat it, too.”

Contrary to my conclusion, Crnič proposes on this basis that pre-vP only should freely take wide scope. He interprets (i) as evidence that only always forms a constituent with its associate (see Section 7): the object DP QRs to the complement position of only, and only DP then undergoes a second step of QR above require. The puzzle then becomes why a parallel derivation cannot derive wide scope of only in (79). For this, Crnič suggests that prosodic factors may be at play rendering the reading, while formally derivable, difficult to detect. My system is in the opposite predicament: it readily captures frozen scope in (79), but leaves wide scope in the ACE data unexplained. If F must attach local to the focus, it should attach just above the object DP in (i), in which case pre-vP only must realize ONLY. This would result in the illicit derivation in (ii). The FP QRs above require to license ACE, but then ONLY does not c-command the focus at LF, which it must.

(iii) a. $[TP \text{we}_1 [vP \text{t}_1 \text{are required } [TP \text{PRO}_1 \lambda 2 [\text{ONLY } [vP \text{t}_2 \text{learn } [TP \text{F DP}]]]]]]$

b. $[TP \text{we}_1 [\text{F DP} \lambda 3 [vP \text{t}_1 \text{are required } [TP \text{PRO}_1 \lambda 2 [\text{ONLY } [vP \text{t}_2 \text{learn } t_1]]]]]]$
5.3 Comparison with type-ambiguity

The proposal does better with pre-vP only than the type-ambiguity theory would. Because that theory lacks the additional F head whose placement interacts with the size of the focus, that theory would not predict any contrast between Taglicht's example and Crnič's example. It should predict either free or frozen scope in each, accounting for one data point, and mishandling the other. To see which data point is mishandled, let us consider Taglicht's example as a test case and ask: does the type-ambiguity theory predict frozen or free scope? The data point is repeated for reference:

(88) Recall: Taglicht's frozen scope
John is required to only learn Spanish.

The most obvious parse for (88) has only operate on the embedded vP interpreted as [only], correctly deriving require > only. Much as we saw with ONLY, deriving inverse scope would require sentential only to covertly move from its base position into the matrix clause — and covert movement of only would be blocked here for the same reasons I noted above with ONLY. The available LF, and unavailable alternative LF with covert movement are shown below:

(89) Only operates on vP as [only]

\[ TP \; \lambda 2 \; [vP \; t_2 \; \text{is required} \; \langle PRO_2 \; \lambda 1 \; [\text{only} \; [vP \; t_1 \; \text{learn Spanish}]]]] \]

(90) Moving only?

\[ TP \; \lambda 2 \; [\text{only} \; [vP \; t_2 \; \text{is required} \; \langle PRO_2 \; \lambda 1 \; [\text{only} \; [vP \; t_1 \; \text{learn Spanish}]]]]] \]

With only type-flexible, there are, however, other derivational possibilities. First, the word order is compatible with only operating on the VP, below the trace position of the subject, rather than on the vP. To compose with the VP, only would be type-lifted to operate on a property. I called that variant of only [only2] in Chapter 2. The LF in (91) converges with the first one in deriving surface scope. To derive inverse scope only would again have to covertly move on its own into the matrix clause, this time attaching to the matrix VP, and not leave a trace, as in (91-b). For the same reasons as before, such movement should be blocked.

(91) Only operates on VP

a. \[ TP \; \lambda 2 \; [vP \; t_2 \; \text{is required} \; \langle PRO_2 \; \lambda 1 \; [vP \; t_1 \; \text{learn Spanish}]]]] \]

b. \[ TP \; \lambda 2 \; [vP \; t_2 \; \text{only} \; [\text{VP is req.} \; \langle PRO_2 \; \lambda 1 \; [\text{only} \; [vP \; t_1 \; \text{learn Spanish}]]]]]] \]

In actual fact, the structure must be (iv-a), where only realizes F, and ONLY is above require; at LF, the object DP must QR above require but below ONLY, as in (iv-b).

(iv) a. \[ TP \; w_1 \; [\text{ONLY} \; [vP \; t_1 \; \text{are required} \; \langle TP \; \text{PRO1} \; \lambda 2 \; [FP \; F \; [vP \; t_2 \; \text{learn DP}]]]]] \]

b. \[ TP \; w_1 \; [\text{ONLY} \; [\text{DP \lambda 3} \; [vP \; t_1 \; \text{are required} \; \langle TP \; \text{PRO1} \; \lambda 2 \; [FP \; F \; [vP \; t_2 \; \text{learn t_3}]]]]]] \]

One possibility is that the locality constraint governing the position of F should be re-envision as a parsing heuristic: parse the string with F as low as possible to create a licit LF. In (79), there is a grammatical parse with F just above the DP and only realizing ONLY at the embedded vP edge. On the other hand, as just shown in (iii), there is no grammatical parse for (i) with F at that site. I leave better working out these ideas to the future.
The final derivation would go like this. Suppose that only attaches to the VP, and that the VP is Montague-lifted from type \(<e, st>\) to type \(<<est, st>, st>\). This is a property of quantifiers. Only could then type-lift to compose with that quantifier, to create a new quantificational meaning of the same type for only VP. These steps are shown:

(92) Montague-lifting VP
   a. \([VP] = \lambda x . \lambda w . x \text{ learned Spanish in } w\)
   b. \([VP_2] = \lambda F_{est, st} . \lambda w . F([VP])(w)\)

(93) Interpreting only
   \([only_0]^{ALT} = \lambda G_{<est, st>, st} . \lambda F_{est, st} . [only]^{ALT}(G(F))\)

(94) Interpreting only VP
   a. \([only \text{ VP}] = [only_0]^{ALT}([VP_2])\)
   b. \(\lambda F_{est, st} . [only]^{ALT}(F([VP]))\) (type \(<<est, st>, st>\))

Starting with the LF in (95-a), inverse scope could now derive not by moving only, but rather by moving the VP with only adjoined to it. Movement would target a scope site in the matrix clause above require and would leave a trace, as in (95-b). In order for movement to shift scope, the trace must be of property type (type \(<e, st>\)). Then, in effect, [only] is interpreted at the final landing site above require, while [VP] is semantically reconstructed to the position of t3. Since the VP contains no scope operator, reconstructing it has no detectable effect, while only > require is derived.

(95) Moving only VP
   a. \([TP \lambda 2 \text{ [vP t2 is required [vP t2 [ PRO₂ } \lambda 1 \text{ [vP learn Spanish]]]]]])\)
   b. \([TP \lambda 2 \text{ [[only [vP learn Spanish]]} \lambda 3 \text{ [vP t2 is required [vPRO₂ } \lambda 1 \text{ [vP t1 t3]]]]]\))

The overall prediction hinges on whether this derivation is truly viable: if so, the type-ambiguity theory would generate surface and inverse scope; if not, just surface scope is predicted. Even if only were type-flexible, I believe that this derivation would not be available. To re-iterate, an element (in this case, the VP) is Montague-lifted, a type-lifted only composes with the Montague-lifted VP, and the effect is for [only] to be interpreted high, while the VP semantically reconstructs to the trace position. This derivation closely resembles one I considered for split scope in Test 1: a DP was Montague-lifted, a type-lifted only composed with it, and the effect was for [only] to be interpreted high, while the basic meaning of the DP semantically reconstructed. I ruled out that derivation empirically on the basis of Condition C effects. Assuming the derivation just sketched for pre-vP only is, likewise, unavailable, the type-ambiguity theory would predict frozen scope — capturing Taglicht's example and mishandling Crnčić's. Although I did not illustrate with Crnčić's example, changing the size of the focus would have no effect on the available derivations for pre-vP only: the derivations discussed here with Taglicht's example would extend wholesale to Crnčić's.
5.4 Local summary

This section has addressed cases where overt only precedes a vP in an embedded clause, focusing on the perceived scope of 'only' relative to the embedding predicate. In doing so, I showed that the proposal makes fine-grained predictions, which are borne out. In the primary profile of data under consideration, the object DP is focused, in which case F attaches local to it, and pre-vP only realizes ONLY, disambiguating scope. When the entire vP is focused, F attaches to it, and pre-vP only may realize F, allowing abstract ONLY to scope higher. The proposal's predictions are not matched by the type-ambiguity theory, which cannot explain differences in scope depending on the location of focus, and most likely predicts frozen scope across cases. The proposal not only captures pre-vP only, but it appears to have more success than the type-ambiguity theory.

6 Cross-linguistic evidence

Now, I turn to cross-linguistic matters with three aims: to review the antecedents for my proposal, to establish independent evidence for the syntactic and phonological assumptions of the system, and to present a unified perspective on variability in 'only' constructions within and between languages to set up future inquiry. The discussion is organized around the goal of justifying assumptions. The key assumptions are enumerated here, based on discussion in Section 4:

1. Only constructions involve two-heads.
2. Those two heads Agree for [ONLY] features.
3. The morphology varies in how it realizes the two heads.

I will bring together arguments from Dutch (Barbiers 2014, Bayer 2016), Vietnamese (Höle 2013, Erlewine to appear), Hungarian (Horvath 2007, Cable 2007, 2010), and Korean (Lee 2004). Section 6.1. will focus on Assumptions 1 and 3, and Section 6.2 will focus on Assumption 2.

6.1 Assumptions 1 and 3

I begin with direct phonological evidence for a bi-partite structure. Whereas English shows either pre-vP only or pre-DP only, but not both together, colloquial Dutch and Vietnamese wear a bipartite structure on their sleeve: only constructions may involve an overt vP-level morpheme and an overt DP-level morpheme at the same time. A unified picture will emerge on how to think about variation in the realization of only, both within and between languages. There is always the same bipartite structure (and, in my view, the semantics). The locus of variation is in the mapping from syntax to phonology. The broad distribution of English only becomes just one piece of a broader puzzle.

6.1.1 Dutch

We saw in English that pre-vP and pre-DP only cannot co-occur, and I proposed that this is due to a constraint of English morphology. If this constraint is not universal, we should find languages which
do simultaneously spell out both heads. Dutch is one such language. As observed in Barbiers (2014) and further discussed in Bayer (2016), only constructions in Dutch can be realized in three ways:

(96) Three possible patterns
   a. Maar een boek ken ik
      only one book know I
      'I know only one book.'
   b. Een boek ken ik maar
      one book know I only
   c. Maar een boek ken ik maar
      only one book know I only

Like English, Dutch can express only with an overt morpheme maar at the vP edge, as in (96-a), or with that same morpheme local to the object DP, as in (96-b). In addition, the third option that was unacceptable in English is reported to be attested in Dutch, at least colloquially: vP-level maar can co-occur with DP-level maar. Barbiers and Bayer take the doubling pattern as evidence for a bipartite structure along the lines of the one I have proposed. My version is (97). Though the relevant movements are not shown in (97), note that the sentences in (96) are verb-second, with the FP (maar) een boek moved to the specifier of the CP, and the verb ken moved to C.

(97) Bi-partite structure common to (96-a)-(96-c)

\[
\begin{align*}
[CP C [TP ik2 [onlyP [vp [fp [fiONLY] [dp een boek]] ken]] v] ONLY[uONLY] T]]
\end{align*}

With this in place, variability both within Dutch and across Dutch and English is simply a matter of surface phonology. Dutch, like English, has a Vocabulary Item which specifies a realization for any interpretable or uninterpretable [ONLY] feature, in this case as maar. The grammaticality contrast between double maar in Dutch and double only in English is due to idiosyncratic phonological properties of the two languages. Whereas both languages require that at least one of the two heads in the bi-partite structure be realized, Dutch does not require that at most one be realized.

(98) Morphological hypothesis (Dutch)
   a. Vocabulary Items
      [alphaONLY] → maar
   b. Constraints on realization
      If two [alphaONLY] features are in a probe-goal relation, at least one is realized.

If maar realizes [ionly] on ONLY, the string is (96-a). If maar realizes [uonly] on F, the string is (96-b). And, if maar realizes both features, the string is (96-c). All three possibilities accord with the constraint in (98-b). In a language with direct evidence for a bi-partite structure, there is still morphological optionality in how the structure is realized.
6.1.2 Vietnamese

If only constructions involve two distinct heads, we would also expect to find a language where the two heads are realized by distinct morphemes. This is the case in Vietnamese, with the relevant paradigm given in (99), due to Høle (2013) and Erlewine (to appear).

(99) **Three possible patterns**

a. Nam **chi** mua [cuon sách].
   Nam only₁ buy CL book
   ‘Nam only bought one book.’

b. Nam mua **moi** [cuon sách].
   Nam buy only₂ CL book

c. Nam **chi** mua **moi** [cuon sách].
   Nam only₁ buy only₂ CL book

The form which occurs pre-vP is **chi**, and the one which occurs pre-DP is **moi**. **Ci** can occur with **moi** left unrealized; **moi** can occur with **chi** left unrealized; and both can occur simultaneously. The doubling pattern in (99-c) is again supportive of a bi-partite structure, as in (100). Note that the FP remains in situ in (99-a)-(99-c), whereas the FP fronted to first position in Barbiers’ Dutch paradigm. The Vietnamese data show that the presence of two heads is not somehow linked to movement.

(100) **Bi-partite structure common to (99-a)-(99-c)**

\[
[Tp \text{Nam}_1 [\text{only}_P \text{ONLY}_{[\text{ONLY}_I]} [\text{vP t}_1 \text{ mua } [\text{FP F}_{[\text{ONLY}_U]} [\text{DP cuon sách}]]]]]
\]

As before, I propose that the only difference between Vietnamese and other languages lies in the phonology. Vietnamese, like English and Dutch, requires at least one of ONLY and F to be realized and, like Dutch, it carries no prohibition against both being realized simultaneously. The difference from English and Dutch is in the vocabulary items: instead of having a single vocabulary item for any \([\text{ONLY}_I]\) feature, there are separate vocabulary items for \([\text{ONLY}_I]\) and \([\text{ONLY}_U]\). \([\text{ONLY}_I]\) on ONLY maps to **chi**, and \([\text{ONLY}_U]\) on F maps to **moi**.

(101) **Morphological hypothesis (Vietnamese)**

a. **Vocabulary Items**
   \([\text{ONLY}_I]\) → **chi**
   \([\text{ONLY}_U]\) → **moi**

b. **Constraints on realization**
   If two \([\text{ONLY}_I]\) features are in a probe-goal relation, at least one is realized.

6.1.3 Local summary

This section has presented direct evidence for a bipartite structure from Dutch and Vietnamese. In those languages, 'only' constructions may exhibit two overt morphemes simultaneously, one at the
vP (the site of ONLY) and one at the DP (the site of F). Variability in how the bipartite structure is realized within a language comes from flexibility built into the language's morphological constraints. For Dutch and Vietnamese, the morphology requires at least one head to be overt — which leaves the speaker the choice to realize ONLY, F, or both.

I proposed that differences between languages again stem from the morphology, and identified two dimensions of cross-linguistic variability. First, some languages have a single vocabulary item for any [ONLY] feature (Dutch, English), while others have separate vocabulary items for [iONLY] and [uONLY] (Vietnamese). Second, some languages require at least one head to be overt (Dutch, Vietnamese), while others require exactly one head to be overt (English).

Of course, I have studied just a tiny fragment of data. It would a worthwhile future project to study a broad typology of cross-linguistic focus constructions from the perspective suggested here. On the one hand, such an investigation could assess certain specific predictions. For instance, we might expect to find a language which requires both heads in the bipartite structure to always be overt. On the other hand, we could ask more generally: can the full range of variability in how only is expressed across languages be understood as different ways of mapping an underlying bipartite structure to overt phonology? I will take that to be the null hypothesis.

6.2 Assumption 2

All of the discussion so far has assumed that the two heads in the bipartite structure Agree for the operator-specific feature [ONLY]. Now, I want to provide direct evidence for that assumption. Horvath (2007) previously provided an argument that operator-specific features exist from focus fronting in Hungarian. In Hungarian, the associate of 'only' obligatorily fronts to a pre-verbal position, as in (102). The string in (102-a) is acceptable, with a fogardásról ('only the reception') pre-verbal. The alternative strings in (102-b) and (102-c), with a fogardásról in situ are unacceptable.

(102) **Obligatory fronting with csak ('only')**

a. Mari csak a fogardásról késett el.
   Mary-NOM only the reception-from late-was away
   'Mary was late only for the reception.'

b. *Mari elkésett csak a fogardásról.
   Mary-NOM away-late-was only the reception-from

c. *Mari csak elkéselt a fogardásról.
   Mary-NOM only away-late-was the reception-from.

First of all, if fronting is feature-driven movement, we can conclude that the construction involves a covert head just above the vP. That head must probe for some feature involved with focus, and bear an EPP feature to necessitate movement. Assuming csak is pre-DP, it must realize a separate head local to the DP. This line of reasoning has corroborated for Hungarian our earlier general conclusion

---

12I present an adaptation of Horvath's argument to the system under development, which bears a close similarity to hers. She took fronting constructions to involve a clausal head and a head just above the focus, each bearing an El-Op feature (i.e. 'exhaustive identification'). I adopt a parallel bipartite syntax, and refer to the feature as [ONLY].
that only constructions involve two heads, ONLY and F. Csak realizes F, and the structure shown is immediately prior to movement of csak a fogardásról to the specifier of ONLY.

(103) **Bi-partite structure for (102-a)**

\[ [\text{TP Mari}_1 \ [\text{ONLY} \ [? \ EPP] \ [\text{VP t}_1 \ \text{késett el} \ [\text{FP F} \ [\text{DP a fogardásról}]]]]] \]

Not only does fronting require two heads, but restrictions on fronting make it possible to diagnose the feature specifications of those heads. Suppose first that ONLY probed for a general focus feature borne directly by the focused constituent. To not be confused with the F head, I indicate that feature as Foc in (104). In that case, ONLY would have no properties specific to csak. It would enter into an Agree relation with the DP bearing Foc, and the EPP feature on ONLY would trigger fronting of the DP. Csak, which realizes F, would be pied-piped, given some suitable mechanism for pied-piping.

(104) **Adding features (to be revised)**

\[ [\text{TP Mari}_1 \ [\text{ONLY} \ [\text{Foc}(), \ EPP] \ [\text{VP t}_1 \ \text{késett el} \ [\text{FP F} \ [\text{DP a fogardásról}][\text{Foc}(+)]]]]] \]

While this captures the paradigm in (102), Horvath's key point is that this set up cannot be correct, since it predicts fronting to be more ubiquitous than it is: it predicts that all focused constituents should front, and this is not the case. The problem is illustrated in (105), which involves még ... is ('even') instead of csak. Despite being focused, the associate of még ... is cannot front.

(105) **Impossibility of fronting with még ... is ('even')**

a. Mari elkésett még az esküvőjéről is.
   Mary-NOM away-late-was yet the wedding-her-from also
   'Mary was late even for her wedding.'

b. *Mari még az esküvőjéről is késett el.
   Mary-NOM yet the wedding-her-form also late-was away

Taken together, the paradigms in (102) and (105) establish a correlation between two properties: (a) whether or not the FP fronts to the specifier of ONLY, and (b) whether or not F is realized as csak or még ... is.\(^{13}\) Since an EPP feature on ONLY triggers fronting, (a) is a property of ONLY, and (b) is a property of F. The correlation between them shows that ONLY must be able to “see” F. In particular, ONLY must be able to see what operator has its realization at F. This can only be achieved if operator-specific features, such as [ONLY] exist; if ONLY and F both bear an operator-specific feature; and if ONLY and F are in a probe-goal relation.

In this way, we arrive at the following conclusion. ONLY in the csak structure does not probe for a general Foc feature. Rather, ONLY probes for [ONLY], and Agrees with F. The structure in (103) must be completed as (106). With Agreement between ONLY and F, the EPP feature on ONLY triggers fronting of the entire FP (csak a fogardásról) to the specifier of ONLY.

\(^{13}\)Although még ... is is clearly itself complex, I will treat it as a simple instantiation of a single lower head to ease discussion. This is admittedly idealized, and determining the composition of még ... is is a worthy enterprise, which I will not undertake. I refer the reader to Crnič (2014) for a discussion of how to decompose 'even'.
Now, there is no expectation that fronting will be triggered with \( m\dot{e}g \ldots \) is. If, contrary to my hypothesis, the lower head is the locus of interpretation, the structure with \( m\dot{e}g \ldots \) is could contain no higher head at all. Alternatively, if it does contain a higher head, it would have to be a different one, \textsc{EVEN}, with a different feature specification from \textsc{ONLY}. \textsc{EVEN} would bear an \([\textsc{EVEN}]\) probe, and there is no reason to expect it would also bear an EPP feature. The structure would be (107):

\[
[TP \text{Mari}_1 [\textsc{EVEN}[\text{EVEN}(\cdot)][VP \text{elkésett} [F[u\textsc{EVEN}(\cdot)] [PP \text{az esküvőjéről}]]]]]
\]

By reasoning about the properties of fronting, we have drawn the following conclusions for \textit{csak} constructions in Hungarian, following Horvath: they must involve two heads, those two heads are specified for \([\textsc{ONLY}]\) features, and the two heads are in an Agree relation. In this way, we have arrived at a structure for \textit{csak} just like the one I proposed for English \textit{only}:

\[
[TP \text{John}_1 [\textsc{onlyP} \textsc{ONLY}[\text{ONLY}][VP \text{t1 learned} [FP \text{one language}]]]]
\]

The sole distinction lies in whether or not the higher head bears an EPP feature. In English, there is no EPP feature, so Agreement between the higher and lower heads does not trigger fronting; in Hungarian, the higher head does bear an EPP feature. In fact, there is evidence from languages which exhibit fronting for cross-linguistic variability in which operator-specific features co-occur with an EPP feature. Overall, I extend to English the syntax for \textit{‘only’} in focus fronting constructions.

### 6.3 A semantic precedent

The proposal adopts the syntactic insights from the authors discussed so far, and combines them with a particular semantic claim: that the higher head is always interpreted, even when the lower, DP-level head is realized. To end our tour of cross-linguistic data, I want to flag Lee’s (2004) proposal for Korean \textit{man} (‘only’), which makes claims not only about syntax, but also semantics, and is the most direct antecedent for my own proposal. Lee observed a scope puzzle with \textit{man}, namely that when an object DP marked with \textit{man} scrambles above the subject, the scope of \textit{‘only’} relative to the subject depends on what other morphological marking the DP bears. Consider (109):

\[
\begin{align*}
\text{Man with case marking} & \\
\text{Mary-\textit{man-ul}}_1 [\text{motun-salam-i \textit{t1 semmwul-ul cwuessta}] } & \\
\text{Mary-\textit{only-Acc} every-person-Nom t1 gift-Acc gave} & \\
\text{‘(Lit.) Only Mary, everyone gave a gift t’}
\end{align*}
\]

The DP of concern is \textit{Mary}, which is marked with \textit{man} and, in addition, bears the accusative case marker \textit{ul}. That DP has scrambled from its base position in the double object construction to a position...
above the universal subject. Despite scrambling, the scope of ‘only’ is frozen below ‘every’ in this data point. The interpretation in (110-a) is available, but not (110-b).

(110) Scope in (109): frozen
a. ‘Everyone gave a gift to Mary and to no one else.’
   \( (\text{every} > \text{only}) \)

\begin{tabular}{ll}
  b. & ‘Mary is the only person that everyone gave a gift to.’ \( (\text{*only} > \text{every}) \) \\
\end{tabular}

If the DP is not marked for case, but rather bears a postposition, the pattern changes. The sentence in (111) illustrates, forming a minimal pair with (110). In this case, the DP with man is marked with ekey (‘to’). Just like in (110), the DP has scrambled above the universal subject. But now, scope ambiguity arises, as in (112). ‘Only’ is free to take scope above or below ‘every’.

(111) Man with post-position
Mary-ekey-man \[ \text{motun-salam-} t_1 \text{ senmwul-ul cwuessta} \]
Mary-to-only every-person-Nom t_1 gift-Acc gave
‘(Lit.) Only to Mary, everyone gave a gift t.’

(112) Scope in (111): free
a. ‘Everyone gave a gift to Mary and to no one else.’
   \( (\text{every} > \text{only}) \)

\begin{tabular}{ll}
  b. & ‘Mary is the only person that everyone gave a gift to.’ \( (\text{only} > \text{every}) \) \\
\end{tabular}

To explain the effect of morphological marking, Lee proposes that man does not itself contribute the meaning of ‘only’. Rather, man is a dummy Agreement marker. The meaning comes from an abstract ONLY heading a functional projection which, in Korean, can occur at different positions in the structure. Morphological marking tracks the scope of ONLY, since the order of morphemes in the word mirrors the order of functional projections in the syntactic hierarchy, per Baker’s (1985) Mirror Principle. In (109), the occurrence of man inside of ul shows that the projection headed by ONLY occurs below the projection in which accusative case is assigned, assuming that accusative case assignment occurs in the specifier of an AgrOP. AgrOP is below the position of the subject and ONLY, being below the AgrOP, must be below the subject as well: ‘only’ takes scope below ‘every’. In (111), man occurs outside of ekey. If postpositions are generated in the vP, this means that ONLY takes scope somewhere above the vP. This is compatible with ONLY attaching above the vP, but below the subject, or at a higher position above the subject. Scope ambiguity is thus predicted.

My proposal for pre-DP only extends Lee’s core insight for Korean to English: pre-DP only is dummy Agreement morphology, while abstract ONLY is interpreted. Note I do formulate the lexical entry for ONLY differently than Lee. Lee’s ONLY corresponded, in effect, to a type-lifted variant of [only], and she assumed that a trace of the scrambled DP composed with ONLY. I take ONLY to uniformly operate on the vP as [only] — in accord with the SIH.

6.4 Concluding cross-linguistic evidence

In this section, we have seen cross-linguistic support for the syntactic and phonological assumptions behind the proposal from Section 2. At least in some languages, ‘only’ constructions involve two
heads (optionally both overt in Dutch and Vietnamese), those two heads Agree for [ONLY] features (diagnosed by restrictions on focus fronting in different languages), and variability in the realization of ‘only’ within and between languages can be understood as variability in which of the two heads gets realized (offers a unified perspective on Dutch and Vietnamese, as well as English).

The proposal I made in Section 2 can thus be seen as importing these cross-linguistic insights about the syntax of only into English, and building around this syntax a semantics which explains the ellipsis and scope facts from Cases 1 and 2. To re-iterate, the key semantic claim is that ONLY is interpreted, while F is inert. This distribution of semantic labor also mirrors an earlier proposal for another language, being Lee’s proposal for Korean man. Overall, the approach I have taken fits English only seamlessly into the broader cross-linguistic typology.

7 Comparison with alternatives

I now compare my proposal to other theories of only which have been advanced in the literature separate from the traditional type-ambiguity theory. Specifically, I will consider the idea that only is uniformly interpreted as a two-place operator (Wagner 2006, building on Horn 1969, Chomsky 1976, Drubig 1994), and the idea that only decomposes into negation and an exceptive, like French ne...que (von Fintel & Iatridou’s 2007)”. I will show that neither theory can handle the full range of data in Tests 1 and 2 in their present form, though it may be possible to modify the second theory to achieve the necessary empirical coverage, and integrate its insights into my own proposal. Still, if my proposal is correct, a number of puzzles arise as to how to account for data which seemingly argue for other theories. I will flag those puzzles, but not solve them.

7.1 Two-place only

Wagner (2006), like this chapter, takes pre-DP only and pre-vP only to be interpreted in the same way. Nonetheless, he makes a different claim about how exactly only is interpreted. Whereas I take the interpreted operator to uniformly be propositional and scope at the vP, he proposes that only uniformly takes two arguments, the first of which is provided by the focused DP. In effect, we disagree on the baseline case: I take pre-vP only to be the baseline, and Wagner takes the baseline to be pre-DP only. Wagner’s approach inherits the problems I raised earlier for the traditional type-ambiguity theory with respect to pre-DP only, and extends them to pre-vP only.

7.1.1 Illustrating the approach

To get a feel for the “two-place only” theory, let us consider for now just the basic sentence pair provided at the outset of the chapter, repeated below.

(113) Pre-vP only vs. pre-DP only
  a. John only learned oneF language. = (1-a)
  b. John learned only oneF language. = (1-b)
Following Horn (1969), Wagner provides a two-place entry for *only*. While the entry he gives is tailored to compose with referential DPs, an appropriate generalization is (114). I will call this denotation $\text{only}_{\text{new}}$. $\text{only}_{\text{new}}$ takes as its first argument an element $x$ of any type $\alpha$. Its second argument is then some $f$ of type $<\alpha,\text{st}>$, i.e. a function which outputs a proposition when saturated with an argument of type $\alpha$. $\text{only}_{\text{new}}$ presupposes that $f(x)$ is true, and negates non-weaker alternatives. $\text{only}_{\text{new}}$ has a polymorphic type, able to compose with different kinds of arguments.\(^{14}\)

\begin{equation}
\text{Two-place only}
\end{equation}

\[\text{only}_{\text{new}} = \lambda x_{\alpha} \cdot \lambda f_{<\alpha,\text{st}>} \cdot \lambda w : f(x)(w) \cdot \forall a \in \text{ALT}(x) \ [f(a)(w) \rightarrow f(x) \subseteq f(a)]\]

This meaning for *only* is well-suited to analyze pre-DP *only*, provided that *only* operates directly on the quantificational DP. The predicted output is somewhat different from what we saw with $\text{only}_{3}$, however. $\text{only}_{3}$ operated on a quantifier to output a new quantifier:

\begin{equation}
\text{Composing with only}_{3}
\end{equation}

a. $\text{only}_{3}(\text{ALT})(\text{one language})$

b. $= \lambda f_{\text{est}} \cdot \text{only}_{3}(\text{ALT})(\text{one language})(f)$

If $\alpha$ in the statement of $\text{only}_{\text{new}}$ is filled in as type $<\text{est},\text{st}>$, the overall type for $\text{only}_{\text{new}}$ is, in effect, type $<\text{est},\text{est},\text{st},\text{st}>$. Saturating the first argument of $\text{only}_{\text{new}}$ with *one language* thus yields an $<\text{est},\text{st},\text{st},\text{st}>$ meaning — a property of properties of quantifiers.

\begin{equation}
\text{Composing with only}_{\text{new}}
\end{equation}

a. $\text{only}_{\text{new}}(\text{one language})$

b. $= \lambda F_{<\text{est},\text{st},\text{st}>} \cdot \lambda w \cdot \forall g \in \text{ALT}(\text{one lang}) \ [F(g)(w) \rightarrow F(\text{one lang}) \subseteq F(g)]$

(defined iff $F(\text{one lang})(w)$)

Because *only* operates on the DP, the LF for (1-a) would be (117), where *only one language* QRs. To create an interpretable structure, two steps of QR are necessary. The first step of QR leaves a trace of type $e \ (t_2)$, which can compose with the verb *learned*. The second step of QR must leave a trace ($t_3$) of quantifier type, i.e. type $<\text{est},\text{st}>$. Abstracting over that trace derives a property of quantifiers (type $<\text{est},\text{st},\text{st}>$) — the right sort of argument for *only one language* in (116).

\begin{equation}
\text{LF for (1-b)}
\end{equation}

\[\text{TP [TP John } \lambda 1_{\text{vp}} \ [\text{only one language}] \lambda 3_{\text{vp}} t_3 \lambda 2_{\text{vp}} t_1 \ \text{learned } t_2]]\]

The property derived with QR is (118), roughly the set of quantifiers containing *that John learned* $x$ at a given world. Applying *only one language* to the derived property triggers the presupposition that the quantifier *one language* is in that set at the evaluation world, i.e. it is presupposed that John learned (at least) one language. The assertion says that alternative quantifiers of the form *n languages* are not in the set, yielding the exclusive entailment. This is the correct meaning.

\footnote{Thanks to Michael Wagner for discussion on this point.}

271
The derived property
\[ \lambda f_{est} . \lambda w . f(\lambda x . \lambda w . \text{John learned } x \text{ in } w)(w) \]

Now, how is (1-a) analyzed? Under the original type-ambiguity theory, pre-vP only was interpreted differently from pre-DP only, as \([\text{only}]\), instead of \([\text{only}_3]\). Whereas pre-DP only operated on the DP, pre-vP only operated on the vP. For Wagner, however, only is always interpreted as \([\text{only}_{new}]\), and operates on the DP in both cases. The base structure for (1-a) is:

Base structure for (1-a)
\[
\text{[TP John } \lambda 1 [\text{only}_P \text{ only } [\text{vP } t_1 \text{ learned one}_F \text{ language}]]]
\]

In this structure, only occurs at its scope position at the vP edge, while the DP one language is inside the vP. One problem with this structure is that \([\text{only}_{new}]\) does not have all the necessary arguments. It requires two arguments, and its sole argument at the moment is the vP. In order to provide a restrictor argument for only, Wagner, building on Chomsky (1976) and Drubig (1994), proposes that the DP one language covertly moves to the complement of only. In this case, the movement must proceed in two steps to yield an interpretable structure. First, one language must QR (leaving a type e trace, \(t_2\)), and then undergo focus movement to the complement of only (leaving a quantifier type trace, \(t_3\)). The structure is (120). Note that, when focus movement takes place, the binder index triggering lambda abstraction is inserted not below one language, but below only one language on the clausal spine.

Structure after focus movement
\[
\text{[TP John } \lambda 1 [\text{only}_P [\text{only one language}] \lambda 3 [\text{vP } t_3 \lambda 2 [\text{vP } t_1 \text{ learned } t_2]]]]
\]

The structure in (120) for (1-a) is identical to the structure in (117) for (1-b) and, accordingly, the example with pre-vP only is interpreted identically to the example with pre-DP only. In both cases, \([\text{only}_{new}]\) operates on the quantificational DP to yield a meaning of type \(<<<\text{est},\text{st}>,\text{st}>,\text{st}>>\).

7.1.2 Problem 1: pre-DP only

Because \([\text{only}_{new}]\) operates on the DP, Wagner’s approach inherits the problems raised in Cases 1 and 2 for pre-DP only under the original type-ambiguity theory. In Test 1, I provided a direct argument that pre-DP only cannot compose with the DP to create a meaning of type \(<<<\text{est},\text{st}>,\text{st}>,\text{st}>>\). I entertained that sort of meaning in Section 2.2.1 as a way of deriving split scope in (121).

Recall: split scope prototype
You have to publish [only three papers].

(only > have > three papers)

Although I did not consider \([\text{only}_{new}]\) itself, I derived the same result in a different way, under the original type-ambiguity theory. Under that theory, variants of only apply to arguments of various types to return a meaning of the same type. I entertained Montague-lifting three papers from type \(<>\text{est},\text{st}>>\) to type \(<<<\text{est},\text{st}>,\text{st}>,\text{st}>>\). \([\text{only}_2]\) then yields a meaning for only one language of type \(<<<\text{est},\text{st}>,\text{st}>,\text{st}>>\). The meaning in (123) is equivalent to the one \([\text{only}_{new}]\) would predict.
Lifting one language

a. \([\text{one language}] = \lambda f_{est} \cdot \lambda w \cdot \exists X [(X)\geq 1 \land \text{language}^\star (X)(w) \land f(X)(w)]\]

b. \([\text{one language}_2] = \lambda F^{<<est,st>,st>} \cdot \lambda w \cdot F([\text{one language}])(w)\]

Interpreting only

\([\text{only}_5] = \lambda G^{<<est,st>,st>,st>} \cdot \lambda F^{<<est,st>,st>} \cdot \text{only}([G(F)])\]

Composing with onlys

a. \([\text{only one language}] = [\text{only}_5][\text{one language}]\]

b. \(= \lambda F^{<<est,st>,st>} \cdot \text{only}([F([\text{one language}])]\)

Interpreted as type \(<<est,st>,st>,st>, only three papers in (121) could QR as a constituent above have and leave a quantifier type trace below have. Only would then be interpreted at the final landing site, while the quantificational DP would be semantically reconstructed down to the trace position.

Possible LF for (121)

\([TP w [\text{vP}_3 [\text{only }3 \text{ papers}]] [\text{vP}_2 \text{ you} 3 \text{ have} [TP w' [\text{vP}_1 \text{ PRO } 3 \text{ publish } t_1]]]])]

I argued that this analysis was not viable, based on Condition C effects, which showed that a DP that is syntactically high as part of only DP cannot semantically scope lower. For the full argument, I defer to Section 2.3. What is crucial here is that the theory with \([\text{only}_\text{prev}]\) seems to commit to only DP being of type \(<<est,st>,st>,st>\) when the DP is quantificational and, as such, inherits the problems raised in Section 2.3. Moreover, even if we granted the \(<<est,st>,st>,st>\) meaning, the ellipsis data in Test 2 would not be predicted. Because only in the first sentence operates on the DP and only DP QRs as a unit, any possible antecedent constituent in the LF for the first sentence either contains both only three papers or neither of them. There is no antecedent of the form publish (at least) three papers and thus no way to license ellipsis of \(\Delta\).

Recall: ellipsis licensing

a. You have to publish only three papers.

b. And now you’re telling me that you can’t \(\Delta\)? \((\Delta = ‘\text{publish (at least) three papers’)\)

Overall, then, with respect to pre-DP only, Wagner’s approach inherits the problems for the original type-ambiguity approach. Only operates on the DP when, in fact, the arguments in Cases 1 and 2 demonstrated that pre-DP only must instead operate on the vP.

Problem 2: pre-vP only

In the original type-ambiguity theory, pre-vP only was analyzed differently from pre-DP only. Whereas pre-DP only operated on the DP, pre-vP only operated on the vP as \([\text{only}]\). As a result, pre-vP only was immune to the problems I raised for pre-DP only. Pre-vP only allows for similar scope readings to pre-DP only, and can also license VP-ellipsis in a similar way — but with pre-vP only operating on the vP, these data were easily handled.
(127) **Scope and ellipsis facts with pre-vP only**

a. You are only required to publish three papers.  
   \[ (\text{only} > \text{have} > \text{three papers}) \]

b. And now you’re telling me that you can’t \( \Delta \)?  
   \[ (\Delta = \text{‘publish (at least) three papers’}) \]

Split scope in (127-a) derived syntactically: *only* attached at the matrix vP, while *three papers*, interpreted separately from *only*, QRed within the embedded clause. There was no need for a type \( <<\langle\text{est, st}\rangle,\text{st},\text{st}\rangle \) meaning. The possibility for a syntactic scope split also explained ellipsis: if *only* attaches alone in the matrix clause, and *three languages* attaches alone in the embedded clause, the embedded clause contains a constituent with the meaning *publish (at least) three papers*.

In Wagner’s approach, pre-vP *only*, like pre-DP *only*, composes directly with the DP and, as such, the problems I raised for pre-DP only extend to pre-vP only, as well. If *only* is always interpreted as \[ [\text{only}_{new}] \], the structure for (127-a) must be (128).

(128) **Possible LF for (127-a)**

\[ [\text{TP w} [\text{vP}_3 [\text{Only}_3 \text{papers}] \text{vP}_2  \lambda_2 [\text{vP}_1 \text{you}_3 \text{have} [\text{TP w'} [\text{vP}_1 \text{PRO}_3 \text{publish}_t]]]]]] \]

*Only* attaches just above the matrix vP. *Three papers* first undergoes a step of QR to the edge of the embedded vP (leaving a type e trace), and then undergoes focus movement up to the complement of *only* (leaving a quantifier type trace) to provide *only* with its restrictor argument. This LF is exactly the same as the one in (125) with pre-DP *only*. \[ [\text{only}_{new}] \] will compose with *three languages* to create a problematic \( <<\langle\text{est, st}\rangle,\text{st},\text{st}\rangle \) meaning — and ellipsis will remain a mystery. By moving up to *only*, *three papers* leaves the embedded VP, which is now *publish t*, which is not the right meaning to serve as antecedent for \( \Delta \) in (127-b).

Wagner (2006) in fact noted the possibility of split scope with pre-vP *only* in a footnote (see his fn. 20; the observation is attributed to Daniel Büring and Yael Sharvit). He suggested that focus movement of the DP obligatorily reconstructs either semantically or syntactically. Having already argued that semantic reconstruction is not viable, I note further that syntactic reconstruction does not seem viable either. If *three papers* were syntactically returned to the embedded clause, *only* would be left without a restrictor, and the structure would not be interpretable. With pre-vP *only*, just like with pre-DP *only*, *only* and the DP must be interpreted at distinct syntactic positions, but the current theory does not allow for a syntactic scope split, as far as I can see.

### 7.1.4 Problem 3: scope of pre-vP only

A further challenge to Wagner’s theory, previously observed in Crnić (2014), comes from the scope of pre-vP *only* in (129). This example, discussed in earlier, is unambiguously interpreted with *only* scopes below *require*. Crnić points out that Wagner’s theory should be able to generate the unavailable inverse scope reading.

(129) **Recall: frozen scope**

John is required to only learn Spanish.
How would the derivation proceed? First, Spanish would covertly move to the complement of only, as in (130), leaving a trace of type e. Assuming Spanish is interpreted referentially, [only\textsubscript{new}] applies to Spanish to yield a basic quantifier meaning for only Spanish (type <est, st>).

\begin{equation}
\text{(130) Step 1: covert focus movement}
\quad [\text{TP} \text{John}_2 \lambda 1 [\text{vp} \ t_1 \text{is required} [\text{PRO}_2 \lambda 1 [\text{vp} [\text{only Spanish}] \lambda 3 [\text{vp} \ t_1 \text{learned} \ t_3]]]]]
\end{equation}

Once this first step of focus movement has taken place, nothing should prevent the quantifier only Spanish from undergoing QR, which could transport only Spanish to a position above require, as in (131). Assuming the trace (\text{t}_4) is of type e, this results in the inverse scope order. As discussed in earlier, my own proposal correctly predicts frozen scope in this data point.\(^{15}\)

\begin{equation}
\text{(131) Step 2: QR of only Spanish}
\quad [\text{TP} \text{John}_2 \lambda 1 [\text{vp} [\text{only Spanish}] \lambda 4 [\text{vp} \ t_1 \text{is required} [\text{PRO}_2 \lambda 1 [\text{vp} \ t_4 [\text{vp} \ t_1 \text{learned} \ t_3]]]]]
\end{equation}

7.1.5 Local summary

Overall, we have seen that analyzing only uniformly as [only\textsubscript{new}] inherits the problems raised for pre-DP only, and generalizes them to vP only. In addition, certain scope facts with pre-vP only are not predicted. Wagner originally argued for his approach based on facts about NPI licensing: he argued for a generalization where NPIs are not licensed in the restrictor argument of [only\textsubscript{new}], but are licensed in its scope. I must postpone to a future occasion the question of how to capture Wagner’s NPI facts under my proposal.\(^{16}\)

7.2 Decomposing only (von Fintel & Iatridou 2007)

Another approach to only comes from von Fintel & Iatridou (2007, hence FI). Their proposal shares a crucial feature of my own: that only is masking two underlying morphemes. In their analysis, those morphemes are both interpreted — one is negation, the other an exceptive — and they attach at different positions in the structure. Examples such as (132-a) and (132-b) have the same underlying structure, and that structure informally corresponds to (132-c).

\begin{equation}
\text{(132) Negation + exceptive}
\quad \begin{align*}
\text{a. } \ & \text{John learned only Spanish.} \\
\text{b. } \ & \text{John only learned Spanish.} \\
\text{c. } \ & \text{‘John didn’t learn anything other than Spanish.’}
\end{align*}
\end{equation}

\(^{15}\)Based on data involving Antecedent Contained Ellipsis, Crnič ultimately concludes that this property of the focus movement approach is, in fact, a desirable feature of the model rather than a bug (see footnote 11). I, however, take it to be a challenge, and have set up my proposal to predict frozen scope in this case. Still, there is a tradeoff: complications arise for data with ACE (as explained in footnote 11).

\(^{16}\)For an additional argument for Wagner’s theory from island effects, see Erlewine & Kotek (to appear), though their arguments are countered in Longenbaugh & Bassi 2017.)
The syntax proposed for *only* is the same as the syntax languages such as French wear on their sleeve in a counterpart construction. A French datum with a meaning parallel to (132-a) and (132-b) is given in (133). Here the negation, \( n(e) \), and exceptive, *que*, are both overt.

(133) **French ne...que**

Jean n’a appris que l’espagnol.

J. NEG has learned but DEF Spanish

‘Jean only learned Spanish.’

Let us return to the English data for a concrete illustration of Fl’s composition. Adopting their proposal, the LF for (132-a) and (132-b) would be (134):

(134) **LF for (132-a) and (132-b)**

\[
[\text{John} \, \lambda 1 [\text{NEG} \, [[\text{EXC Spanish}] \, \lambda 2 [\text{t1 learn t2}}]]]]
\]

Fl state a meaning for the covert exceptive, by which *EXC Spanish* is interpreted as a quantifier, per (135). *EXC Spanish* composes with a predicate \( P \). The presupposition is that the extension of \( P \) is non-empty, and the assertion is that the extension of \( P \) contains some entity that is not Spanish. Applied to (134), the presupposition is that John learned something, which projects over negation. The assertion is that John didn’t learn anything that isn’t Spanish. With the presupposition and assertion combined together, the inference obtains that John learned Spanish and nothing else.

(135) **Interpreting EXC Spanish**

a. \((EXC \text{ Spanish}) P\)

b. \(A: \exists y [y \neq \text{Spanish} \land P(y) = 1]\)

c. \(P: \exists x [P(x) = 1]\)

It is immediately clear that Fl’s proposal has a certain commonality with my own: they propose that only constructions involve two heads, and that is a central feature of my proposal as well. The difference between our views lies in how compositional labor divides between heads. For me, the higher head is interpreted as a propositional operator, while the lower head is semantically inert. For them, the higher head is negation and the lower head is an exceptive.

Let us consider: how does Fl’s approach fare with the data discussed in this chapter? How might split scope derive? Suppose negation attaches in the matrix clause above *have*, while the exceptive phrase *EXC three papers* QRs within the embedded clause below *have*, as in (137).

(136) **Split scope prototype**

You have to publish only three papers.

(137) **NEG > have > EXC three papers**

\[
[\text{NEG} \, [\_P \, \text{you1 have} \, [\text{PRO1} \, \lambda 2 [\_P \, [\text{EXC 3 papers}] \, \lambda 3 [\_P \, \text{publish t3}}]]]]]
\]
The schema for interpreting EXC requires the complement of EXC to denote an entity, as Spanish does. The schema could be modified to accommodate a quantifier like three papers, but instead I show three papers QRred out of the exceptive phrase, as in (138). QR leaves a type e trace.

(138) **Final LF**

\[
\begin{array}{l}
\text{[you} \text{1} \text{NEG [vP t1 have [PRO1} \text{2 [vP [3 papers]} \lambda \text{4 [vP [EXC t4]} \lambda \text{3 [vP t2 pub t3]]]]]]}
\end{array}
\]

What meaning is predicted, given the schema in (135)? EXC in the embedded clause triggers the presupposition that you publish something. Assuming that presupposition projects universally from have, the predicted presupposition is (139): that at every require-world there is something which you publish. The assertion is (140): that not every world require-world contains both a plurality of three papers and something non-identical to it that you publish.

(139) **Presupposition of (138)**

\[\forall w' \in R(w) [\exists X [\text{you published } x \text{ in } w']]\]

(140) **Assertion of (138)**

\[\neg \forall w' \in R(w) [\exists X [\#(X)=3 \land \text{papers*(X, w')} \land \exists X' [X' \neq X \land \text{publish*(you, X', w')]}]\]

This meaning is not right. It predicts, for instance, that the perfectly felicitous sequence in (141) should come out as a contradiction. 17

(141) **Adding set-up sentence**

Chomsky has published six papers. You have to publish only three.

The set-up establishes that Chomsky published six papers. Granting that, the underlined portion of (140) comes out true at any world where you published anything, since then there will be a plurality of three papers (containing three of Chomsky’s papers) and something non-identical to it that you published. To satisfy the assertion, there must be a require-world at which you published nothing — but, an entailment that you be allowed to publish nothing is problematic in two ways. First, it is intuitively wrong: the sentence conveys that you must publish three papers. And, second, it contradicts the presupposition. It may be possible to re-formulate EXC to avoid the problem — but the ellipsis facts also pose a prima facie challenge. The data are repeated:

(142) **Ellipsis licensing**

a. To be considered for tenure, you have to publish only three papers.

b. And now you’re telling me that you can’t A.

The LF for (142-a) would be (138) above. The key observation is that there is no constituent which contains three papers, but not the exceptive head. The most likely candidate AC would be the underlined constituent, which contains both. Its interpretation is informally publish something other than some plurality of three papers. The detailed version is (143), with informal paraphrases of the

---

17Thanks to Roger Schwarzschild for pointing this out to me.
presupposition and assertion interleaved with the notation. Clearly, this is not semantically identical to $\Delta$, which is interpreted publish (at least) three papers.

(143) $[\text{AC}]$

\[ = \lambda x : \exists X [\text{publish}^*(x, X, w')] \]

\[ (x \text{ published something}) \]

\[ . X [\#(X) = 3 \land \text{papers}^*(X, w') \land \exists X' [X' \neq X \land \text{publish}^*(\text{you}, X', w')]] \]

\[ (x \text{ published something other than some plurality of three papers}) \]

Fl’s proposal differs from my own in that there is still an operator which composes with the DP, being the exceptive. This creates a problem for ellipsis licensing. The AC must contain the DP, but no component of meaning associated with only — not even the exceptive. I have suggested that the entire meaning is encoded in [only], which localizes at the vP level. Do the challenges raised for Fl genuinely rule out a $\text{NEG} \ldots \text{EXC}$ decomposition, or merely the particular formulation considered? To assess this definitively, cross-linguistic work is needed, testing whether the key data replicate in languages which overtly show a $\text{NEG} \ldots \text{EXC}$ decomposition. I leave this as a project for future work. For now, I note that English exceptive constructions argue that at least the ellipsis problem does independently require a solution. Consider ellipsis in (144), with an overt exceptive:

(144) **Ellipsis with an exceptive**

a. To get an $\Lambda$, you don’t have to read anything but this short paper.

b. And you’re telling me that you can’t $\Delta$.

The only VP apparent in (144-a) is *read anything but this short paper*, which contains exceptive *but*. Yet, $\Delta$ in (144-b) is interpreted as *read this short paper*, without any hint of the exceptive. The antecedent for $\Delta$ must come from (144-a) — but it is not clear how. Just as the first argument of EXC in (135) is of type e, von Fintel (1993) provides a semantics for *but* according to which its first argument is of type e (see also Moltmann 1995, Gajewski 2008, 2013, Hirsch 2015, Crnič 2016). For *but* to compose, (144-a) requires this structure:

(145) **Most obvious syntax for (144-a)**

\[ [\text{TP} \text{ you don’t have to read } [\text{DP} \text{ any } [\text{NP} \text{ thing } [\text{ExcP} \text{ but } [\text{TP} \text{ this short paper}]]]]]] \]

The ellipsis fact, however, suggests that there may be more syntax than meets the eye. Ellipsis would be predicted if the complement of *but* were a full clause. Following an earlier idea in Reinhart (1991), the structure just for the exceptive phrase would be (146). The elided clause contains the VP *read this short paper*, which is an appropriate antecedent for $\Delta$.

(146) **A hidden clause**

\[ [\text{ExcP} \text{ but } [\text{TP} \text{ you have to read this short paper}]] \]

Another familiar diagnostic for covert syntax converges. In (147), the complement of *but* contains a conjunction, and the second conjunct houses the adverbial *just yesterday*. Since *just yesterday* must
attach on the clausal spine (see Chapters 2, 3, and 4), its presence in the complement of but diagnoses a hidden clause. The structure in (147) is supported: *and* scopes at the vP within a hidden clause, and *just yesterday* adjoins to the second vP conjunct.

(147) **Corroborating with an adverb**  
John didn’t see anyone but [Bill and, just yesterday, Sue].

(148) **Again, a hidden clause**  
[Excp but [TP John₁ [&p [vP t₁ saw Bill] [and [yesterday [vP t₁ saw Sue]]]]]]

Modifying the semantics of *but* to compose with a full clause, while still replicating the results of von Fintel (1993), is a matter of importance, but I cannot do it justice in this thesis. For now, I want to return to *only* and note that if FI’s EXC takes a full clause as its complement, as it appears *but* does, then their decomposition of *only* into negation and EXC would be consistent with ellipsis in (149). Example (149-a) would include in its LF the exceptive phrase in (150), which contains the VP *publish (at least) three papers*. That VP could then serve as antecedent for Δ in (149-b).

(149) **Ellipsis licensing**  
  a. To be considered for tenure, you have to publish only three papers.  
  b. And now you’re telling me that you can’t Δ.

(150) **Exceptive phrase in (149-a)**  
[Excp EXC [TP you have to publish three papers]]

Overall, FI’s decomposition is not immediately satisfactory, but I believe their decomposition may still be promising. If EXC actually takes a clausal complement, ellipsis would follow. This syntactic change requires re-thinking the semantics of EXC, and that may be done in such a way as to resolve the scope problem. I leave this as a broad open avenue to pursue. If successful, their decomposition has the explanatory advantage of assigning meaning to both heads.

### 7.3 Section summary

This section has compared my proposal — where the entire semantic contribution associated with *only* comes from a vP-level operator interpreted as [only] — with two alternatives: a uniform two-place *only*, and decomposition of *only* into negation and an exceptive. I showed that two-place *only* faces the same problems as the original type-ambiguity theory, and generalizes them to further cases. Decomposition does not immediately capture the data, but a modification of the analysis seems to have promise, and is left to future work.

### 8 Conclusion

This chapter has addressed the focus operator *only*, with specific attention paid to cases where the focused constituent is contained within an object DP. In those data, overt *only* may appear before the
vP, or before the focused DP. Rooth (1985) proposed that this phonological variability has a semantic correlate: pre-vP only operates on the vP as [only], while pre-DP only operates on the DP as type-lifted [only,]. The chapter began with two problems for that analysis of pre-DP only:

1. **Split scope:** When only precedes an object DP, only may take scope above an operator at the same time that the DP takes scope below that same operator.

2. **Ellipsis licensing:** A VP of the form V only DP may serve as antecedent to license ellipsis of another VP which does not contain only, i.e. V-\(\phi\).

In Chapter 3, I noted similar scope and ellipsis puzzles with and in apparent object DP conjunction, and showed that they resolve if and actually makes its semantic contribution at the vP. This chapter came to the same conclusion for only. Despite appearances, pre-DP only does not compose with the DP as [only,]. Rather, the interpreted operator is [only], at the vP. Variability in where overt only appears in the linear string is orthogonal to the semantics.

On this basis, I proposed that 'only' constructions involve hidden syntax such that there are two underlying heads. One head, ONLY, attaches just above the vP, while the other head, F, attaches local to the object DP. Variability in the position of only reflects optionality in how the morphology maps the bipartite structure to an overt string: only optionally realizes either ONLY (placing only pre-vP) or F (placing only pre-DP). The semantics, conversely, is invariant, with the interpreted operator always ONLY. ONLY contributes the meaning of E-only, while the F head is semantically inert. I reviewed a range of cross-linguistic data supporting the proposal's main syntactic assumptions. With respect to the cross-linguistic picture, I pursued the idea that the bipartite structure is universal, with all variability in the realization of 'only' due to differences in morphology.

### 8.1 A program for further inquiry

In this chapter, I was once again judicious in circumscribing the empirical domain of inquiry. This chapter has focused almost exclusively on one focus operator (only) in one profile of data (cases where an object DP is focused, or properly contains the focus). I believe the theory I have proposed is, however, general enough to extend to other environments and operators. The theory provides a null hypothesis of broad applicability for further study of focus constructions. To close the chapter, I identify three specific avenues for further inquiry.

#### 8.1.1 Relating focus and questions

As noted briefly earlier, Cable (2007) proposed an analysis of questions with a bipartite structure similar to the one I have pursued for focus. The question in (151-a) has the underlying syntax in (151-b): there is an interrogative complementizer at the C head position, and another head Q more local to the wh-word. Those two heads Agree for Q features, and the QP is what gets attracted to spec-CP. Wh-movement is re-envisioned as QP movement.

280
(151) **Cable’s syntax for questions**

a. Which language did John learn?

b. \[CP C_{[Q(0)]} [TP John1 [vP t1 learned [DP Q_{Q(+)}} \text{ which language}]]\]

An analogy may be drawn between ONLY and the interrogative C, and between F and Q. Indeed, Cable extended his syntax for questions to Hungarian focus fronting, building on Horvath’s (2007) insights. In Section 5, I noted a distributional parallel between where F can attach and where Q can attach. I argued that F must attach “sufficiently local” to the focused constituent. When the object DP is focused, I operationalized locality such that F can attach to that DP, but cannot attach as high as the vP. The data in (152), repeated from earlier, revealed just this profile with Q:

(152) **Distribution of Q**

a. Which language did John learn.

b. *Learn which language did John.

The DP which language is analogous to the focus. That DP may move in (152-a), showing that Q may attach just above that DP. On the other hand, the vP cannot move in (152-b), showing that Q cannot attach that high. My claims about the distribution of F receive independent corroboration from the distribution of Q. The pair in (153) establishes a further parallel:

(153) **Another parallel of F~Q**

a. John talked only to FRED.

b. To whom did John talk?

If the focused constituent is a DP complement to a preposition, only may precede the preposition, indicating that F attaches just above the PP. Pied piping of the PP in (153) shows that Q can attach above the PP in a similar configuration, with a wh-DP in place of the focus. This being said, there are also certain instances where the parallel breaks down:

(154) **A non-parallel between F~Q**

a. (John met a bunch of diplomats, but a single author.)

He met only [the author of War and Peace].


The position of only in (154-a) signals that F can adjoin to the DP when the object DP properly contains the focus. The focused element is War and Peace, inside the DP. In (154-b), Q does not seem to be able to attach at that same site, when the focus is replaced with a wh. The pied piping in (154-b) does, however, improve in matrix clauses and non-restrictive relative clauses, as in (155). Cable further observes that data along the lines of (154-b) are acceptable in other languages, including Tlingit. There are, then, still parallels for the distribution of F from Q.
A parallel with other environments

a. The author of which book did John meet.
b. War and Peace, the author of which John met, is a good read.

It would take us too far afield to fully explore the distribution of F relative to Q, but I hope to have established that correlating these heads is a promising direction to pursue. The investigation here has already turned up independent support for my assumptions about where F can attach when the object DP is focused, and further inquiry may help shed light on the general locality constraints that govern the relationship between F and the focus. Cable proposes a number of parameters on which languages differ with respect to where Q can attach. I suspect that understanding the distribution of F in English and cross-linguistically is a matter of fitting F correctly into the typology of Q.

8.1.2 Extending to other environments

The proposal is sufficiently general to analyze only across environments, not just in cases where the object DP is focused. To illustrate, let us consider an example where ‘only’ associates with a subject DP. In this case, F attaches local to the DP John. ONLY, which must c-command the focused constituent, attaches just above the top clausal node. The structure is (157). Only may realize either F or ONLY, and both options result in the same surface string.

(156) Subject DP focused: initial only
     Only John\_F came.

(157) LF for pre-subject only
     [ONLY [TP [v\_P F John] λ.1 [v\_P t\_1 came]]]

Another derivation which would yield the right meaning with a focused subject would have ONLY attach just above the vP, while the FP reconstructs at LF from its overt position in spec-TP back to its base position within the vP, below ONLY. The surface structure, prior to reconstruction, would be (158-a). If only realized F, the surface string in (156) would derive again. However, if only realized ONLY, a new string would derive, being (158-b).

(158) “Backwards association”

a. [TP [v\_P F John] λ.1 [v\_P t\_1 came]]
b. *John\_F only came.

In fact, as Jackendoff (1972) observed, (158-b) cannot be interpreted as equivalent to (156): ONLY cannot “backwards associate” with the subject via reconstruction. Still, backwards association is possible with even in English, and with only in many other languages, so a derivation like (158) is possible in principle. Limits on backwards association constitute a general puzzle for any theory of only (see Erlewine 2014 for a recent analysis, and references therein).

Given the availability of the structure in (157), with ONLY above the TP, a question arises as to when ONLY can attach at the TP-level versus the vP-level. When the object DP contains the focus, ONLY cannot attach as high as the TP. This is shown by the contrast between (159) and (160).
(159-a) is a familiar example with pre-\(vP\) only, and it derives from the structure in (159-b) if ONLY is realized. (160-a) is an ungrammatical counterpart with only in initial position, which would result from realizing the TP-level ONLY in (159-b).

(159) **Focus in object DP: pre-\(vP\) only**
   a. John only learned one\(F\) language.
   b. \([TP \text{ John}_1 \ [\text{ONLY} [vP \text{ t}_1 \text{ learned} [FP \text{ F one language}]枉]]]\)

(160) **Focus in object DP: initial only**
   a. *Only John learned one\(F\) language.
   b. \([\text{ONLY} [TP \text{ John}_1 \ [[vP \text{ t}_1 \text{ learned} [FP \text{ F one language}]枉]]]\)

One possibility is that ONLY, like \(F\), is subject to a locality constraint. To capture the contrast between (156) and (160), we might posit a constraint that ONLY attach at the sentential node closest to the focus (following Jacobs 1983, 1986; Büring & Hartmann 2001). When the object DP contains the focus, the \(vP\) is the most local sentential node. When the subject DP contains the focus, given that it moves to spec-TP and cannot reconstruct below ONLY, the TP is the most local viable sentential node. The situation is more complicated, however, when embedding is considered. As we saw at various points in the chapter, when a sentence containing a focused DP is embedded, ONLY can attach to the matrix \(vP\) or the embedded \(vP\). Recall Taglicht’s ambiguous example:

(161) **Recall: Taglicht’s ambiguity**
You are required to learn only Spanish.
   a. \([TP \text{ you}_1 [vP \text{ t}_1 \text{ are req.} [TP \text{ PRO}_1 \lambda 2 \ [\text{ONLY} [vP \text{ t}_2 \text{ learn} [FP \text{ F Spanish}]]]]]]\)
   b. \([TP \text{ you}_1 [\text{ONLY} [vP \text{ t}_1 \text{ are req.} [TP \text{ PRO}_1 \lambda 2 [vP \text{ t}_2 \text{ learn} [FP \text{ F Spanish}]]]]]]\)

The embedded \(vP\) is more local to Spanish than the matrix \(vP\), but the locality constraint must not be so strict as to rule out attachment to the matrix \(vP\). Erlewine (2016) proposes a requirement for Vietnamese that the sentential focus particle attach as low as possible within a phase. This could perhaps account for the English data, as well. In the bi-clausal example in (161-a), the matrix and embedded \(vPs\) should be in separate phases, assuming that the clause boundary is phasal. If so, they might not compete with respect to locality. That said, I leave it as an open issue to properly study the distribution of ONLY once the theory is generalized across environments.

### 8.1.3 Extending to other operators

Although I built up the theory with only, it seems promising to extend the theory not only across environments in which only appears, but also across focus operators. Recall that even, like only, patterns as a Free Operator. Extending the proposal directly from ‘only’ to ‘even’, these strings would share the bipartite syntax in (162). The higher head is now EVEN, which attaches just above the \(vP\) and is interpreted as the sentential operator \([\text{even}]\). \(F\) is still just a dummy functional element, here attached to the DP a cooking contest. The two heads Agree, this time for \([\text{EVEN}]\) features, and overt even may realize either head, deriving the two word orders.
(162) **Distribution of even**

a. John won even a cooking contest.

b. John even won even a cooking contest.

(163) **Structure for (162)**

\[\text{[TP John] [EVEN[iEVEN()] [vP t1 won [FP uEVEN(+) [DP a cooking contest]]]]}\]

Some initial evidence for a bipartite syntax with 'even' comes from Colloquial Singapore English ('CSE'), as discussed in Quek & Hirsch (2016). CSE borrows its lexicon primarily from English, but has grammatical features in common with the southern Chinese dialects and Malay (Platt & Weber 1980). In CSE, certain focused constituents, including the associates of 'even', optionally front to a pre-verbal or pre-subject position. One example of fronting is (164):

(164) **Fronting with even**

[Even difficult integral] also my little sister got solve.

'My little sister solved even this difficult integral.'

In Hirsch & Quek (2016), we proposed the syntax in (165), just like the syntax for fronting seen earlier, but now adapted for 'even'. EVEN probes for an [EVEN] feature, and Agrees with the FP. EVEN further bears an EPP feature, triggering fronting of the FP to the specifier of EVEN.

(165) **Underlying syntax for (164)**

\[\text{[EVEN[iEVEN(), EPP] [TP my little sister got solve [FP F a cooking contest]]]}\]

In (165), two heads are overtly realized. One head, realized even, forms a constituent with difficult integral. The other, realized also, attaches on the clausal spine. We argued that even realizes F, while also is a special realization of EVEN when it bears an EPP-feature. By linking also to the EPP, we could explain why overt also does not co-occur with even in CSE without fronting:18

(166) **Without fronting**

*Also my little sister got solve [even difficult integral].

'My little sister even solved this difficult integral.'

Evidence for an operator-specific [EVEN] feature comes from restrictions on fronting. Whereas Hungarian allows associates of 'only', but not associates of 'even' to front, CSE exhibits the opposite pattern. Example (167) illustrates the impossibility of fronting with 'only'. Accordingly, we concluded that EVEN, when it triggers fronting, must probe for an [EVEN] feature.

---

18 Note that fronting also occurs with universal quantifiers and results in a distributive interpretation (reminiscent of Mandarin *dou*; see Liu 2016 and Xiang 2017 for recent treatments, and references therein for further discussion; on CSE, see Yeo & Tsoulas 2013). When a universal fronts, it is again followed by also. In Quek & Hirsch (2017), we assumed that fronting of distributive quantifiers was a separate process to focus fronting. The works just mentioned do, however, attempt unification, and their proposals involve encoding meaning in the head that triggers fronting distinct from the meaning of EVEN. If so, the EVEN head in CSE may in fact be separate from the head triggering fronting. We are not certain yet.
(167) **No fronting with only**

*Only difficult integral* also my little sister got solve.

‘My little sister only solved this difficult integral.’

Standard English ’even’ constructs might involve the same syntax as CSE fronting constructions, except with no EPP feature on EVEN. The two heads are present and Agree, but movement of the FP is not triggered. Comparing CSE and Hungarian, it is clear that which operator-specific features can co-occur with an EPP feature depends on the language: in Hungarian, [ONLY] can co-occur with an EPP feature, but [EVEN] cannot, and vice versa in CSE. I leave it as a future project to directly probe for a bipartite syntax in ’even’ constructions without fronting, and to test the idea that the meaning of ’even’ is localized at the higher EVEN head.
Chapter 7
Conclusion

1 Overview of the thesis

The starting point for the study was a division between sentential operators. Some sentential operators always occur syntactically with constituents associated with sentence meanings, while other sentential operators seem to show a broad distribution. These Free Operators (‘FOs’) were the main concern of the thesis, specifically the coordinators and or, and the focus operator only.

The traditional approach to FOs takes them to be systematically ambiguous in their semantic type. I showed in Chapter 2, however, that at least certain mechanisms that have been employed to derive systematic ambiguity over-generate, since they predict that all sentential operators should be FOs. The hypothesis I have pursued takes all sentential operators to have an inflexible semantics by which they compose with sentential meanings: the Semantic Inflexibility Hypothesis (‘SIH’). And is interpreted as [and], which operates on truth-values, and only is interpreted as [only], which operates on a proposition. The SIH makes a prediction that what is special about FOs is the syntax. In any case where a FO appears to operate on a constituent with something other than a sentential meaning, there must be covert syntax present, including a scope site for the FO.

In Chapter 3, I considered cases where and appears to operate on object DPs. In these data, I argued that and, when interpreted as logical conjunction, can and must be interpreted as [and], not the type-lifted variant [andas], which composes with quantifiers. I proposed that apparent DP conjunction is underlying vP conjunction, with vPs having type t meanings. Chapter 3 developed tests to diagnose hidden syntax which formed an important backbone of the thesis, as they were adapted for other constructions in subsequent chapters. The first test involved the distribution of adverbs: adverbs that can only operate on sentential meanings can adjoin to an apparent DP conjunct, diagnosing hidden syntax, including a sentential scope site. The second involved licensing of VP ellipsis: V DP1 and DP2 could license ellipsis of another VP having the form V DP2, indicating an underlying verbal projection in the second conjunct. The third involved split scope: and could scope higher than its apparent DP conjuncts, unexpected if and directly composed with those DPs. These tests showed that a parse with vP conjunction must be available, and further scope data were most readily understood if direct DP conjunction is unavailable.

Chapters 4 and 5 addressed two constructions which prima facie posed special problems for the SIH. These chapters had two parts. Part I reconciled the data with the SIH. Part II fit the results of Part I into a general analysis of the constructions, of interest independent of the SIH.

Chapter 4 considered pseudo-clefts. I started from the observation that DP and DP can occur as the pivot in a pseudo-cleft. If pseudo-clefts are a constituency diagnostic and pick out a DP and DP constituent, they are problematic for the SIH, assuming and can be interpreted as logical conjunction. I provided novel evidence for an idea previously developed to explain connectivity effects: that
pseudo-clefts contain covert syntax in the form of a full clause post-copular. That clause contains appropriate type t scope sites for [and], and its existence received support from the distribution of adverbs and licensing of VP-ellipsis, two diagnostics from Chapter 3.

In the remainder of Chapter 4, I probed more deeply into the semantics of pseudo-clefts, and developed a new perspective on their composition, according to which their compositional backbone is a covert Exh operator akin to only. The pre-copular constituent is interpreted as a question (i.e. a set of propositions), and provides an overt restrictor for Exh. The post-copular constituent is Exh’s prejacent. This LF captured the meaning of basic data, and novel examples with an adverb within the post-copular constituent. As part of this discussion, I focused special attention on cases where epistemic possibly occurs post-copular (What Obama approved was this bill and possibly that bill). This brought to the fore an “Over-negation Problem” which arises when Exh or only is combined with possibly. Exh and only are traditionally defined to exclude any alternative not entailed by the prejacent. As a result, Exh(ϕ ∧ ∨ψ) or only(ϕ ∧ ∨ψ) are wrongly predicted to convey (ϕ ∧ ∨ψ ∧ ¬ψ), a meaning similar to Moore’s paradox (I believe ψ, but ψ is false). In actual fact, Exh and only do not negate ψ in this case. I proposed a revision to the lexical semantics of only and only to account for that fact: ψ is negated only if [¬ψ] is consistent with the prejacent.

Chapter 5 turned to coordinating questions, both conjunction and disjunction. If and and or inflexibly operate on truth-values, question coordination is prima facie problematic, since questions, unlike declaratives, are not obviously associated with truth-values. I adopted an LF for a matrix question with a covert performative prefix. Performative sentences have been analyzed as being of type t, and the performative prefix creates an appropriate high scope site for and. I adopted a Hamblin semantics for questions and a compositional derivation for the Hamblin meaning with a type t node above the interrogative complementizer. I proposed that or scopes at that lower site.

In the remainder of the chapter, I focused closely on question disjunction. The existence of question disjunction has been a significant point of controversy. Adding to recent observations by other authors, I provided new evidence that question disjunction is attested. The intuition generally reported is that a disjunction Q1 or Q2 offers the responder a choice between two questions. I proposed that Q1 or Q2 is better analyzed as a single question which offers a choice between answers to that question. Specifically, I proposed that Q1 or Q2 is interpreted as a single mention-some question, which derived the choice intuition. This result followed from the LF in the first part of the chapter, combined with an independent theory of mention-some questions.

Chapter 6 shifted from coordination to focus operators. Much as Chapter 3 did with and, Chapter 6 focused on cases where only appears to operate on an object DP. The split scope and VP ellipsis tests from Chapter 3 were adapted to only and showed that pre-DP only does not compose with the DP, despite appearances. Having demonstrated the need for covert syntax, I proposed that pre-DP only is semantically inert Agreement morphology, while the interpreted operator is a covert ONLY, which scopes at the vP and is interpreted as [only]. The proposal made fine-grained predictions about scope, and received extensive cross-linguistic support.

The case studies converge on one and the same result: in cases where the SIH predicts covert syntax, that syntax is present. The tests from Chapter 3 — involving the distribution of sentential adverbs, licensing of VP-ellipsis, and split scope — are powerful tools to diagnose that covert syntax
in a variety of environments. It is important to re-iterate that all of the syntactic machinery I have drawn upon was needed independently. This is especially clear in Chapters 4 and 6. Chapter 3 analyzed apparent DP conjunction with the same mechanism as gapping, and Chapter 6 analyzed only with Agreement mechanisms widely deployed throughout the syntax and specifically deployed in a similar way with negation in negative indefinites. Pursuing the SIH has not taken burden off of the semantics only to place new burdens on the syntax. It has taken burden off of the semantics and made use of syntactic mechanisms that are required anyway.

As noted at the outset of the thesis, my aim has not been to provide a final argument that the SIH is correct. Rather, I have tested its predictions in a variety of specific cases and shown that they are borne out. To support the SIH in its full generality, covert syntax must be present in every single case where the SIH predicts it. At the end of Chapter 4, I laid out ideas on what covert syntax might look like in further environments, and left testing these ideas to future projects. At the end of Chapter 5, I noted that the Agree-based theory of focus operators is itself very general, though again, I left testing more predictions to the future. By showing that the SIH makes correct predictions in many environments, this thesis has, I hope, shown the promise of the research program.

2 Updating the typology of sentential operators

Before closing, let us take a step back and re-consider the two classes of sentential operators identified in Chapter 2. Class 1 operators were those which appear to be cross-categorial based on their surface distribution, and Class 2 operators were those that do not. Concretely, I categorized an operator as Class 1 if it appeared to operate on DPs, and an operator as Class 2 otherwise.

(1) Classification of sentential operators

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinators (and, or)</td>
<td>Modals (e.g. can)</td>
</tr>
<tr>
<td>Focus operators (only, even)</td>
<td>Modal adverbs (e.g. possibly)</td>
</tr>
<tr>
<td>not(?)</td>
<td>Evidential adverbs (e.g. allegedly)</td>
</tr>
<tr>
<td>Evaluative adverbs (e.g. luckily, sadly)</td>
<td>Aspectual adverbs (e.g. again)</td>
</tr>
<tr>
<td>Temporal adverbs (e.g. yesterday)</td>
<td></td>
</tr>
</tbody>
</table>

By diagnosing covert syntax in a range of cases where coordinators and focus operators seem to operate on DPs, Chapters 3-6 have taken steps towards a re-classification of those two core cases of Class 1 operators as Class 2. Once the underlying syntax is properly diagnosed, coordinators and focus operators actually operate on sentence meanings, like Class 2 operators.

Nonetheless, a question remains: if there are mechanisms which obscure the underlying syntax of coordinators and focus operators, why do they not also apply with other Class 2 operators to result in them having a deceptively broad surface distribution, as well? In fact, the mechanisms deployed in the central cases discussed in this thesis are not expected to generalize.

289
Conjunction Reduction in Chapter 3 involved ellipsis in a second conjunct, which would not be operative in the mono-clausal structures at play with other Class 2 operators. Moreover, in a CR structure, it is worth noting that the other Class 2 operators actually did precede a DP. One of the key arguments for hidden structure in apparent object DP conjunction came from contrasts such as:

(2) **Distribution of yesterday**
   a. ?*John saw yesterday you.
   b. John saw me and, yesterday, you.

The bipartite structure invoked with focus operators could in principle extend to other Class 2 operators, but I believe that there are principled reasons for why it does not. First, *only*:

(3) **Recall: only**
   a. John only read THREEm papers.
   b. John read only THREEm papers.

It is crucial to re-iterate that the interpretation of *only* depends on an element within the DP: *three* is focused and the make-up of the set of alternatives that *only* quantifies over is constrained by focus. In the bipartite structure that I proposed, the lower F head is a "focus marker", occurring in a local relationship with the focus. With other Class 2 operators, such as *yesterday* in (4), there would be nothing for a separate DP-level head to mark:

(4) **Compare with: yesterday**
   a. Yesterday, John read three papers.
   b. ??John read yesterday three papers.

We could envision the syntax in (5) for *yesterday*, where there is a head on the clausal spine and another local to the DP. If that syntax were right and *yesterday* could realize either head, *yesterday* should show a broad distribution, parallel to focus operators. The interpretation of *yesterday* in (4), however, does not depend on some specific property of *three papers* the way the interpretation of a focus operator depends on the focus. It would, then, make sense that this syntax is not correct.

(5) **Hypothetical bipartite syntax**
   \[ \{TP \text{John}_1 [\text{YESTERDAY} [\wp t_1 \text{read} [\wp F \text{three papers}]]]\}\]

One case to consider further in this connection is negation. As discussed in Chapter 5, analyses of negative indefinites and negative concord take DP-level negative morphology to be Agreement morphology with a higher interpreted negation. Does negation have any relation to a DP the way *only* does? Perhaps negation should be analyzed as a focus-sensitive operator, as suggested, for instance, in Jackendoff (1972) and Partee (1993). Partee provided illustrative contrasts such as (6), attributed to Hajíčková (1973, 1984):
Negation interacting with focus

a. This time John’s COUSINF didn’t cause our victory. (i.e. someone else did)
b. This time John’s cousin didn’t cause our VICTORYF. (i.e. she caused something else)

3 The big picture

As discussed in detail in Chapter 2, the overarching hypothesis that natural language cannot imbue an operator with flexible argument structure has far-reaching consequences for our understanding of the semantic module, and the mapping from syntax to semantics. In this final section, I want to re-iterate and expand on the general perspective that emerges from this work. In general, the aim has been to show that the semantics is, in one way or another, less powerful than commonly believed.

3.1 Blocking flexibility

In Chapter 2, we entertained a number of possibilities for how exactly flexible argument structure would come about. Perhaps the lexicon have sufficient expressive power to state a single meaning for FOs which allows them to compose with different arguments (as the \( \cap \) operator did for and). Alternatively, perhaps FOs have a basic meaning as sentential operators, while a family of other meanings are derived through type-shifting rules. If so, type-shifting rules could localize in different places in the grammar. They could be lexical rules mapping from a “core” lexicon containing the basic meaning to a “derived” lexicon containing the type-lifted meanings. Alternatively, just the sentential operator meaning might be stored in the lexicon, with type-shifting achieved through unary semantic rules which operate in the semantic module without syntactic or phonological correlate. Finally, type-shifters could be grammatical morphemes.

To facilitate exposition, I framed the thesis in terms of type-shifting, with a conception of type-shifting as unary semantic rules, following Jacobson. Although the SIH was stated for convenience in those terms, the point is more broad. Assuming the absence of flexible argument structure with FOs is not accidental, but principled, natural language must be unable to systematically imbue FOs with a flexible argument structure in any of the ways just enumerated. The lexicon must be built in a circumscribed way such that any single meaning has inflexible argument structure, as in the type system sketched at the very beginning of the thesis. Moreover, natural language must lack at least the Geach Rule as a lexical rule, a unary semantic rule, or encoded in a morpheme.

Pursuing the point with respect to type-shifting, calling into question the Geach Rule should cause us to ask more generally: does the grammar allow for any type-shifting rules and, if so, what kind? In assessing this question, the next natural step is to consider the nominal type-shifters of Partee (1987). The Montague Lift is perhaps most suspect, since it has a similar working to the Geach Rule: it shifts type \( \langle a, (s)t \rangle \rightarrow <a, (s)t> \), much as the Geach Rule does.\(^1\) If argument structure cannot be re-organized through type-shifting via the Geach Rule, perhaps it cannot via the Montague Lift either. While I leave proper consideration of these matters to the future, I would note that the results in this

\(^1\)The Ident operation \( [\text{Ident}] = \lambda x . \lambda y . y = x \) has a rather different character: while it does change type (type e to type \( <e, t> \)), it also introduces a new predicate — the identity predicate — which the Geach Rule and Montague Lift do not. It is possible that the grammar allows Ident, but not the Geach Rule or Montague Lift.
thesis have already served to undermine one of the classical arguments for the Montague Lift. That argument involves conjunction:

(7) Classical argument for Montague Lift
    I saw John and every student.

In (7), John, which has a basic meaning of type e, apparently conjoins with a quantifier. To parse the sentence, John is Montague Lifted from type e to type <et,t>. \([\text{and}_3]\) then operates on John and every student, just as it would operate in the quantifier conjunction every student and every professor. The Montague Lift is required to transform John into an appropriate argument for \([\text{and}_3]\). If the SIH is right, however, DPs are never logically conjoined, and the force of this argument evanesces. The example in (7) specifically would have a vP-level CR structure of the sort proposed in Chapter 3:

(8) CR-based LF for (7)
    \[
    [TP \lambda \lambda [vP t_1 \text{ saw John}] \ (\ [vP [every student] \lambda 2 [vP t_1 \text{ saw t_2}]]))]
    \]

John and every student are the objects of separate occurrences of saw within distinct vPs. John can be interpreted as type e within its conjunct, while every student is a quantifier in its conjunct.

### 3.2 Towards a general principle

Given that flexibility must be blocked in a number of different places, we are led to a deeper level of questioning: is there a very general principle of UG which renders flexibility impossible? In discussing this question, I will fall back on the habit of viewing flexibility as involving multiple related meanings. I said at the very beginning of the thesis that the SIH does not rule out that a meaning such as \([\text{and}_2]\) could be lexicalized in a language other than English, or even that \([\text{and}_3]\) might, in rare cases, be accidentally homophonous with \([\text{and}]\). I took this not to be so in English, but did not ban it. We might entertain, however, a stronger hypothesis that meanings such as \([\text{and}_2]\) are banned. If so, type-shifting principles to derive them should be banned, too.

I am aware of two principled ways to block higher-type meanings, which I outline below. By restricting the availability of higher-type meanings, the views developed below converge in important ways with a line of research being pursued by Pietroski (e.g. 2017) and Szabó (2016), as well as early work in Higginbotham (1985). As Pietroski notes, the recursive definition of semantic types allows for meanings of arbitrarily complex types to be stated. He and Szabo question whether those meanings are in fact possible in natural language. While I will not eliminate complex types altogether, the principles that I entertain do limit the extent to which they are attested.

#### 3.2.1 Option 1: an economy constraint on semantic type

The first possibility is based on Heim (2015). Heim's primary concern lies not with and or only, but with “raising” verbs such as seem. To build up to the idea, some background is in order on seem. Heim provides the example in (9), where seem occurs with a subject (John) and a complement clause apparently missing an argument (to be obnoxious).

292
(9) **Syntactic frame for seem**

John seems to be obnoxious.

*Seem* constructions have the property that although there may be a matrix subject, its interpretation is still associated with the embedded clause. (9) says that it seems (to the speaker, for instance) that John is obnoxious, not that it seems to John that someone is obnoxious. One way to make this clear is by noting that B does not contradict A in the following dialog:

(10) **“Faultless disagreement”**

A: John seems to be obnoxious.
B: No, he doesn’t.

As noted in the literature on predicates of personal taste (e.g. Lasersohn 2005, 2008, Stephenson 2007, Pearson 2012, 2013), disagreements like that in (10) are “faultless”. Speaker A conveys “John seems to be obnoxious to me” and Speaker B conveys “John doesn’t seem to be obnoxious to me”. These are mutually compatible, since the attitude holder is different in the two cases: A in the first versus B in the second. To set the attitude holder, an additional argument may be introduced into the syntax of A’s utterance, and then the two speakers contradict one another:

(11) **Setting the attitude holder**

A: John seems to me to be obnoxious.
B: #No, he doesn’t.

For our purposes, it will be sufficient to state a context-dependent denotation for *seem*, with the speaker the attitude holder. The denotation is stated in (12), where ‘S(SP(C))(w)’ picks out the set of worlds compatible with what seems to be the case to the speaker in the evaluation world, w.

(12) **Analysis 1: lexical entry**

\[ [\text{seem}_1]^C = \lambda p_x . \lambda w . \forall w’ \in S(\text{SP(C)}(w)) [p(w’)] \]

Given a meaning for *seem* as a sentential operator, the seminal raising analysis of Rosenbaum (1967) and Postal (1974) invokes the syntax in (13). *John* originates in the embedded clause, and moves from there to the matrix spec-TP. The trace of *John* saturates the argument of *obnoxious* to manufacture a proposition, and \[[\text{seem}_1]\] operates on that proposition, as sketched in (13). The trace is abstracted over and bound by *John* from the matrix spec-TP, above *seem*.

(13) **Analysis 1: syntax**

\[ [TP_2 \text{John } \lambda 1 [VP \text{seems } [TP_1 t_1 \text{ to be obnoxious}]]] \]
(14) **Sketch of composition**

a. $\llbracket TP_1 \rrbracket^{g,C} = \lambda w \cdot g(1)$ is obnoxious in $w$

b. $\llbracket VP \rrbracket^{g,C} = \llbracket [seem]^{g,C}(\llbracket TP_1 \rrbracket^{g,C})$

$b. = \lambda w \cdot \forall w' \in S(sp(C))(w) [g(1)$ is obnoxious in $w']$

c. $\llbracket TP_2 \rrbracket^{g,C} = \lambda w \cdot \forall w' \in S(sp(C))(w) [John is obnoxious in $w']$

The alternative is to adopt a higher-type meaning for *seem*, as in (15-a). $[seem_2]$ composes with a property $f$ (type $<e,st>$) and a quantifier $F$ (type $<est,st>$). It assembles those arguments into a proposition by computing $F(f)$, and asserts that the proposition is true in all *seem*-worlds. Like the different variants of *and* and *only*, the new meaning for *seem* is definable based on $[seem_1]$.²

(15) **Analysis 2: lexical entry**

a. $[seem_2]^C = \lambda f_{est} \cdot \lambda F_{<est,st>} \cdot \lambda w \cdot \forall w' \in S(sp(C))(w) [F(f(w'))]$

b. $= \lambda f_{est} \cdot \lambda F_{<est,st>} \cdot [seem_1]^C(F(f))$

Interpreting *seem* as $[seem_2]$ allows for the more transparent syntax in (16). Rather than raising from the embedded spec-TP, *John* originates in the matrix clause. The argument of *obnoxious* is left unsaturated in the embedded clause which therefore denotes a property, rather than a proposition. $[seem_2]$ composes with that property and a quantifier meaning for *John*. Internal to itself, $[seem_2]$ applies $[John]$ to the property, creating the proposition *that John is obnoxious*. I sketch the syntax and semantics in abridged form, omitting vP structure.

(16) **Analysis 2: syntax**

$[TP_2] \begin{array}{c} \text{John} \end{array} [VP \text{ seems } [TP_1 \text{ to be obnoxious}]]$

(17) **Sketch of composition**

a. $\llbracket TP_2 \rrbracket^{g,C} = \llbracket [seem_2]^{g,C}(\llbracket TP_1 \rrbracket^{g,C})(John)$

b. $= \llbracket [seem_1]^{g,C}(\lambda x \cdot \lambda w \cdot x \text{ is obnoxious in } w)(John))$

c. $= \lambda w \cdot \forall w' \in S(sp(C))(w) [John is obnoxious in $w']$

Heim provides arguments against Analysis 2. I will review just one of these arguments, which is based on Jacobson (1990). Jacobson observes an asymmetry between raising and control predicates in whether or not they allow null complement anaphora (‘NCA’, Grimshaw 1976). Many control predicates do, as illustrated in (18). Yet, no raising predicates do, as in (19).

(18) **Baseline: NCA with control predicates**

John tried/forgot/remembered/refused to take out the garbage, and I think that Bill tried/forgot/remembered/refused, too.

²Note that $[seem_2]$ is not directly derivable from the Geach Rule. The Geach Rule could derive a meaning of type $<est,<est,st>$. $[seem_2]$ is of type $<est,<est,st>,st>$ — which involves the same arguments, but in a different order. In other words, $[seem_2]$ is a different Currying of a meaning derivable from the Geach Rule.
Jacobson argues that the asymmetry in NCA between control and raising cannot be understood if *seem* is interpreted as *[seem]_2*. I present a modified version of her argument, based on Heim.

Heim assumes that NCA involves a null pronoun of property or proposition type. How, then, does NCA arise with control? Although control predicates occur in a superficially similar syntactic environment to raising predicates, the status of the subject is different. Whereas the matrix subject is semantically an argument of just the embedded predicate in raising, it is associated with both the matrix and embedded predicates in control. The first conjunct in (18) with *try*, for instance, would paraphrase: “John tried to make it such that he himself took out the garbage”. As such, the meaning of *try* cannot be that of a sentential operator. It must directly compose with *John*.

One option is to analyze *try* as (20). *Try* denotes a relation between an entity and a proposition, and asserts that the proposition is true in all worlds in which the entity’s efforts in the evaluation world are successful. The restrictor of the universal quantifier is picked out by ‘$T(x)(w)$’. The syntax for (18) may, then, be (21).

The embedded clause contains a covert PRO, and expresses the proposition that PRO took out the garbage. PRO is eventually bound by *John*, which also saturates the argument of [[try]], resulting in the overall proposition that *John took out the garbage in all worlds at which John’s efforts are successful*. In this case, the embedded clause (underlined) may be replaced with a proposition type pronoun to create the NCA observed in (18).

Having seen how NCA may come about with control, let us now proceed to the matter of main concern: how does the impossibility of NCA bear on Analysis 2 of “raising” constructions? To answer this, it is sufficient to examine its syntax:

Assuming that the pronoun in NCA can be either proposition type or property type, NCA could straightforwardly derive by the embedded TP in (22) (underlined) with a property type pronoun. To predict that raising predicates never allow NCA, Analysis 2 must be unavailable. The impossibility of NCA is an expected consequence of Analysis 1, repeated below. Because the matrix subject originates in the embedded clause and moves out of it, the embedded clause must contain sufficient
internal structure to include the base position of the subject. As such, it is not possible to replace the underlined material with a pronoun.\(^3\)

\[(23)\] **Recall: syntax for Analysis 1**

\[\text{TP} \quad \text{John } \lambda \text{t} [\text{VP seems } \text{TP} \quad \text{t} \text{ to be obnoxious}]\]

Taking \([\text{seem}_1]\) to be attested, and \([\text{seem}_2]\) not, Heim asks how would a child acquiring English converge on \([\text{seem}_1]\) over \([\text{seem}_2]\)? The impossibility of NCA would not help cue the learner, since it involves negative evidence. Heim's tactic, therefore, is to constrain the space of possible meanings through an economy constraint — one which sanctions \([\text{seem}_1]\), but rules out \([\text{seem}_2]\). To appreciate her account, it is crucial to re-iterate that \([\text{seem}_2]\) can be defined based on \([\text{seem}_1]\):

\[(24)\] **Recall: \([\text{seem}_2]\) from \([\text{seem}_1]\)**

\[\text{a. } [\text{seem}_2]^C = \lambda \text{f}_{\text{est}} \cdot \lambda \text{F}_{\text{est},\text{st}} \cdot \lambda \text{w} \cdot \forall \text{w}' \in \text{S(sp(C))(w)} [\text{F(f)(w')}]\]

\[\text{b. } [\text{seem}_2]^C = \lambda \text{f}_{\text{est}} \cdot \lambda \text{F}_{\text{est},\text{st}} \cdot [\text{seem}_1]^C(F(f))\]

Heim's idea is that a given meaning is disallowed when it can be defined based on a meaning of a simpler semantic type. She defines complexity in terms of the number of symbols in the type, as in (25). In the case at hand, \([\text{seem}_1]\) has a simpler type than \([\text{seem}_2]\), because \(<\text{st, st}>\) contains fewer symbols than \(<\text{est, est, st, st}>\). She then states the constraint as (26), and \([\text{seem}_2]\) is blocked because it is definable from \([\text{seem}_1]\).

\[(25)\] **Defining complexity**

\[\text{F has a simpler type than } \text{F}' \text{ iff the type-label of } \text{F} \text{ is shorter than the type-label of } \text{F}'.\]

\[(26)\] **Type Economy Principle**

\[\text{If } \text{F is definable in terms of an } \text{F'} \text{ with a simple type than } \text{F}, \text{ then } \text{F is not a possible denotation for a morpheme in natural language.}\]

As Heim notes, her constraint predicts that operators such as *and* and *only* must be lexicalized with their basic meaning, and not as one of the higher-type variants. As discussed at length in Chapter 2, all of the higher-type variants are definable based on the basic meanings \([\text{and}]\) and \([\text{only}]\). Among other effects, Type Economy requires any expression that *can* be interpreted as a sentential operator to only be listed as such in the lexicon.

Where does Heim's proposal situate with respect to the different conceptions of type-shifting? In addition to blocking *lexicalization* of higher-type meanings would it prevent them from being *derived*?

\[^3\text{The argument as presented differs in detail from Jacobson. I have assumed that NCA is licensed whenever the matrix subject does not originate in the embedded clause. Analysis 2 is thus incompatible with NCA, and Analysis 1 compatible, due to their syntax. Jacobson takes NCA to have a semantic licensing condition: the complement clause can elide if the embedding predicate takes it as an argument in Functional Application. She proposes that the complement clause never contains covert syntax, but rather always denotes property. Control predicates apply to that property, while *seem*, interpreted as \([\text{seem}_1]\), must compose through a different composition rule (Function Composition).}\]
through a type-shift? Because Type Economy references a *morpheme*, it is a principle about building a lexicon. As a result, type-shifting should be precluded as a lexical rule. Assuming that “morpheme” is understood to mean “an entry in either a core or derived lexicon”, the outputs of type-shifting would be ruled out. Moreover, Heim notes that her system prevents a type-shifter for FOs from being encoded in a morpheme. She observes that a shifter like (27) is definable based on the identity function of type <t,t>, as in (28). Because the identity function has a simpler type than the shifter, Type Economy blocks the shifter.

(27) **Shifter morpheme.**
\[
\begin{align*}
\text{[[SHIFT]]} & = \lambda F_{t,t} \cdot \lambda f_{et} \cdot \lambda g_{et} \cdot \lambda x \cdot F(f(x))(g(x))
\end{align*}
\]

(28) **Defining from identity function**
\[
\begin{align*}
a. \quad [[ID]] & = \lambda p_t \cdot p \\
b. \quad [[SHIFT]] & = \lambda F_{t,t} \cdot \lambda f_{et} \cdot \lambda g_{et} \cdot \lambda x \cdot [[ID]](F(f(x))(g(x)))
\end{align*}
\]

Although two conceptions of type-shifting are ruled out, it seems to me the third is still allowed. If “morpheme” is understood as “lexical entry”, Type Economy would not block higher-type meanings from being derived through a unary semantic rule, outside of the lexicon. It could be that unary rules must be blocked separate from Type Economy, in this conception. Alternatively, it might be possible to re-formulate Type Economy to apply not just to morphemes in the lexicon, but rather more liberally — perhaps to syntactically non-complex expressions, throughout the derivation.

To summarize, Heim’s system blocks through a general economy principle lexicalization of higher-type meanings for FOs, and blocks type-shifting in two places it is not attested. Blocking unary rules would require re-formulation of the constraint to apply outside just the lexicon.

### 3.2.2 Option 2: a strong theta criterion

A simplicity bias in semantic type is not, however, the only way to understand why higher-type meanings are not attested. In this subsection, I sketch an alternative perspective, which powerfully blocks type-lifted meanings, however they come about. As noted in Chapter 2, the effect of type-lifting *and* is to allow conjunction to take place before a sentence meaning is actually derived in the composition. At that point, I provided derivations comparing \[\text{[and]}\] with \[\text{[and}_2\]\]. I do so again, but with updated structures, annotated for semantic types:

---

4This perspective was suggested by Danny Fox (p.c.).
The derivation in (29) is a baseline involving the basic meaning $\text{[and]}$. Within each conjunct, the argument of the intransitive predicate is saturated by a trace bound by John to create a sentence meaning, and $\text{[and]}$ operates on those sentence meanings. In (30), on the other hand, $\text{[and}_2\text{]}$ operates directly on the intransitive predicates. As we have seen, it introduces a variable $x$ to saturate them, and applies $\text{[and]}$ to the resultant sentence meanings:

(31) **Interpreting the $\&P$ in (30)**

a. $\text{[&P]} = \text{[and}_2\text{]}(\text{[laughed]})(\text{[danced]})$

b. $= \lambda x. \text{[and]}(\text{[laughed]}(x))(\text{[danced]}(x))$

Since $x$ is abstracted over, its value is filled in by the trace of John and, finally, by John later in the composition. Looking at the tree bottom-up, $\text{[and}_2\text{]}$ enables conjunction before the arguments of the intransitive predicates have actually been saturated. In this subsection, I suggest that the grammar may include constraints which prevent an operator from intervening between a predicate and the argument saturating that predicate. Type-lifted variants of FOs would, then, be blocked, because they interrupt the required local syntactic relationship between predicate and argument.

The idea that the arguments of a predicate must occur in a particular syntactic configuration has its origin in work on linking theta roles to syntax. Predicates assign theta-roles (e.g. agent, theme, goal) to their arguments and linking principles require an argument bearing a specific theta role to occur in a specific syntactic position. Baker (1988) formulated in this connection the Uniformity of Theta Assignment Hypothesis (‘UTAH’): the principle that identical thematic relationships between items are represented by identical structural relationships between those items in the underlying syntax. In terms of the structural assumptions made in this thesis, the theme of a transitive verb is the complement within the VP, and the agent is the specifier of vP.

I suggest a different perspective from the UTAH, involving a strict linking between semantic and syntactic argument structure. As a starting idea, let us suppose that the semantic arguments of a predicate must be immediately saturated in the syntax. Any predicate requiring an argument of a particular type must be sister to an expression whose denotation is of that type:
Strict Linking Condition (‘SLC’; to be revised)

If an expression \( \alpha \) has a meaning of type \(<\sigma,\tau>\),
then it must be sister to an expression \( \beta \) with a meaning of type \( \sigma \).

To illustrate, consider the simple intransitive predication in (33). *Laughed* denotes a predicate of type \(<e,t>\) and, as such, the SLC requires that *laughed* be sister to an expression of type \( e \). Assuming that the linking condition is imposed at LF, laughed is sister to the trace of *John*, which is of type \( e \) and, thus, the condition is satisfied with respect to *laughed*. The SLC also requires that the \(<e,t>\) predicate derived with movement of *John* be sister to an expression with a type \( e \) meaning, and it is: its sister is *John*, which may be interpreted as type \( e \).

(33) A simple intransitive predication

\[
\text{TP} \\
\begin{array}{c}
\text{John} \\
\text{\( e \)} \\
\lambda 1 \\
\rightarrow \text{P} \\
\begin{array}{c}
t_1 \\
\text{laughed} \\
\text{\( e \)} \\
\end{array}
\end{array}
\]

Although the attested structure in (33) respects the SLC, other simple sentences show that the SLC is too restrictive in its current form, and requires immediate revision. Consider the structure in (34), which replaces *John* with a transparently quantificational DP, *every student*:

(34) A simple intransitive predication

\[
\text{TP} \\
\begin{array}{c}
\text{every student} \\
\text{\( \langle e,t \rangle \)} \\
\lambda 1 \\
\rightarrow \text{P} \\
\begin{array}{c}
t_1 \\
\text{laughed} \\
\text{\( e \)} \\
\end{array}
\end{array}
\]

The SLC is satisfied with respect to *laughed*, which is a sister to a type \( e \) trace, this time of the raised quantifier. Yet, a problem arises higher up. The SLC requires two conflicting things at the top node. First, it require *every student*, having a type \(<e_t,<e_t,t>>\) meaning, to be sister to an expression with a type \(<e,t>\) meaning. That is the case: its sister denotes the \(<e,t>\) predicate derived with Predicate Abstraction. In addition, however, it should require that derived \(<e,t>\) predicate to be sister to an
expression with a meaning of type e. That, however, is not the case: its sister is the quantifier. As a result, the SLC is incorrectly predicted to rule out this basic derivation.

Generalizing from the configuration in (34), the SLC will be violated whenever two expressions merge, each of which has a complex type. To prove this, suppose that an expression $\alpha$ with $[\alpha]$ of type $<a,b>$ is sister to an expression $\beta$ with $[\beta]$ of type $<c,d>$. The SLC requires that type a be type $<c,d>$, and it also requires that type c be type $<a,b>$. If $a$ is $<c,d>$, the type for $[\alpha]$ may be re-written $<c,d>,b>$. But, if $c$ is $<a,b>$, that type may be re-written $<<a,b>,d>,b>>$. Since a new occurrence of type a is introduced, this process of re-writing would continue to infinity. No actual complex types for $[\alpha]$ and $[\beta]$ can meet the specifications the SLC imposes.

If the SLC is on the right track, it must be constrained. One way to avoid a violation in (34) is to re-state the SLC to be only apply when a predicate externally merges with its sister, as in (35). In effect, the revised SLC says that every predicate introduced through external merge must be sister to an appropriate argument.

(35) **Strict Linking Condition (‘SLC’; updated)**
If an expression $\alpha$ has a meaning of type $<\sigma,\tau>$
and is externally merged with its sister $\beta$, then $\beta$ must have a meaning of type $\sigma$.

I continue to assume that the SLC applies at LF. Moreover, I will assume that two nodes count as externally merged with one another, even if one of the externally merged nodes is converted to a trace at LF. Because laughed and its sister externally merged in (34), the SLC requires that the sister of laughed have a meaning of type e, exactly as before. That requirement is met, since laughed is sister to the type e trace of John at LF. No SLC violation can, however, arise at the top node, since it is created through internal merger, i.e. movement, of the quantifier.

Of course, restricting the SLC to sites of external merger only somewhat weakens the prediction. The revised SLC still disallows any two expressions both with meanings of complex types from externally merging, unless one moves at LF to leave a lower type trace. Although this prediction is quite strong, I believe it is worth pursuing. In fact, it leads to some of the same ideas that I speculated on earlier, at the end of Chapter 5, with respect to the internal structure of a DP:

(36) **Most obvious syntax**
$[_{DP} \text{every } [_{NP} \text{linguist}]]$

The most obvious syntax for every linguist is the one just above, where every (as a D head) and linguist (as an NP) externally merge with one another. The SLC rules out this configuration. Since linguist externally merges with its sister and has a type $<e,t>$ meaning, its sister must be of type e, and every, being of type $<e,t>$ is not. The structure can, however, become viable if every originates within the NP and moves to the D-head position leaving a type e trace:

(37) **Derivation with D-movement**
$[_{DP} \text{every } \lambda I [_{NP} t_1 \text{linguist}]]$
In (37), *linguist* still counts as having externally merged with its sister, and its sister has a type \( e \) meaning, as the SLC demands. The quantifier has internally merged with its sister, so the SLC does not apply. Overall, then, the structure in (37) is SLC-respecting. In fact, I suggested a structure exactly like (37) at the end of Chapter 4 to explain how *and* can scope inside DPs. Saturating the \(<e,t>\) predicate with the type \( e \) trace creates a type \( t \) scope site for \([\text{and}]\) at the NP-level. I illustrated at that point with the DP *the linguist and philosopher*, which may pick out the unique person who is both a linguist and a philosopher. The suggested LF for that DP was convergent with (37):

(38) **Conjunction inside DP**

\[ [\text{DP every } \lambda x \ [\&P\ [NP \ t_1 \ \text{linguist}] \ [NP t_1 \ \text{philosopher}]]] \]

A related issue arises with Predicate Modification configurations, illustrated in (39). The two \(<e,t>\) predicates, *happy* and *dog* appear to be externally merged with one another, in violation of the SLC.

(39) **Predicate Modification configuration**

the happy dog

Building on the suggestions made about DP structure, one possibility is that the determiner originates within both the AP and the NP in (39), with the determined ATB-moved to the D head position, as in (40). Both \(<e,t>\) predicates are then sister to a type \( e \) trace. The externally merged AP and NP are both of type \( t \) and thus non-functional. In this way, the SLC is respected.

(40) **Possible structure for (39)**

\[ [\text{the } \lambda x \ [[\text{AP } t_1 \ \text{happy}] \ [NP t_1 \ \text{dog}]]] \]

How is (40) interpreted? Predicate Modification, as traditionally defined, elevates one of the type-lifted variants of \([\text{and}]\) — specifically, \([\text{and}_2]\) — into the grammar as a composition rule. Predicate Modification takes two \(<e,t>\) predicates and intersects them, just as \([\text{and}_2]\) does. If \([\text{and}_2]\) does not exist as a possible predicate, I am suspicious whether Predicate Modification should exist as a composition rule either. The structure in (40) leads to a re-formulation of Predicate Modification to parallel type \(<t,<t,t>>\ [\text{and}]\), as the AP and NP are both of type \( t \).

In general, the SLC leads to more movement than we may be accustomed to, but that movement, at least in some cases, receives support from the distribution of sentential operators, given the SIH. Let us assume the SLC and see how it rules out higher-type variants of sentential operators, returning to the conjunction derivations from the beginning of this subsection. First, the baseline with \([\text{and}]\):
Recall: derivation with [and]

Laughtered and danced are externally merged, and the SLC requires that each be sister to an expression of type e. This requirement is satisfied, since they are both sisters to a type e trace of John within their respective conjuncts. Moreover, since and is externally merged and interpreted as type <t,<t,t>>, it must be sister to an expression with a type t meaning — and it is: the right vP conjunct. In turn, the &’ has a type <t,t> meaning, and, again, it must be sister to an expression with a type t meaning — and it is: the left vP conjunct. The SLC is satisfied for every externally merged predicate, and the derivation is licit. The situation in (42) is, however, very different:

Recall: derivation with [and2]

Laughtered and danced are still externally merged, but they are not sister to an expression with a type e meaning. Danced is sister to and, and laughed is sister to the &’. And is interpreted as [and2] of type <et,<et,et>> and, in turn, the &’ is of type <et,et>. Accordingly, the SLC is not satisfied, and the derivation is ruled out. Given the SLC, the argument of a predicate must be immediately saturated and, in (42), the conjunction intervenes between the <e,t> predicates and their type e argument. All
of the higher-type variants of sentential operators would be ruled out in the same way. Because the SLC is a general constraint on the locality of arguments, not a constraint operative specifically in the lexicon, it would block higher-type meanings for sentential operators regardless of how they come about — whether they are lexicalized, or derived (anywhere in the grammar).

4 Final remarks

We now reach the end of the thesis. The goal has been to show that a key corollary of the SIH — that sentential operators are inflexibly interpreted as such — is supported in a range of cases. The thesis has undertaken detailed study of underlying syntax and syntax-phonology mapping to understand how modules interfacing with the semantics create the illusion of semantic flexibility. If the SIH is correct, it has far-reaching consequences for our understanding of how the semantics is constrained, and how the syntax-semantics mapping is constrained. In this chapter, I have sketched perspectives on some of these issues — issues which this thesis has imbued with new urgency.
References


Bayer, Josef. (1999). Bound focus in German or how can association with focus be achieved without going semantically astray? In G. Rebuschi & L. Tuller, (eds.), The Grammar of focus, pages 55-82, Amsterdam, John Benjamins.


Bjorkman, Bronwyn & Zeijlstra, Hedde. *Upward Agree is superior.* Ms. Queens University/Gottingen.


Cable, Seth. (2007). *The grammar of Q: Q-particles and the nature of wh-fronting, as revealed by wh-questions of Tlingit.* (Doctoral dissertation), MIT.


Heim, Irene. (2000). Notes on interrogative semantics. Lecture notes, MIT.

Heim, Irene. (2012). Advanced semantics. Lecture notes, MIT.

Heim, Irene. (2015). Constraints on meaning. Handout from the LF Reading Group. MIT.


Kuroda, S-Y. (1969). Remarks on the notion of subject with reference to words like also, even, or only. Reprinted in Papers in Japanese Linguistics 11, 98-120


Lin, Vivian. (2002). *Coordination and sharing at the interfaces.* (Doctoral dissertation). MIT.


van Rooij, Robert. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy* 26, 727-763


321


