

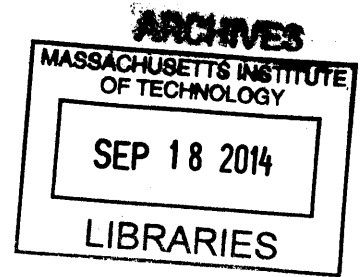
Movement Out of Focus

by

Michael Yoshitaka Erlewine

A.B., University of Chicago (2007)

A.M., University of Chicago (2007)



Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2014

© 2014 Michael Yoshitaka Erlewine. All rights reserved.

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

Signature redacted

Author

Department of Linguistics and Philosophy

August 29, 2014

Signature redacted

Certified by

Irene Heim

Professor of Linguistics

Thesis Supervisor

Signature redacted

Certified by

David Pesetsky

Ferrari P. Ward Professor of Linguistics

Thesis Supervisor

Signature redacted

Accepted by

David Pesetsky

Head, Department of Linguistics and Philosophy

Movement Out of Focus
by
Michael Yoshitaka Erlewine

Submitted to the Department of Linguistics and Philosophy
on August 29, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This dissertation investigates the consequences of overt and covert movement on association with focus. The interpretation of focus-sensitive operators such as *only* and *even* depends on the presence of a focused constituent in their scope. I document the complex conditions under which operators are able to associate with a focused constituent which has moved out of their scope. In particular, I concentrate on the ability of English *even* but not *only* to associate “backwards” in this configuration.

I propose a theory based on the Copy Theory of movement which predicts the attested patterns of backwards association. When an operator gives the appearance of associating backwards, it is in fact associating with focus in the lower copy of the movement chain, within its scope. This is possible with *even* but not *only* due to independent differences in their compositional semantics: *only* uses focus alternatives to compute new truth conditions, whereas *even* uses the alternatives to introduce a presupposition without modifying the truth conditions. I furthermore argue that neither syntactic reconstruction nor covert movement of *even* (the scope theory) are adequate as a general solution to the problem of backwards association. This analysis supports a view where focus is represented in the narrow syntax and then interpreted at the interfaces.

The analysis is built upon a general framework for focus interpretation based on Kratzer (1991) which I apply to structures involving copy chains, combined with new facts regarding the projection behavior of the scalar inference of *even*. After presenting my proposal, I discuss its implications for the internal structure of DPs and show that it offers a new structural diagnostic for the derivational path of movement. Moreover, the inability of scope reconstruction to feed focus association in English motivates a new approach to syntactic reconstruction. The proposal developed here explains a range constraints on patterns of focus association, and more generally contributes to our understanding of the interaction of syntactic operations such as movement with the semantic and information-structural notion of focus.

Thesis Supervisor: Irene Heim
Title: Professor of Linguistics

Thesis Supervisor: David Pesetsky
Title: Ferrari P. Ward Professor of Linguistics

Preface

There's a part of me that is amused by the fact that I work on this phenomenon called "focus," as focus may be the skill I have lacked the most as a grad student. I'm glad, though, that my dissertation concerns a problem which has bothered me throughout my entire time at MIT. The problem I investigate here originated with a terribly confusing squib that I wrote for Kai von Fintel in my first year, on the interpretation of Japanese sentences with two *only* particles stacked on the same argument. It's transformed quite a bit over the years, sometimes as my most active project and many times not. I suspect that this dissertation has a certain quality in common with my MA thesis: when Chris Kennedy read it, his first comment was "I can tell that you thought about this for a year, and wrote it in a month." In this case, just multiply that through by three or so.

This result is largely due to the helpful and generous contributions of others, as well as their encouragement, to stay on track and to push forward. First and foremost I thank my committee: Irene Heim, David Pesetsky, Martin Hackl, and Kai von Fintel. It frankly seems odd that they are not co-authors; whether they like it or not, this is as much theirs as it is mine. This is especially true of my co-chairs, Irene and David. Any and all good theoretical ideas here probably came out of a meeting with Irene, while any and all interesting empirical observations probably came out of a meeting with David. Martin and Kai's feedback added important, complementary perspectives: Martin constantly pushed me towards the trees while Kai made me look at the forest. They were an ideal committee to push me in every direction at once, towards the data and the theory and the big and the small. This wouldn't have happened without them. Thank you.

The work here has been greatly influenced by discussions with colleagues at MIT, especially Isaac Gould, Patrick Grosz, Aron Hirsch, Hadas Kotek, Miriam Nussbaum, Juliet Stanton, Yasu Sudo, and Coppe van Urk, as well as comments from Noam Chomsky, Danny Fox, Sabine Iatridou, Shigeru Miyagawa, Rick Nouwen, and Maziar Toosarvandani. From beyond the leaky walls of Stata, I received important comments on this work at various stages from Roni Katzir, Chris Kennedy, Jason Merchant, Chris Tancredi, Raj Singh, and Michael Wagner. We might have spoken just once or twice, but each of these conversations were incredibly important to me and to the resulting ideas presented here. I also thank audiences at a 2011 GLOW workshop on association with focus, Sinn und Bedeutung 2013, and the 2014 LSA.

And now here comes everything else: during the past five years, I did many other things besides, well, focus. I begin with the linguists. Much of my taste in linguistic theory developed by observing Norvin Richards' head-dances and Donca Steriade's vigorous nodding during other people's talks. I thank my co-conspirators, Hadas Kotek, Isaac Gould, Ted Levin, and Coppe van Urk—here's to many more! Thanks also to Sam Alxatib, Athulya Aravind, Bronwyn Bjorkman, Anthony Brohan, TC Chen, Hyesun Cho, Noah Constant, Jessica Coon, Youngah Do, Maria Giavazzi, Iain Giblin, Claire Halpert, Edwin Howard, Natasha Ivlieva, Gretchen Kern, Hrayr Khanjian, Junya

Nomura, Sasha Podobryaev, Omer Preminger, Daeyoung Sohn, Sam Steddy, Ayaka Sugawara, Wataru Uegaki, Michelle Yuan, Suyeon Yun, Sam Zukoff, and many others for good company, food, and judgments. I also learned so much from Gennaro Chierchia, Jim Huang, Masha Polinsky, Louis Liu, Edwin Tsai, and Yimei Xiang, down the street.

One regret I have about the work I present here is the utter lack of Mandarin, Kaqchikel, and Atayal, which have been important pillars of my work. I thank Ana López de Mateo and Pedro Mateo Pedro for our many discussions of Kaqchikel; the Atayal speakers who graciously worked with me in Taiwan, especially Uncle Taya; and Pamela Pan and Ning Tang for their impeccable Mandarin judgments. I also thank Edith Aldridge, Henry Chang, Daniel Hole, and Walli Paul for their continued encouragement.

My life in Cambridge was further enriched by other pursuits. I was glad to be a part of the Boston Awesome Foundation and the broader Mozilla and WordPress communities, and I hope I was able to give back in some way. Thanks also to the MIT square and contra dancers, as well as the friendly staff at the Biscuit, Darwin's, Dwelltime, and 1369. I also thank my non-linguist friends (they exist!), especially Evan, Stephanie, Noah, Bailey, Sarah, for being there.

Finally I thank my parents, Ron and Minako, for providing all the opportunities which brought me this far; and to Naomi, for her unbounded support. And to Hadas, too, for all of the above.

I dedicate this work to the memory of my grandparents, Yoshiro and Masako Oka.

Contents

Notation	8
1 Introduction	10
1.1 The question	10
1.2 The idea	12
1.3 Previous work on association with traces	14
1.4 The plan	15
2 The framework	17
2.1 The interpretation of focus	17
2.1.1 Focus basics	17
2.1.2 Interpreting <i>only</i>	19
2.1.3 Computing focus	22
2.2 The Copy Theory and the A/ \bar{A} distinction	28
2.3 Interpreting copies of focus	35
2.4 The requirements of the lower copy: evidence from ellipsis	41
2.5 Summary	49
3 Refining the semantics of <i>even</i>	50
3.1 The established semantics of <i>even</i>	51
3.2 Projection through quantification	55
3.3 The projection of <i>even</i> through quantification	59
3.4 Determining the domain of quantification	64
3.4.1 Domain restriction by the lower copy and wholesale late merger	65
3.4.2 Cross-dimensional domain anaphora	70
3.5 Proposal	72
3.6 Summary	74
Appendix 3: A similar scalar inference	74

4	Proposal	78
4.1	The importance of the lower copy	80
4.2	A short note on syntactic reconstruction	87
4.3	<i>Only</i>	91
4.4	<i>Even</i>	101
4.5	Associating with the higher copy of focus	105
4.6	Focus association and covert movement	108
4.7	Detecting lower copies	114
4.7.1	<i>It...that</i> clauses	116
4.7.2	Raising to object and object control	117
4.7.3	<i>Tough</i> constructions	118
4.8	Summary	120
5	Varying the associate	122
5.1	Association with quantifiers	122
5.2	Association with the moved DP	127
5.3	Backwards association and DP structure	134
5.4	Association with <i>wh</i> -phrases	137
5.5	Summary	142
	Appendix 5: Backwards association with <i>also</i>	142
6	<i>Even</i> in downward-entailing contexts	148
6.1	The lexical ambiguity theory	149
6.2	The scope theory of <i>even</i>	152
6.2.1	Covert movement of <i>even</i> alone	153
6.2.2	Covert movement of <i>even</i> and its associate	156
6.3	Summary	159
	Appendix 6: Nakanishi's (2012) argument for the scope theory	160
7	Reconstruction and association	165
7.1	Reconstruction and association with F-marking in the restrictor	166
7.2	Reconstruction feeds backwards association with determiners	171
7.3	Reconstruction via Inverse Trace Conversion	177
7.4	Reconstruction feeds association with <i>only</i> in German and Dutch	187
7.5	Summary	191
	References	193

Notation

Exposition

CAPITALS	pronounced with stress
<i>italics</i>	emphasis (metalinguistic)
bold	emphasis (metalinguistic)
○	yes, possible
×	no, impossible

[...] _F	F-marked
$\rightsquigarrow p$	introduces the non-assertive inference p
\equiv	defined as

see page 20

Semantic types

t	truth value
e	individual
s	world
a	assignment function
σ, τ	common metavariables over types
$\langle \sigma, \tau \rangle$	function with signature $D_\sigma \rightarrow D_\tau$

Variables

D_τ	the domain of values of type τ
w_0	the actual world
\mathbb{N}	{1, 2, 3, ...}, the domain of variable indices
i, j	arbitrary indices in \mathbb{N}
\mathbb{N}_F	{F1, F2, F3, ...}, the domain of focus indices, following Kratzer (1991) \mathbb{N} and \mathbb{N}_F are disjoint.
g	an ordinary assignment function

h	a focus assignment function, following Kratzer (1991)	see page 24
h_0	the preadjacent focus assignment function	see page 24
i	a variable with index i	

Functions

$[[\alpha]]$	returns the semantic denotation of the syntactic object α	see pages 23–
$\text{typeof}(\alpha)$	returns the semantic type of the object α	
$f g$	the combination of the functions f and g ; “ f or else g ”	see page 23
$[x \mapsto y]$	the function with domain $\{x\}$, mapping x to y	

Chapter 1

Introduction

1.1 The question

This dissertation investigates the interaction of *movement* and *focus association*, two operations that have been widely adopted in contemporary theories of syntax and semantics.

The interpretation of operators such as *only* and *even* depends on the placement of focus in their scope. The sentences in (1a–b) demonstrate this sensitivity. The choice of focus has been shown to affect the interpretation of different focus-sensitive operators in a systematic way by Jackendoff (1972); Rooth (1985, 1992b) and much subsequent work.

(1) Association with focus: (modified from Beaver and Clark, 2008)

- a. David will $\left\{ \begin{array}{l} \text{only} \\ \text{even} \end{array} \right\}$ wear a BOW TIE when teaching.
- b. David will $\left\{ \begin{array}{l} \text{only} \\ \text{even} \end{array} \right\}$ wear a bow tie when TEACHING.

The question that this dissertation addresses is: *can the focused constituent move out of the scope of its focus-sensitive operator?* This configuration is schematized below:

(2) Moving a focus operator's associate out of the operator's scope:

$\alpha_F \dots [Op [\dots \underline{\quad} \dots]]$ (with α interpreted as the associate of the operator)

I will show that association in this configuration—which I term “backwards association”—is possible under a specific and quite limited set of circumstances. An important observation is that focus-sensitive operators differ in whether or not they allow this pattern of focus association. In the simple case of English topicalization, below in (4), we see that backwards association is possible

with *even* but not *only*. That is, it seems that movement disrupts the ability of *only* to associate with the moved constituent, but association with *even* is unaffected.

(3) **Baselines:**

✓ John $\left\{ \begin{array}{l} \text{only} \\ \text{even} \end{array} \right\}$ met [Mary]_F at the party.

(4) **Backwards association with *even* but not *only*:**

- a. * [Mary]_F, John *only* met ___ at the party. (based on Tancredi, 1990, ex. 57b)
 b. ✓ [Mary]_F, John *even* met ___ at the party.¹

We observe a similar contrast with the ability of these focus adverbs to associate backwards with the subject. As first noted by Jackendoff (1972), *even* is able to associate with a leftward subject, while *only* cannot. This observation has been a long-standing puzzle in the focus literature.

(5) **Backwards association with the subject:**

- a. * [John]_F will *only* pass the exam.
 b. ✓ [John]_F will *even* pass the exam.

This contrast also extends to cases of covert movement, such as QR. If *only* associates with part of a quantificational DP, the DP cannot take scope outside of the scope of *only*. This restriction does not hold of *even*:²

(6) **Association with *only* restricts the scope of “every boy”:** (based on Aoun and Li, 1993)


- a. Someone wants to *only* meet [every [boy]_F in the room]. $\checkmark \exists > \forall, * \forall > \exists$
 b. Someone wants to *even* meet [every [boy]_F in the room]. $\checkmark \exists > \forall, \checkmark \forall > \exists$

A central question addressed in this dissertation concerns the variation in the behavior of different focus-sensitive operators with regard to backwards association. I will develop a view under which the grammar is in principle able to generate backwards association configurations with all focus-sensitive operators, but only certain operators and not others are able to operate on such constructions and yield a usable meaning. The dissertation will provide a detailed investigation of the configurations under which backwards association is possible, and present a principled explanation of the rather complex set of facts that we will uncover.


²Aoun and Li (1993), following Tancredi (1990), makes the claim that focus association with *only* restricts QR height, but their example does not conclusively show this. See discussion in section 4.6.

1.2 The idea

I propose that the “backwards” association of *even* with focused material that has moved out of their scope is illusory. In such cases, *even* is actually associating with a *lower copy of movement*, within the scope of *even*. I adopt the Copy Theory of movement and assume that the “trace” position includes an unpronounced copy of the moved material. There is therefore an instance of the focus-marked constituent “Mary” within the scope of *even*, even though it is part of the copy of the subject that is unpronounced. *Even* in (7) is associating with this material in the lower copy of movement.

- (7) ✓ [Mary]_F, John *even* met [Mary]_F at the party. (=4b)
- 

The same configuration with *only* (8) yields an uninterpretable structure, due to independent differences in the semantic contributions of *even* and *only*. The crucial difference will be that *even*'s semantic contribution is a projective inference that does not modify the assertion, whereas *only*'s semantic contribution modifies the assertion. The semantics of *even* will be a point of detailed discussion.

- (8) * [Mary]_F, John *only* met [Mary]_F at the party. (=4a)
- 

This proposal also offers a principled explanation for the ability of VP-*even* but not VP-*only* to associate leftwards with a subject. This contrast was first observed in Jackendoff (1972), and has persisted as a long-standing puzzle in the syntax/semantics of focus association.

- (9) VP-*even* can associate with leftward subject, but not VP-*only*:
- a. ✓ A [professor]_F will *even* come to the party.
 - b. * A [professor]_F will *only* come to the party.

While the ability of *even* to associate with moved material as in (4) has been noted in some corners of the literature, its relation to the phenomena of leftward subject association with VP-*even* (9) has not been made explicit. This apparently exceptional association of VP-*even* with subjects as in (9) is a special case of the general difference between *even* and *only* we will observe in configurations such as (2). Leftward subject association in (9a) is possible due to the VP-internal subject: *a professor* was generated within the scope of *even* and then moved out.

Under my proposal, the focus contained in the *lower copy* of movement is essential for backwards association constructions. Support for this view comes from the contrast between the raising and control structures in (10), where backwards association with the subject is possible across the raising embedding “seem” but not across the control embedding “want.”

(10) **Backwards association across raising vs control:**

- a. ✓ A [professor]_F seems to *even* be at the party. *raising*
- b. * A [professor]_F wants to *even* be at the party. *control*

This contrast is explained under the common view that raising involves a movement chain, but the control construction does not, as illustrated in (11). In the raising derivation in (11a), there is a copy of the focused constituent in the scope of *even*. *Even* can associate with this focus and introduces its scalar inference. In contrast, in the control derivation in (11b), there is no instance of the focused constituent within the scope of *even*, explaining the unavailability of backwards association.

- (11) a. ✓ A [professor]_F seems to *even* [[a [professor]_F be at the party]. *raising*
- b. * [A [professor]_F]_i wants to *even* [PRO_i be at the party]. *control*

More generally, my proposal predicts that backwards association with *even* requires the moved constituent to include the focused material in its lower copy. Evidence from Binding Condition C will show this to be the case.

The proposal I develop here provides a principled solution to the difference between *only* and *even*—as well as more fine-grained facts to be presented later—based on their compositional semantics, while preserving the hypothesis that focus affects the semantics of Association with Focus operators in a uniform manner (Rooth, 1985, 1992b). There is no special technology of backwards association; instead, instances of apparent backwards association are actually association with (unpronounced) focused material in the operator's scope.

Two types of alternative proposals will be considered. The first alternative we might imagine is that backwards association forces syntactic scope reconstruction, so that the focused constituent is in the scope of *even* at LF. This alternative is schematized below:

(12) **Backwards association by syntactic reconstruction:**

- a. PF: [DP Quantifier ... α_F ...] ... [Op [...] ...]
- b. LF: ... [Op [... [DP Quantifier ... α_F ...] ...] ⇒ *Op* associates with *α*

This alternative faces a number of problems. First, it's not clear how this approach to backwards association would distinguish between *even* and *only*. Second, it predicts the DP in question to have obligatory narrow scope in backwards association configurations, contrary to fact:

(13) **Backwards association is compatible with different scopes for the DP:**

[_{DP} Every [student]_F] didn't *even* come to the party.

- a. $\checkmark \forall > \text{Neg}$: \Rightarrow No student came.
- b. $\checkmark \text{Neg} > \forall$: \Rightarrow Not every student came, but some may have.

The second alternative I consider is to extend the scope of *even*, so that the focused constituent outside of *even*'s surface c-command domain is actually within its interpreted scope at LF. This possibility is schematized in (14) below. At first glance, this possibility seems promising, as it has been independently proposed that *even* can covertly move in the "scope theory" of *even* (Karttunen and Peters, 1979; Wilkinson, 1996; Guerzoni, 2004; Nakanishi, 2012, a.o.).

(14) **The scope theory as potential solution to backwards association:**

- a. PF: A [professor]_F will [*even* [come to the party]]
- b. LF: *even* [A [professor]_F will [[come to the party]]]

However, movement of *even* should not distinguish between the raising and control contrast (10) above or, more generally, contrasts where backwards association with *even* requires the focused constituent to have originated within its scope. I will ultimately conclude that the scope theory of *even* cannot account for the correct pattern of backwards association with *even* and defend the view that focus-sensitive adverbs are interpreted in their pronounced position (Rooth, 1985).

1.3 Previous work on association with traces

The inability of *only* to associate backwards with material which has moved out of its scope has motivated previous authors to propose a general ban on backwards association configurations. This literature begins with Tancredi (1990), which proposes the Principle of Lexical Association (PLA): "An operator like *only* must be associated with a lexical constituent in its c-command domain" (p. 30). From the text of Tancredi (1990), it is clear that this PLA was meant to apply to both *only* and *even*.³

The PLA has since been invoked to bear on questions of LF and covert movement, e.g., in Aoun and Li's (1993) discussion of *wh*-in-situ interpretation. Beaver and Clark (2008) follows Tancredi's (1990) PLA generalization, based on *only*. They claim that traces can't be F-marked because F has to be realized prosodically, predicting no backwards association to be possible.

³To wit, Tancredi (1990, fn 1) states "In the remainder of this paper, I will ignore this complication [= that VP-*even* but not VP-*only* can associate leftward with subjects], and deal exclusively with sentences in which *even* and *only* behave identically."

The empirical landscape is complicated in two ways. First, as we have observed above, backwards association is possible with *even* but not *only*, which is unpredicted by both Tancredi's PLA and by Beaver & Clark's explanation for the PLA. Second, there is cross-linguistic variation in the ability of *only* to associate backwards, with Barbiers (1995) presenting grammatical cases of such association in Dutch.

English *also* is also able to associate backwards, similar to what is presented for *even* above. This was first described by Krifka (1998). Rullmann (2003) then proposed that in such cases, *also* is associating with an *F-marked trace* of the moved constituent. Rullmann's proposal is similar in spirit to what I will propose here. However, I will show that the distribution of backwards association with *even* and *also* differ in ways that cannot be explained without the adoption of the Copy Theory.

1.4 The plan

The dissertation is structured as follows: In chapter 2 I present my basic syntactic and semantic assumptions which will form the basis for my theory. In particular, I adopt the Copy Theory of movement (Chomsky, 1993) and associated work on the interpretation of chains (Sauerland, 1998; Fox, 2002) and present an adaptation of the semantics of focus from Kratzer (1991). I will show that this framework is able to model basic cases of focus association with copies and also discuss the semantic contribution of lower copies of movement.

In chapter 3 I investigate the projection behavior of the scalar inference of *even*. The arguments in this chapter support the view that scalar inference of *even* projects generically through quantification. This will become important for explaining the ability of *even* to associate backwards.

In chapter 4 I present my proposal for the patterns of backwards association of *even* and *only* with material which has moved out of their scope. I argue that backwards association with *even* is illusory: when this is possible, *even* is in fact associating with focus in the restrictor of the lower copy of movement, within its scope. Evidence for this view comes from the distribution of backwards association with *even*. This same configuration with *only* leads to an uninterpretable structure, which introduces a unsatisfied presupposition, or else is a non-contingent assertion (always true or always false). Such structures cannot meaningfully contribute a conversation and are judged as ungrammatical (Gajewski, 2002). This difference between *even* and *only* is explained by the different semantic contributions of *only* and *even*: *even* uses focus-alternatives to introduce a non-assertive inference, without affecting the truth conditions of the utterance, while *only* uses focus alternatives to introduce new truth conditions.

I extend these findings to patterns of backwards association and quantifier scope which can be explained under the same proposal using covert movement as the vehicle for covert scope shift.

Finally, I present the use of backwards association with *even* based on the theory developed in this chapter as a structural diagnostic for lower copies of movement. This is able to detect whether a moved constituent originated within a particular domain.

In chapter 5 I investigate the patterns of backwards association with focused determiners and with entire focused DPs. I show that in these cases, both *even* and *only* are unable to associate backwards. I argue that this inability to associate backwards is due to different reasons: backwards association with a determiner is generally impossible, while backwards association with a DP may be possible, depending on the semantics of the operator. I discuss grammatical cases of association of the focus-sensitive operator *also* with the whole DP. I also discuss patterns of backwards association with *wh*-phrases.

In chapter 6 I look at the scale-reversal behavior of *even* in downward-entailing contexts. There are broadly two approaches to this behavior in the literature. I show that the lexical ambiguity view, that the meaning of *even* is polarity sensitive, is able to correctly model the distribution of backwards association with *even* in downward-entailing contexts. In contrast, the scope theory of *even*, which proposes that *even* covertly moves and is interpreted with wider scope, will either overgenerate or undergenerate patterns of backwards association.

Finally, in chapter 7 I discuss the relation between reconstruction and focus association. I show that syntactic reconstruction cannot be the general mechanism by which *even* is allowed to associate backwards, but nonetheless syntactic reconstruction feeds focus association in certain configurations. This complex set of facts motivates a new conception of syntactic reconstruction. I will also discuss facts from Dutch and German which contrast with the English pattern.

Chapter 2

The framework

In this chapter I will present the syntactic and semantic foundation that my dissertation will be built upon. Section 2.1 presents an introduction to the *Alternative Semantics* approach to the compositional semantics of focus (Rooth, 1985, et seq), together with the formal mechanisms of focus evaluation that I will use here. The formalism I use is based on Kratzer's (1991) "focus indices," but with certain differences in notation. Section 2.2 will present my primary syntactic assumption, namely the Copy Theory of movement and associated work on the interpretation of movement chains. Section 2.3 will present the basics of how focus interpretation can occur together with the Copy Theory of movement. Finally, I investigate the semantic contribution of the lower copy of movement.

2.1 The interpretation of focus

2.1.1 Focus basics

It has long been recognized that the semantic interpretation of a linguistic utterance is determined not just by the words we put together into a sentence, but by *how* those words are uttered. In particular, work in the past 50 years has led to the recognition of the variable placement of sentential stress—"focus"—as a core grammatical function which is involved in a variety of syntactic, semantic, and pragmatic phenomena. Krifka (2006) provides a succinct, unified definition for the use of focus in natural language:

- (15) "Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions." — Krifka (2006)

We can see the connection between *focus* and *alternatives* more concretely through the following example. Consider example (16) below, where the word "turtle" is pronounced with a heavy stress.

The focus on “turtle” conjures up other potential alternatives to “turtle,” based on the current discourse context. Each of these local alternatives then corresponds to alternative propositions, as shown in (16). The meaning of the proposition without the contribution of focus-sensitive operators is called the *prejacent*.

(16) Alex took the TURTLE to school.

Prejacent proposition: Alex took the turtle to school.

Focused constituent: turtle

Alternatives to “turtle”: frog, pig...

Alternative propositions: Alex took the *frog* to school, Alex took the *pig* to school...

These same alternatives can then be used for a variety of purposes (Rooth, 1992b, and much literature since). For example, focus is used to evaluate the congruence of questions and answers (Rooth, 1992b, and references therein):

(17) **Evaluating question-answer congruence with focus:**

a. question: What did Alex take to school?

i. ✓ Alex took the TURTLE to school.

ii. # ALEX took the turtle to school.

b. question: Who took the turtle to school?

i. # Alex took the TURTLE to school.

ii. ✓ ALEX took the turtle to school.

The effect of focus on the interpretation of a sentence is not merely pragmatic but belongs squarely in the domain of semantics. No clearer is this than in the use of focus alternatives by quantificational operators. These are called *focus-sensitive operators* and include words like *even*, *only*, and *also*. In English, focus-sensitive operators can come in both adverb and constituent-marking variants. In this dissertation I will be studying the adverbial variants.

(18) **Adverbial and constituent-marking variants of focus operators in English:**

a. Alex $\left\{ \begin{array}{l} \text{only} \\ \text{even} \\ \text{also} \end{array} \right\}$ [_{VP} took the TURTLE to school].

b. Alex took $\left\{ \begin{array}{l} \text{only} \\ \text{even} \\ \text{*also} \end{array} \right\}$ [_{DP} the TURTLE] to school.

Each of these operators relates the prejacent proposition to alternative propositions in a different way. Here I give informal paraphrases of the meanings introduced by the different focus-sensitive operators:

- (19) Alex $\left\{ \begin{array}{l} \text{only} \\ \text{even} \\ \text{also} \end{array} \right\}$ took the TURTLE to school.
- a. *only*: the prejacent “Alex took the turtle to school” is true, but all the alternative propositions (“Alex took the frog to school,” “Alex took the pig to school”...) are false.
 - b. *even*: the prejacent proposition is “Alex took the turtle to school” was less likely than the alternative propositions, e.g. “Alex took the frog to school,” “Alex took the pig to school”..., but the prejacent is nonetheless true.
 - c. *also*: at least one of the alternative propositions (“Alex took the frog to school,” “Alex took the pig to school”...) is true, and additionally the prejacent “Alex took the turtle to school” is true.

Note that the semantics of each of these operators makes reference to the same prejacent proposition and the same set of alternative propositions. The computation of the prejacent and the alternatives occurs in a uniform manner, independent of which focus-sensitive operator—if any—is then used.

As a corollary to this, we see that if a different part of the sentence is focused instead, for example as in (20) below, the meanings introduced by *only*, *even*, and *also* varies in a systematic manner.

- (20) Alex $\left\{ \begin{array}{l} \text{only} \\ \text{even} \\ \text{also} \end{array} \right\}$ took the turtle to SCHOOL.

Next I will present the technical mechanism by which the choice of focus placement is related to the generated set of alternative propositions.

2.1.2 Interpreting *only*

In this dissertation I will be adopting the highly influential *Alternative Semantics* approach to focus developed by Mats Rooth in Rooth (1985, *et seq.*). In this approach, focused constituents are interpreted in-situ at Logical Form (LF), without movement to the focus-sensitive operator.⁴ In *Alternative Semantics*, just as every syntactic node has an ordinary semantic value, we can similarly com-

⁴Here I will not consider approaches to Association with Focus which involve covert focus movement.

pute a set of alternative denotations for that node, which we derive by swapping out any focused constituents with their contextually-determined alternatives.

Here for concreteness I will be demonstrating the system using the *only* variant of example (19). The uppercase text indicates pitch accent on “turtle.” Text in italics (and elsewhere, bold) is not meant to indicate any linguistic feature.

(21) Alex *only* took the TURTLE to school.

I follow Jackendoff (1972) and much subsequent work in modeling the effects of focus by the addition of a formal “F” feature to focused constituents in the narrow syntax.⁵ This abstract F-marking can be thought of as a syntactic annotation which mediates between the observed prosodic realization and its semantic consequences.⁶ The LF for (21) which I consider here is in (22).

(22) LF: *only* [Alex took the [turtle]_F to school]

Note that in this section I take the entire proposition, including its proper name subject, to be within the scope of the adverb *only*. This is done to simplify the initial presentation; in the next section I will present a derivation with the subject interpreted outside of the scope of *only*.

Let’s begin by defining our semantics for *only*. *Only* is focus-sensitive, which in Alternative Semantics means that its interpretation will make reference to the semantic value of its complement (the prejacent proposition) as well as a set of alternative propositions. Here I will use *Prejacent* to represent the prejacent proposition (type $\langle s, t \rangle$) and *Alternatives* to stand in for the set of alternative propositions. (I will reformulate this definition in terms of a unified semantic denotation function $[-]$ later in this section.) Following Horn (1969), *only* presupposes the prejacent proposition and asserts that all other alternatives are false:⁷

(23) **Semantics for *only* (preliminary):**

$\widehat{\text{only } \alpha}$ in world w is defined if *Prejacent*(w) is true

$\widehat{\text{only } \alpha}$ in world w is true $\iff \forall \varphi \in \text{Alternatives} (\varphi \neq \text{Prejacent} \rightarrow \neg\varphi(w))$

⁵I use the term “narrow syntax” to refer to the portion of the derivation that feeds both Phonological Form (PF) and Logical Form (LF), in the T-model of grammar (Chomsky, 1995).

⁶This idea of an abstract feature underlying both the placement of pitch accent and association with focus-sensitive operators seems to go back to the “[+Prominent]” feature of Fischer (1968), discussed in Anderson (1972). The details of the phonological realization of F-marking will not be crucial in this dissertation. See Selkirk (1984) for discussion of this question of exactly where the pitch accent of F-marking is placed—the so-called “focus projection” problem.

Note that in the approach to focus I adopt here, based on Kratzer (1991) and presented in the next section, “F” will be thought of as an index on syntactic nodes rather than a feature.

⁷Here the presuppositional part is encoded as a definedness condition on the denotation (Heim and Kratzer, 1998).

Let's see how this results in the correct truth conditions for (21). Given the LF in (22), I assume that the complement of *only* has the prejacent and alternative propositions in (24). Note that the *Alternatives* include the *Prejacent* proposition. This will always be true of alternative sets.

- (24) LF: *only* [α Alex took the [turtle]_F to school]
- a. *Prejacent* = $\lambda w . \text{Alex took the turtle to school in } w$
- b. *Alternatives* = $\left\{ \begin{array}{l} \lambda w . \text{Alex took the turtle to school in } w, \\ \lambda w . \text{Alex took the frog to school in } w, \\ \lambda w . \text{Alex took the pig to school in } w \end{array} \right\}$

Let's assume in the context of evaluation that the prejacent is true—Alex did take the turtle to school. The definedness condition of [*only* α] is therefore met in the actual world w_0 . The semantic value of [*only* α] in world w_0 —and hence, the truth conditions of the entire sentence (21) in the actual world—will be the following:

- (25) (21) in world w_0 is true *if and only if* [*only* α] in world w_0 is true *if and only if*
- $$\forall \varphi \in \left\{ \begin{array}{l} \lambda w . \text{Alex took the turtle...in } w, \\ \lambda w . \text{Alex took the frog...in } w, \\ \lambda w . \text{Alex took the pig...in } w \end{array} \right\} (\varphi \neq (\lambda w . \text{Alex took the turtle...in } w) \rightarrow \neg\varphi(w_0))$$
- which, in turn, is true *if and only if*
- $$\neg(\text{Alex took the frog to school in } w_0) \wedge \neg(\text{Alex took the pig to school in } w_0)$$

This accords with our understanding of the sentence, which was paraphrased above in (19a): (21) is true *if and only if* Alex took the turtle to school, and Alex did not take any other animal to school. The semantics for *even* will be discussed in detail in chapter 3.

Before moving to the formal mechanism by which these alternatives are computed, I note an important characteristic of the Alternative Semantics approach to focus alternatives adopted here. That is, the alternatives that are considered by a focus-sensitive operator are those of the *complement* of the operator (Rooth, 1985).⁸ Therefore we predict that the semantics of operators such as *only* will be sensitive to the placement of F-marking *within* the operator's scope (complement), but not outside of it.

I assume that adverbial focus-sensitive operators take scope in their surface position, and therefore we predict them to be able to associate with material in the verb phrase, but not the subject which is outside of their surface scope. This prediction is supported by the possible patterns of association of *only*, first discussed by Jackendoff (1972).

⁸Note that, in the Rooth (1992b) variant of the theory, which I do not follow here, the relationship between the alternatives considered by a focus-sensitive operator and the denotation of its complement is more indirect.

- (26) *Only* must associate with a constituent in its scope: (Jackendoff, 1972, pp. 248–250)
- a. * [John]_F *only* gave his daughter a new bicycle.
 - b. John *only* [gave]_F his daughter a new bicycle.
 - c. John *only* gave [his]_F daughter a new bicycle.
 - d. John *only* gave his [daughter]_F a new bicycle.
 - e. John *only* gave his daughter a [new]_F bicycle.
 - f. John *only* gave his daughter a new [bicycle]_F.

However, as I have mentioned above in chapter 1, it is not the case that *all* focus-sensitive operators must associate with material in their surface scope. For example, as Jackendoff (1972) notes, *even* is able to associate leftward with the subject in the *even* variant of (26a):

- (27) ✓ [John]_F *even* gave his daughter a new bicycle. (Jackendoff, 1972, p. 248)

It is important to note at this point that *only* appears to behave as predicted by the semantics of focus, while *even* appears to be misbehaved. This is the main puzzle which I will address in this thesis.

2.1.3 Computing focus

Now it's time for the real work. I will present the definition of the semantic denotation function $[\cdot]$ which will be used here. The formulations I adopt are based on those presented in Kratzer (1991)—in turn based on a discussion in Rooth (1985, p. 12)—but with some notational differences. The use of Kratzer's (1991) formulation simplifies certain aspects of the analysis when we later combine it with the Copy Theory of movement. This payoff will come in section 2.3.

I begin by adopting a basic system of compositional semantics where the denotation of all nodes take an assignment function variable and a world variable, as in Montague's (1970) *Universal Grammar* system.⁹ The extensional type τ corresponds to a "superintension" of type $\langle a, \langle s, \tau \rangle \rangle$ in this system.¹⁰ Denotations specified in the lexicon will be of superintensional type. A couple examples are given here:

⁹This model for logical forms is also adopted with extensive discussion in the "Formalism" section of Rooth (1985, pp. 45–59). See also Novel and Romero (2009) for a more modern discussion advocating for the use of semantic denotations which are assignment-dependent in this way.

¹⁰The term "superintension" was suggested by Irene Heim (p.c.). Note that a superintension of type $\langle a, \langle s, t \rangle \rangle$, corresponding to extensional type t , can be thought of as a set of assignment-world pairs, which is the representation of a "context" as proposed in Heim (1983).

(28) $\llbracket \text{turtle} \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a turtle in } w$ type $\langle a, \langle s, \langle e, t \rangle \rangle \rangle$

(29) $\llbracket \text{take} \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e . \lambda y_e . y \text{ takes } x \text{ in } w$ type $\langle a, \langle s, \langle e, \langle e, t \rangle \rangle \rangle \rangle$

The following gives a recursive definition of the semantic value denotation function $\llbracket \cdot \rrbracket$ for constituents which do not include F-marking.

(30) **Definition: $\llbracket \cdot \rrbracket$ for non-F-marked constituents:**

a. **Terminal Nodes (TN):** (Heim and Kratzer, 1998, p. 43)

If α is a terminal node, $\llbracket \alpha \rrbracket$ is specified in the lexicon.

b. **Functional Application (FA):**

$$\left[\left[\begin{array}{c} \alpha_{\langle a, \langle s, \tau \rangle \rangle} \\ \beta_{\langle a, \langle s, \langle \sigma, \tau \rangle \rangle} \quad \gamma_{\langle a, \langle s, \sigma \rangle \rangle} \end{array} \right] \right] \equiv \lambda g_a . \lambda w_s . \underbrace{\llbracket \beta \rrbracket (g)(w)}_{\text{type } \langle \sigma, \tau \rangle} \underbrace{\llbracket \gamma \rrbracket (g)(w)}_{\text{type } \sigma}$$

c. **Predicate Modification (PM):**

$$\left[\left[\begin{array}{c} \alpha_{\langle a, \langle s, \langle e, t \rangle \rangle} \\ \beta_{\langle a, \langle s, \langle e, t \rangle \rangle} \quad \gamma_{\langle a, \langle s, \langle e, t \rangle \rangle} \end{array} \right] \right] \equiv \lambda g_a . \lambda w_s . \lambda x_e . \llbracket \beta \rrbracket (g)(w)(x) \text{ is true and } \llbracket \gamma \rrbracket (g)(w)(x) \text{ is true}$$

d. **Predicate Abstraction (PA):**

$$\left[\left[\begin{array}{c} \alpha_{\langle a, \langle s, \langle \sigma, \tau \rangle \rangle} \\ \lambda i_\sigma \quad \beta_{\langle a, \langle s, \tau \rangle \rangle} \end{array} \right] \right] \equiv \lambda g_a . \lambda w_s . \lambda x_\sigma . \llbracket \beta \rrbracket ([i \mapsto x] \parallel g)(w)$$

where $[i \mapsto x]$ is the function with domain $\{i\}$ which maps i to x .¹¹

(31) **Definition: function combination**

$$f \parallel g \equiv \lambda x . \begin{cases} f(x) & \text{if } x \in \text{domain}(f) \\ g(x) & \text{otherwise} \end{cases}$$

Read “ f or else g .”

Focus-sensitive operators quantify over a set of *alternatives*, which is intuitively the set of possible denotations if we allow F-marked constituents to be swapped out with contextually determined alternatives. We first compute a semantic value as if the choice of F-marked portions is underdetermined, occupied by a placeholder, and then these placeholders are filled in later. This is implemented using the general technology of *indices* and *assignment functions* necessary in compositional semantics for the interpretation of pronouns and bound variables—see Heim and Kratzer (1998) for a standard treatment.

¹¹“ $[i \mapsto x] \parallel g$ ” is my notation, using function combination defined in (31), for what Heim and Kratzer (1998) calls “ $g^{x/i}$.” I introduce function combination as a more general notation, as it will be used in cases besides the substitution of individual values.

F-marked constituents will bear a *focus-index* or *F-index*. If there are multiple F-marked constituents in a structure which are logically distinct, different F-indices will be used. This notion of “logical distinctness” will be discussed later in section 2.3. In our example, (21), there is only one constituent with F-marking. We’ll give it the index F1.

(32) LF: only [Alex took the [turtle]_{F1} to school]

A *focus assignment function* h is function which takes an index $F_i \in \mathbb{N}_F = \{F1, F2, F3, \dots\}$ and returns an alternative denotation for the constituent marked F_i , according to the context. In our example, there will be three distinct focus assignment functions, in (33) below. Both ordinary assignment functions and focus assignment functions can be *partial*, following Heim and Kratzer (1998). By convention, when defining contextually relevant assignment functions for a given example, I will assume that assignment functions are undefined for indices that do not have an explicitly defined value.

(33) **Different focus assignments map F1 to different local alternatives:**

- a. $h_0(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a turtle in w
- b. $h_1(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a frog in w
- c. $h_2(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a pig in w

The set H is the set of these relevant focus assignment functions. So in the example above, $H = \{h_0, h_1, h_2\}$. The focus assignment functions in H must always map each focus-index to an object with a particular semantic type. This is given as a principle in (34). This assumption that local alternatives will always be of the same semantic type is a standard one, going back to Rooth’s definition for the set of alternatives (*p-set* in his 1985 terms) in Rooth (1985, p. 14).

(34) **Focus Assignment Type Uniformity:**

For any F-index $F_i \in \mathbb{N}_F$ and any two $h', h'' \in H$, the set of relevant focus assignment functions, either $\text{typeof}(h'(F_i)) = \text{typeof}(h''(F_i))$ or both $h'(F_i)$ and $h''(F_i)$ are undefined.

I will always use h_0 to represent the focus assignment function that yields the prejacent (35). The value of $[[\alpha_{F_i}]]$ interpreted with the ordinary assignment function g and the focus assignment function h_0 always yields the same semantic value as interpreting node α with assignment g and no F-marking.

(35) **The prejacent focus assignment function:**

h_0 is the focus assignment function such that, for any F-index $Fi \in N_F$,

$$h_0(Fi) = \begin{cases} \text{if there is a node } \alpha \text{ in the workspace with index } Fi, \\ \llbracket \alpha \rrbracket, \text{ as computed by its constituent parts}^{12}; \\ \text{otherwise, undefined.} \end{cases}$$

The prejacent focus assignment function h_0 must be an element in H , the set of relevant focus assignment functions.

The denotation of a constituent with F-marking is dependent on the focus assignment function used (36). Note that the output of the function h —here, $h(Fi)$ —will be ordinary-assignment-dependent. Therefore $h(Fi)$ takes as an argument the ordinary assignment function g , the part of the argument $h \parallel g$ covering indices in N . Here I use the function combination notation (31) in “ $h \parallel g$ ” (“ h or else g ”) to indicate the combination of a focus assignment function h and an ordinary assignment function g .¹³

(36) **Definition: $\llbracket \cdot \rrbracket$ for an F-marked node, given g and h**

Given an ordinary assignment function g and focus assignment function h ,

$$\llbracket \alpha_{Fi} \rrbracket (h \parallel g) \equiv \underbrace{(h(Fi))}_{\text{type } (a, (s, \tau))} (g)$$

where τ is the extensional type of α .

Given the above definitions for $\llbracket \cdot \rrbracket$, we can compute a focus-assignment-dependent meaning for the entire complement of *only* in our example sentence. Each node in the tree below is illustrated with extensional types, evaluated using the focus assignment function h and the ordinary assignment function g (given as the combined argument $h \parallel g$) and the world w .

¹²To have $\llbracket \alpha \rrbracket$ “computed by its constituent parts,” means to compute the denotation of the node α without using the rule for the interpretation of $\llbracket \cdot \rrbracket$ for F-marked nodes, which will be presented in (36). In other words, the denotation we would derive if the node did not have a focus-index.

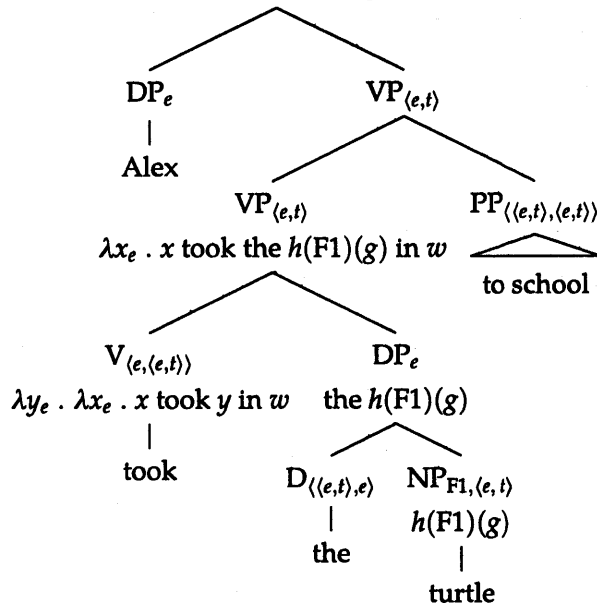
For example, if α is a lexical item, $\llbracket \alpha \rrbracket$ is the denotation of α as given in the lexical entry; if α has two sisters, a functor β and an argument γ of appropriate types, compute $\llbracket \alpha \rrbracket$ using the Functional Application rule; etc.

¹³Here for presentation’s sake, wherever possible, I will separate the ordinary assignment (with domain $\subseteq N$) and focus assignment (with domain $\subseteq N_F$) parts of an assignment function argument, as g and h , which are then passed in as a single function, $h \parallel g$. We could instead think of an assignment function argument as a single function, k , with domain $\subseteq N \cup N_F$. The definition for $\llbracket \cdot \rrbracket$ for F-marked nodes would then be:

$$(i) \llbracket \alpha_{Fi} \rrbracket (k) \equiv \underbrace{(k(Fi))}_{\text{type } (a, (s, \tau))} (k)$$

The definition in (i) is equivalent to that in (36) as long as the value of $k(Fi)$ is not itself focus-assignment-dependent; that is, as long as there are no F-marked nodes that dominate other F-marked nodes. Whether the system should allow for nested F-indices of this form is an empirical issue. Here I will not consider any configurations with this form.

- (37) $\llbracket \text{Alex took the [turtle]}_{F1} \text{ to school} \rrbracket (h \parallel g)(w) =$
 $\text{true} \iff \text{Alex took the } h(F1)(g) \text{ to school in } w$



By evaluating this fragment using different $h \in H$, we can compute each relevant alternative proposition, in (38) below. The resulting values here do not make reference to the ordinary assignment function g , reflecting the fact that the fragment computed does not include any ordinary indices.

- (38) Given an ordinary assignment function g ,
- $\llbracket \text{Alex took the [turtle]}_{F1} \text{ to school} \rrbracket (h_0 \parallel g) = \lambda w . \text{Alex took the turtle to school in } w$
 - $\llbracket \text{Alex took the [turtle]}_{F1} \text{ to school} \rrbracket (h_1 \parallel g) = \lambda w . \text{Alex took the frog to school in } w$
 - $\llbracket \text{Alex took the [turtle]}_{F1} \text{ to school} \rrbracket (h_2 \parallel g) = \lambda w . \text{Alex took the pig to school in } w$

For convenience, if the domain of the assignment function given to $\llbracket \alpha \rrbracket$ does not include the set of focus indices N_F , the prejacent focus assignment function h_0 will be used as a default:

- (39) **Notation:** $\llbracket \cdot \rrbracket$ given a partial assignment whose domain does not cover N_F
 Given an assignment function k with $N_F \not\subseteq \text{domain}(k)$,
 $\llbracket \alpha \rrbracket (k) \equiv \llbracket \alpha \rrbracket (k \parallel h_0)$

The set of alternatives that are used by a focus-sensitive operator is the set of denotations of the operator's scope, computed using different focus assignment functions $h \in H$, for a particular ordinary assignment function g . The set of alternatives for the complement of *only* in (22) above—"Alex took the $\llbracket \text{turtle} \rrbracket_{F1}$ to school"—is in (40):

(40) The alternative set for “Alex took the [turtle]_{F1} to school” given ordinary assignment g :

$$\{[\alpha](h \parallel g) \mid h \in H\} = \left\{ \begin{array}{l} \lambda w . \text{ Alex took the turtle to school in } w, \\ \lambda w . \text{ Alex took the frog to school in } w, \\ \lambda w . \text{ Alex took the pig to school in } w \end{array} \right\}$$

This procedure has yielded the desired set of propositional alternatives, presented above in (24).

This set of alternatives for a particular object α is often called the “focus semantic value” of α , following Rooth (1992b), in contrast to the “ordinary semantic value.” In Rooth’s system, these are two “dimensions” of meaning, with the focus semantic value computed using a different denotation function, $[\cdot]^f$. Definitions for the ordinary semantic value $[\alpha](g)$ and focus semantic value $[\alpha]^f(g)$ are provided in (39) and (41), respectively, for convenience, to relate the system presented here to more familiar systems (Rooth, 1985, 1992b, a.o.).

(41) **Definition:** $[\cdot]^f$

Given an ordinary assignment function g and the set of focus assignments H ,

$$[\alpha]^f(g) \equiv \{[\alpha](h \parallel g) \mid h \in H\}$$

However, in the formalism used here, both the ordinary semantic value and the focus semantic value of α are products of the superintension $[\alpha]$ and the focus assignment functions in H . This has the advantage of letting us restate the semantics of *only* from (23) purely in terms of this superintension, in (42) below.¹⁴

(42) **Semantics for *only* (final):**

$$[\text{only}]_{(a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle)} \equiv \lambda g_a . \lambda w_s . \lambda P_{(a, \langle s, t \rangle)} : \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\text{definedness condition}} \cdot \underbrace{\forall h \in H (h \neq h_0 \rightarrow P(h \parallel g)(w) \text{ is false})}_{\text{truth condition}}$$

In the notation presented here, what makes an operator “focus-sensitive” is that it quantifies over the contextually determined set of focus assignments, H . In order for the operator to meaningfully quantify over these focus assignments, the denotation of the focus-sensitive operator’s argument(s) must be focus-assignment-dependent.¹⁵ I state this requirement as follows:¹⁶

¹⁴The colon-dot notation for indicating definedness or domain conditions from Heim and Kratzer (1998, p. 34–34, 41) is used in (42).

¹⁵The requirement in (43) is stated here only for one-place focus-sensitive operators, such as focus-sensitive adverbs.

¹⁶In Rooth (1992b), these requirements are stated as a presupposition introduced by the squiggle operator \sim , requiring that the set of alternatives for ϕ , the complement of \sim , “contains both the ordinary semantic value [prejacent] of ϕ and an element distinct from the ordinary semantic value of ϕ .”

(43) **Focus Assignment Dependency Requirement:**

If Op is a focus-sensitive operator with complement α , the denotation of α must be focus-assignment-dependent; that is, there must be two focus assignment functions $h', h'' \in H$ such that, for all ordinary assignment functions g with $\text{domain}(g) \subseteq \mathbb{Z}$, $\llbracket \alpha \rrbracket (h' \parallel g) \neq \llbracket \alpha \rrbracket (h'' \parallel g)$.

Note that *only* in (42) will not compose using the general definition of Functional Application in (30) above, but instead must use the following variant of Functional Application which allows the operator to evaluate its complement using different assignments and evaluation worlds (44). Note that such a mode of semantic composition will generally be a necessary addition to (30) for the evaluation of intensional operators such as modals—see for example discussion of “Intensional Functional Application” in von Stechow and Heim (2005).

(44) **Superintensional Functional Application:**¹⁷

$$\left\| \begin{array}{c} \alpha_{\langle a, \langle s, \tau \rangle \rangle} \\ \swarrow \quad \searrow \\ \beta_{\langle a, \langle s, \langle \langle a, \langle s, \sigma \rangle \rangle, \tau \rangle \rangle} \quad \gamma_{\langle a, \langle s, \sigma \rangle \rangle} \end{array} \right\| \equiv \lambda g_a . \lambda w_s . \underbrace{\llbracket \beta \rrbracket (g)(w)}_{\text{type } \langle \langle a, \langle s, \sigma \rangle \rangle, \tau \rangle} \underbrace{\llbracket \gamma \rrbracket}_{\text{type } \langle a, \langle s, \sigma \rangle \rangle}$$

In the next section I present the syntactic theory of movement which I adopt: the Copy Theory of movement. Then in section 2.3 I will present what happens when we combine our system of focus interpretation with the Copy Theory.

2.2 The Copy Theory of movement and the A/\bar{A} distinction

The proposal that I develop here is based on the Copy Theory of movement. Under the Copy Theory of movement, syntactic movement does not replace its target with a new object, a “trace,” but instead merges another “copy” of the targeted object elsewhere in the structure (Chomsky, 1993). At Phonological Form (PF), one copy in each movement chain is chosen for pronunciation: in cases of overt movement, the head of the chain is pronounced while in cases of covert movement, a lower copy is chosen for pronunciation.

A number of syntactic arguments have been given for the Copy Theory, in work such as Chomsky (1993); Sauerland (1998); Fox (1999). The most common of these arguments comes from Binding Condition C, which requires that R-expressions not be bound. Condition C accounts straight-

¹⁷We could now write a lexical entry for the λ -abstractor which composes with Superintensional Functional Application, instead of using the syncategorematic Predicate Modification rule in (30c). The entry would be as follows:

(i) $\llbracket \lambda i_\sigma \rrbracket = \lambda g_a . \lambda w_s . \lambda P_{\langle a, \langle s, \tau \rangle \rangle} . \lambda x_\sigma . P([i \rightarrow x] \parallel g)(w)$

forwardly for the contrast in (45), as the R-expression "John" is c-commanded by the coindexed pronoun "he" in (45b) but not (45a).

(45) **Binding Condition C:**

- a. John_i believes [my argument that he_i is a genius].
- b. *He_i believes [my argument that John_i is a genius].

Now consider the *wh*-question in example (46). Here we have a configuration where "John" is not c-commanded by the coindexed pronoun "he" in the surface representation, but the question is degraded. This degradation is explained if a *copy* of the moved constituent is actually present at the movement gap. Under the Copy Theory of movement, the output of narrow syntax would be a structure as in (47), where "he" does c-command a coindexed "John," leading to a violation of Condition C.

(46) **Condition C and \bar{A} -movement:**

??/* [Which argument that John_i is a genius] does he_i believe ___? (Lebeaux, 1988)

(47) **Structure of (46) in the narrow syntax:**

[Which argument that John_i is a genius] does he_i believe
 [which argument that John_i is a genius]?

The movement involved in example (46) was *wh*-movement, an instance of \bar{A} -movement. On the other hand, A-movement is known to bleed Condition C: that is, the trace positions of A-movement chains are apparently never active for Condition C. This is illustrated through the grammaticality of the following examples:

(48) **Condition C and A-movement:**

- a. [The claim that John_i was asleep] seems to him_i to be correct. (Chomsky, 1993, p. 37)
- b. [John_i's mother] seems to him_i to be wonderful. (Lebeaux, 1998, p. 23)
- c. [Every argument that John_i is a genius] seems to him_i to be flawless. (Fox, 1999, p. 192)

I'll concentrate on example (48c) here. In the derivation of (48c), the matrix subject has A-moved from a base position as the subject of the nonfinite embedding, across the experiencer "him." When the DP is overtly in the lower subject position, as in the similar baseline sentence (49), coindexation of the experiencer "him" and "John" incurs a Condition C violation.¹⁸

¹⁸This is true despite the fact that the experiencer is apparently inside a prepositional phrase and therefore does not c-command "John" below. This is a more general issue regarding the structural configuration relevant for Condition

(49) Baseline: *It seems to him_i that [every argument that John_i is a genius] is flawless.

This behavior of A-movement and Condition C has been raised as a challenge for the Copy Theory of movement. As Takahashi and Hulseley (2009) discuss, if A-movement left a full copy as in (50a), we would yield a configuration parallel to (49) and therefore expect a Condition C violation. An alternative would be to use the Copy Theory for \bar{A} -movement but not for A-movement. We could then avoid a Condition C violation for the derivation in (50b) for (48c), but at the cost of complicating the basic operation of movement.

(50) **Two simple options for A-movement:** (based on Takahashi and Hulseley, 2009, p. 392)

a. With simple Copy Theory:

* [Every argument that John_i is a genius] seems to him_i to be
 [every argument that John_i is a genius] flawless]

b. Without the Copy Theory:

✓ [Every argument that John_i is a genius] seems to him_i to be [t flawless]

Takahashi and Hulseley (2009) propose that A-movement chains have the option of being base-generated as a determiner without a restrictor (here, “every”), with the restrictor *wholesale-late-merged* into the DP in its higher position (51). Because the restrictor does not exist in the lower position, we avoid the Condition C violation in this derivation. This approach preserves the central tenet of the Copy Theory, that all movement is copying (cf 50b), while allowing for A-movement to circumvent a Condition C violation.

(51) **Copy Theory with wholesale late merger:**¹⁹ (Takahashi and Hulseley, 2009, p. 400)

[Every] seems to him_i to be [[every] flawless]

↓ wholesale late merger

✓ [Every [argument that John_i is a genius]]₁ seems to him_i to be [[every]₁ flawless]

For Takahashi and Hulseley (2009), wholesale late merger of the restrictor must occur at least by the position where the DP receives Case, explaining the difference between the Condition C profiles of A- and \bar{A} -chains. In an \bar{A} -chain, the DP receives Case in its base position, therefore wholesale late C—recently taken up in Bruening (2014)—and is orthogonal to the discussion at hand. This same point is noted in Takahashi and Hulseley (2009, fn. 4).

¹⁹Here for simplicity I do not illustrate the intermediate movement step in the edge of the nonfinite clause, which Takahashi and Hulseley (2009) does illustrate.

merger cannot be used in \bar{A} -movement. The restrictor in an A-chain, on the other hand, could be generated in the base position of movement, but it need not be, assuming that DPs in A-movement chains receive Case at the head of their chain.²⁰

The Copy Theory of movement with wholesale late merger for A-movement will be adopted here. Together they are able to account for a range of data on the structure and interpretation of chains and differences between A- and \bar{A} -movements, all while unifying the core syntactic operation of movement as copying or remerge. Note that the facts surrounding binding and the interpretation of chains is more complicated than what is sketched here. I refer the reader to Fox (1999), Takahashi and Hulse (2009), and references therein for extensive discussion.

This Copy Theory approach to movement must be reconciled with our understanding of the semantic consequences of movement. Having multiple (coindexed) instances of an object at LF does not compositionally yield the expected truth conditions. Trace positions are crucial in the interpretation of movement, in particular as variables bound by the predicate abstraction step of movement (Heim and Kratzer, 1998). A solution that has been proposed is to tweak the lower copy at LF in order to interpret these copy-based movement chains. The lower copy is converted into a definite description with the restriction that it be equal to the variable in question through a process of *Trace Conversion* (Rullmann and Beck, 1998; Fox, 2002).

Consider example (52), which has a quantifier in object position. I assume that the quantificational DP undergoes Quantifier Raising (QR) through copying, resulting in a narrow syntax output with a chain of coindexed “every book” DPs.²¹ At LF the lower copy of “every book” will undergo Trace Conversion, resulting in the definite description “the book x ”—formally $\iota y . (y \text{ is a book and } y = x)$.

(52) **Interpreting copies in a movement chain:**

“John read every book.”

- a. Quantifier Raising as copying: $\underbrace{[\text{every book}] \text{ John read } [\text{every book}]}$
- b. LF after Trace Conversion: $[\text{every book}] \lambda x \text{ John read } [\text{the book } x]$

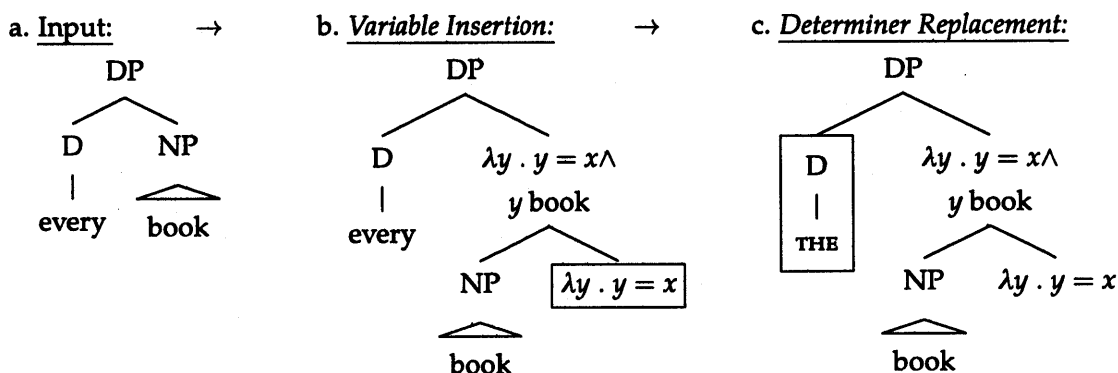
More technically, Trace Conversion can be decomposed into two independent manipulations of the lower copy. First is *Variable Insertion*, the adjunction of the predicate $\lambda y . y = x$ to the restriction of the quantifier, where x is the variable bound by the relevant predicate abstraction step.

²⁰One situation in which the restrictor of an A-chain must be generated lower is when the A-chain undergoes scope reconstruction (Takahashi and Hulse, 2009, based on Fox 1999), which will be discussed in chapter 7.

²¹I assume that quantifiers (extensional type $\langle\langle e, t \rangle, t\rangle$) in non-subject position must covertly move (Quantifier-Raise) to a position of (extensional) propositional type t , in order to resolve a type mismatch. Predicate Abstraction will turn the landing site into an expression of the necessary type (see Heim and Kratzer, 1998).

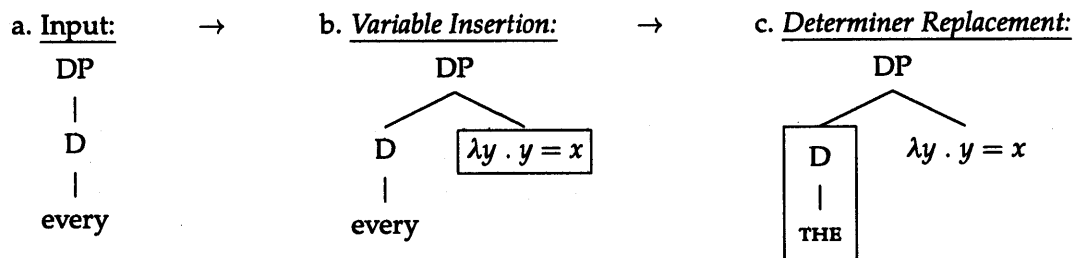
Second is the replacement of the quantificational determiner with a definite determiner **THE**, called *Determiner Replacement*.²² This Determiner Replacement step in particular will become important later in §5.1. These subparts of Trace Conversion are illustrated here:

(53) Trace Conversion for a DP with a restrictor (“every book”):



This Trace Conversion operation also applies to the lower copy of movement chains which do not include a restrictor, as happens in an A-movement chain where the restrictor is wholesale late-merged in a higher position. For example, in (51) above, the determiner “every” is base-generated in the nonfinite subject position, and then the restrictor “argument that...” is wholesale late-merged in the matrix subject position. The lower copy will then be the definite description “the x ,” where x is the variable abstracted over by the corresponding λ -binder above (Takahashi and Hulsey, 2009, fn. 8; see also Bhatt and Pancheva 2004, 2007).

(54) Trace Conversion for a DP without a restrictor (“every”):



Finally, before concluding this section, I will additionally review one striking argument for the Copy Theory of movement from Sauerland (1998, 2004), which comes from *antecedent-contained deletion* (ACD). ACD is a form of VP ellipsis, so-called because the VP ellipsis gap (represented by Δ) is, on the surface, properly contained by the constituent corresponding to the antecedent. Sauerland noted the contrast in (55) below, based on an observation by Kennedy (1994, 1997).

²²Variable Insertion and Determiner Replacement, to my knowledge, can be thought of as simultaneous operations or at least not dependent on a particular order of application. That is, the “first” and “second” in the text is not a claim of temporal or logical sequence.

(55) ACD requires matching head nouns:

- a. Polly visited every town that's near the town/one that Eric did Δ .
 $\Delta =$ "visit"
- b. *Polly visited every town that's near the lake that Eric did Δ .
Intended: $\Delta =$ "visit"

Sauerland argues that the contrast in (55) is explained by the general requirement that an elided VP and its antecedent must be semantically identical (Sag, 1976; Williams, 1977), together with the Copy Theory of movement and the interpretation of lower copies using Trace Conversion. We'll first see how this contrast is unexplained using simple variable traces, and then see how it is explained using the Copy Theory. Sauerland adopts a common approach to ACD, first proposed by Sag (1976) and May (1985) and supported more recently by Kennedy (1997) and Fox (2002), which posits that ACD always involves covert movement (QR) of a DP containing the ellipsis site to a position outside of the antecedent VP. (56) below illustrates such an LF configuration for (55a), using simple variables for traces:

(56) LF for example (55a) without the Copy Theory:

[every town [that's near [the town λx [that Eric did [ellipsis visit x]]]]]
 λy Polly PAST [antecedent visit y]

With the DP "every town..." QRed out of the antecedent VP, the elided VP is no longer contained in the antecedent VP. Ellipsis of the VP labeled "ellipsis" is licensed because that VP has the interpretation "visit x " which is equivalent to that of the antecedent VP, "visit y ."²³ Notice that changing the head noun of the relative clause "the town that..." in (56) to "lake" should also allow for the same ellipsis to be licensed. This LF is illustrated below as (57). This would predict example (55b) to be grammatical, contrary to fact.

(57) LF for example (55b) without the Copy Theory:

[every town [that's near [the lake λx [that Eric did [ellipsis visit x]]]]]
 λy Polly PAST [antecedent visit y]

\Rightarrow predicts ellipsis to be licensed and (55b) to be grammatical ☹

Sauerland argues that this contrast is explained by adopting the Copy Theory of movement and the Trace Conversion proposal for the interpretation of copy chains, reviewed above. Consider first

²³For current purposes, I follow the informal discussion section of Sauerland (1998) by assuming that the names of the variables in the VPs are not considered for evaluating the identity condition for ellipsis. See Takahashi and Fox (2005) for discussion of the notion of semantic identity required.

the Copy Theory-based LF for the grammatical example (55a), in (58) below. In (58), the elided VP includes the Trace Converted lower copy of the relative clause head, “the town x ,” and the antecedent VP includes the Trace Converted lower copy of the QRed DP, “the town y .” I assume here that the adjunct relative clause “that’s near...” is late-merged (Lebeaux, 1988) after QR of the host DP “every town.” (See discussion in Fox (2002).) Ellipsis is licensed in this example as the elided VP “visit the town x ” is semantically identical to the antecedent VP “visit the town y .”

(58) LF for example (55a) with the Copy Theory:

[every town [that’s near [the town λx [that Eric did [ellipsis visit the town x]]]]]
↑ λy Polly PAST [antecedent visit the town y]

In contrast, consider the Copy Theory LF for (55b), in (59) below. In this example, the head of the embedded relative clause is changed to “lake” and, under the Copy Theory, this has a corresponding change in the elided VP, which now refers to “the lake x .” In this example the ellipsis target “visit the lake x ” is *not* semantically identical to the antecedent VP “visit the town y ,” explaining the ungrammaticality of (55b).²⁴

(59) LF for example (55b) with the Copy Theory:

[every town [that’s near [the lake λx [that Eric did [ellipsis visit the lake x]]]]]
↑ λy Polly PAST [antecedent visit the town y]

⇒ ellipsis identity condition does not hold, predicting (55b) to be ungrammatical ©

By adopting the Copy Theory of movement, Sauerland is able to explain his contrast as an instance of the general requirement for semantic identity for ellipsis licensing. Note further that the determiners on the DPs moved in the grammatical ACD example (55a) are not the same: “every” on the QRed DP and “the” on the relative clause. This mismatch does not cause a problem for the licensing of ellipsis. This is explained by the process of Trace Conversion, which makes each lower copy be headed by the same determiner, “the,” but does not modify the restrictor of the lower copy. See Sauerland (1998, 2004) for further discussion of this and related ACD data.

With this background on the Copy Theory of movement, the A/ \bar{A} distinction, and the interpretation of chains, we now turn to the interpretation of focus using such a syntax.

²⁴Furthermore, because both QR and relative clause formation are \bar{A} -movements, wholesale late merger cannot be used to avoid this mismatch in the restrictors of the lower copies. When the movements are A-chains instead, a similar configuration with mismatched head nouns *does* allow for ellipsis, due to the option of not including the restrictors in the base positions of movement. See Takahashi and Hulsey (2009, §5) for presentation of such data and discussion.

2.3 Interpreting copies of focus

How does the Copy Theory of movement affect the interpretation of focus? The issue of focus interpretation within a Copy Theory of movement has not been addressed in the literature. In this section I will show how the interpretation of focus works out straightforwardly in examples involving copy-movement, using the framework for focus presented above, based on Kratzer (1991).

I begin this section by considering what happens if a DP containing F-marking in its restrictor moves within the scope of a focus-sensitive operator. We predict this configuration to be perfectly grammatical both for *only* and for *even*, as the F-marked constituent stays within the operator's scope. Under the Copy Theory, moving a constituent with an F-index will result in a structure with two such nodes with the same F-index. This configuration is schematized in (60):

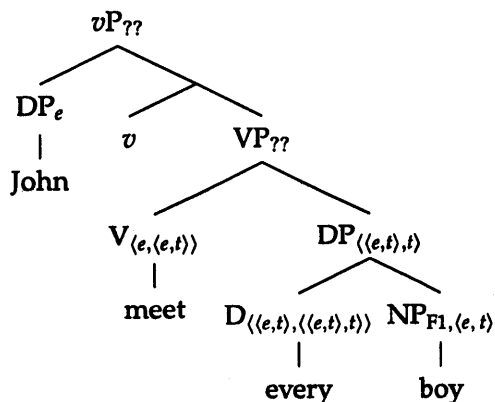
- (60) Moving a DP containing F-marking within the scope of an operator, with Copy Theory:
 $\checkmark \dots \text{only/even} [\dots \underbrace{[\text{DP} \dots \alpha_{F1} \dots]}_{\uparrow} \dots \underbrace{[\text{DP} \dots \alpha_{F1} \dots]}_{\uparrow} \dots]$

Here I will demonstrate how such a structure is interpreted using the simple example in (61):

- (61) John *only* met every [boy]_{F1}.

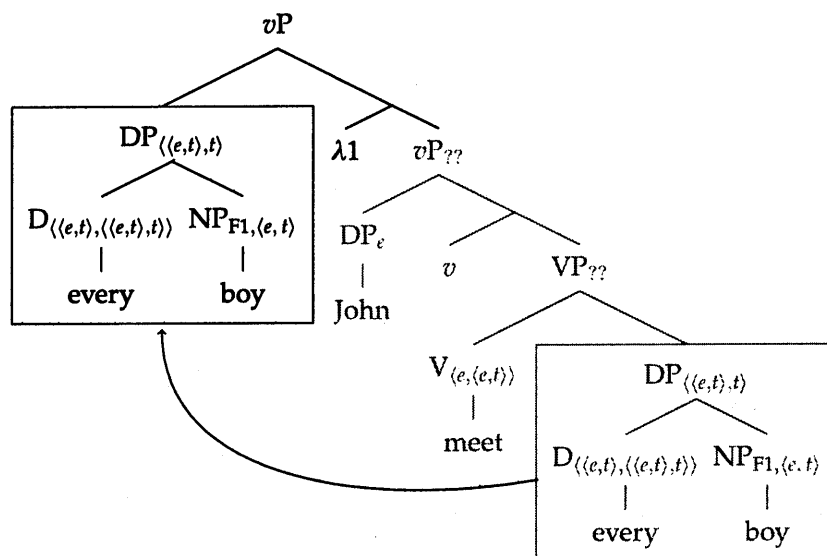
Although there is no overt movement in the derivation of (61), I assume that there is a covert movement step of "every boy" for interpretational reasons. Let's consider the derivation and interpretation of (61) step by step. We first build the *vP* up to the subject in Spec,*vP*. Here each node in the tree is annotated with its extensional type.

- (62) Building the *vP*:



Note that the verb of type $\langle e, \langle e, t \rangle \rangle$ does not compose with its quantificational object of type $\langle \langle e, t \rangle, t \rangle$. This motivates QR of the object to adjoin to vP .²⁵ Under the Copy Theory, this QR step will involve copying the quantifier to its adjoined position. This results in the structure in (63), with two instances of the predicate “boy,” each with the F-index F1. A lambda binder is introduced right below the landing site of movement (Heim and Kratzer, 1998, ch. 7).

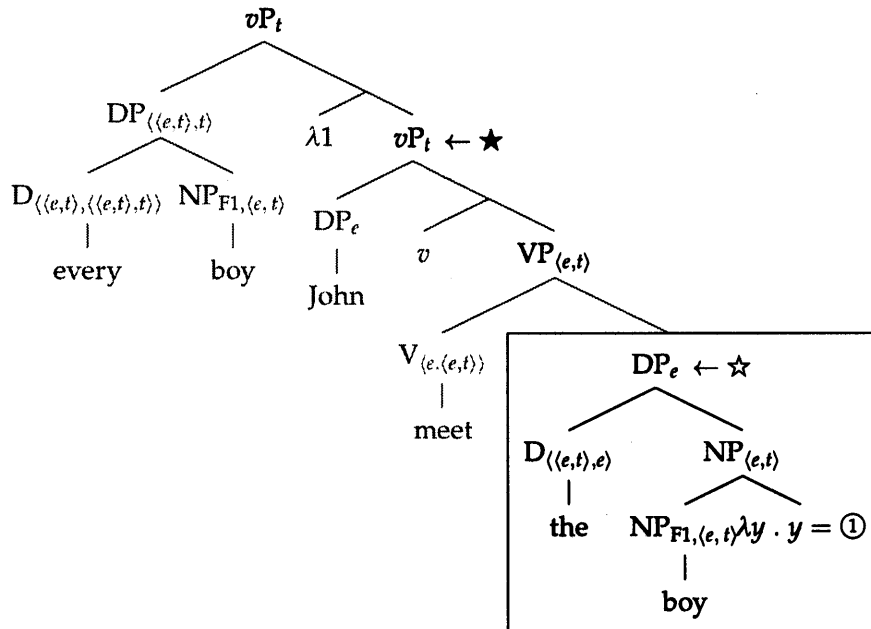
(63) QR as copying:



The lower copy will be transformed into a variable of extensional type e through Trace Conversion. (① is used to represent the variable with index 1.) As a result, the type mismatch will be resolved and the verb and the modified object will be able to compose.

²⁵In order to resolve this type mismatch, the object could conceivably move to a higher position of propositional type. It is critical here that QR target a position within the scope of *only*. Configurations where DPs QR to a position outside of a focus-sensitive operator will be discussed in section 4.6.

(64) Trace Conversion:



Using the definition of $\llbracket \cdot \rrbracket$ (30), we can now compute the supertension for the full vP . I begin with the denotation of the DP indicated by \star in (64).

$$(65) \llbracket \star \rrbracket (h \parallel g)(w) = \iota y . \underbrace{(h(F1))}_{\text{type } \langle a, \langle s, \langle e, t \rangle \rangle}} (g)(w)(y) \wedge y = g(1)$$

type $\langle e, t \rangle$

It is helpful at this point to understand what exactly this denotes, given a particular choice of focus assignment. Assume that the contextual alternative to the F-marked predicate “boy” is the predicate “girl.” These alternatives correspond to two different focus assignment functions, h_0 and h_1 , in H :

(66) Focus assignments for (61):

- a. $h_0(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a boy in w
- b. $h_1(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a girl in w

The DP \star in (65) then has the following denotations under these two focus assignment functions. Here I periphrastically use the metalanguage “the predicate individual” akin to English expressions of the form “the linguist Noam Chomsky.”

- (67) a. $\llbracket \star \rrbracket (h_0 \parallel g)(w) =$ the boy $g(1)$
- b. $\llbracket \star \rrbracket (h_1 \parallel g)(w) =$ the girl $g(1)$

Next I work up to the *vP* labeled \star in (64). I give the superintension in (68) as well as the two alternatives that they correspond to under each focus assignment function:

- (68) $\llbracket \star \rrbracket (h \parallel g)(w) = \text{true} \iff \text{John meets the } \underline{(h(\text{F1}))(g)(w) \ g(1)} \text{ in } w$
 a. $\llbracket \star \rrbracket (h_0 \parallel g)(w) = \text{true} \iff \text{John meets the boy } g(1) \text{ in } w$
 b. $\llbracket \star \rrbracket (h_1 \parallel g)(w) = \text{true} \iff \text{John meets the girl } g(1) \text{ in } w$

Next the λ binder abstracts over the index 1, using the Predicate Abstraction rule. Note that the denotation given by $h(\text{F1})$ is not dependent on the ordinary index 1, allowing for the simplification on the final line:²⁶

- (69) $\llbracket \lambda 1 . \star \rrbracket (h \parallel g)(w) = \lambda x_e . \text{John meets the } \underline{(h(\text{F1}))([1 \mapsto x] \parallel g)(w) \ ([1 \mapsto x] \parallel g)(1)} \text{ in } w$
 $= \lambda x_e . \text{John meets the } \underline{(h(\text{F1}))([1 \mapsto x] \parallel g)(w) \ x} \text{ in } w$
 $= \lambda x_e . \text{John meets the } \underline{(h(\text{F1}))(g)(w) \ x} \text{ in } w$

Finally we compose this with the higher, unmodified copy of the quantifier “every boy.” Recall that this copy of “every boy” also includes the predicate “boy” with focus index F1, and therefore will be focus-assignment-dependent.

- (70) $\llbracket \llbracket \text{DP every } [\text{boy}]_{\text{F1}} \rrbracket \rrbracket (h \parallel g)(w) = \lambda Q_{(e,t)} . \forall x(((h(\text{F1}))(g)(w))(x) \rightarrow Q(x))$
 (71) $\llbracket vP \rrbracket (h \parallel g)(w) = \text{true} \iff \forall x(((h(\text{F1}))(g)(w))(x) \rightarrow \text{J meets the } \underline{(h(\text{F1}))(g)(w) \ x} \text{ in } w)$

The denotation for $\llbracket vP \rrbracket$ in (71) has two portions that are focus-assignment-dependent, corresponding to the two copies of the F-marked predicate “boy” that were made through QR. When we evaluate the superintension of *vP* using any particular focus assignment function $h \in H$, we notice that the two instances of this predicate *covary*:

- (72) a. $\llbracket vP \rrbracket (h_0 \parallel g)(w) = \text{true} \iff \forall x (x \text{ is a boy} \rightarrow \text{John meets the boy } x \text{ in } w)$
 b. $\llbracket vP \rrbracket (h_1 \parallel g)(w) = \text{true} \iff \forall x (x \text{ is a girl} \rightarrow \text{John meets the girl } x \text{ in } w)$

Now we can evaluate *only*. For illustration purposes, I will interpret the non-quantificational subject “John” in *Spec,vP*, rather than outside of the scope of *only*. Assuming that the preajacent proposition holds, we yield the following truth conditions for the sentence (61):

²⁶This simplification will generally hold for cases where the F-marking is a proper subpart of the moved constituent. Different configurations will be discussed much later, in chapter 5.

- (73) $\llbracket (61) \rrbracket (g)(w_0) = \llbracket [only\ vP] \rrbracket (g)(w_0)$ is true if and only if
 $\forall h \in H (h \neq h_0 \rightarrow \llbracket vP \rrbracket (h \parallel g)(w_0)$ is false)
 $\iff \llbracket vP \rrbracket (h_1 \parallel g)(w_0)$ is false
 $\iff (\forall x (x \text{ is a girl} \rightarrow \text{John meets the girl } x \text{ in } w_0))$ is false
 \iff John did not meet every girl in w_0

This yields the correct truth conditions for the example in (61).

There are two steps in the above derivation that are worth highlighting. First is the effect of *copying* an F-marked constituent. In the representation of focus employed here, F-marking is encoded as an index on constituents. Movement creates a complete copy of the targeted object, including the indices.²⁷ This results in two syntactic objects in the final logical representation with identical F-indices.

Second is how multiple instances of identical F-indices are interpreted. Focus alternatives used by focus-sensitive operators such as *only* are computed by choosing a particular focus assignment function h and evaluating the whole complement of the operator with respect to that focus assignment. Therefore multiple instances of identical F-indices will always covary in the alternatives considered by a focus-sensitive operator, as in (72) above.

The covarying property of F-indices is independently motivated by Kratzer's (1991) discussion of VP ellipsis.²⁸ Kratzer (1991) considers the following conversation:

(74) Kratzer (1991, p. 830):

A: What a copy cat you are! You went to Block Island because I did. You went to Elk Lake Lodge because I did. And you went to Tanglewood because I did.

B: I *only* went to [Tanglewood]_F because you did Δ .

The question is how Speaker B's reply in (74) is interpreted. Speaker B's reply involves VP ellipsis, indicated by Δ . Intuitively, the ellipsis site is resolved by a VP of the form "go to [Tanglewood]_F." The question is how these two instances of focus—the explicitly pronounced "Tanglewood" in (74B) and the "Tanglewood" hypothesized in the interpretation of the ellipsis site—are interpreted by *only*. Kratzer (1991) argues that the correct interpretation of (74B) is to have these two instances of "Tanglewood" *covary* in the interpretation of *only*. Assuming that "Block Island" and "Elk Lake Lodge" are the alternatives to "Tanglewood," the following is the correct interpretation of (74B):

²⁷This follows from the No Tampering Condition on movement, which is also part of the original motivation for the Copy Theory itself (Chomsky, 2005).

²⁸See also Wold (1996) for additional arguments for a focus-index system and Sauerland (2007) for discussion of more complex cases involving ellipsis.

(75) Correct interpretation of (74B):

- ¬ (B went to Elk Lake Lodge because A went to Elk Lake Lodge) ∧
- ¬ (B went to Block Island because A went to Block Island)

In contrast, if we allowed the two instances of “Tanglewood” to vary independently, Kratzer (1991) points out, we predict the following, unattested readings. (The eight conjuncts here come from the nine possible alternative propositions, less the single prejacent combination.) The problematic components here are the conjuncts which involve *mismatches* between the two F-marked positions, indicated by skulls on the right.

(76) Incorrect interpretation of (74B):

- ¬ (B went to Elk Lake Lodge because A went to Elk Lake Lodge) ∧
- ¬ (B went to Elk Lake Lodge because A went to Block Island) ∧
- ¬ (B went to Elk Lake Lodge because A went to Tanglewood) ∧
- ¬ (B went to Block Island because A went to Elk Lake Lodge) ∧
- ¬ (B went to Block Island because A went to Block Island) ∧
- ¬ (B went to Block Island because A went to Tanglewood) ∧
- ¬ (B went to Tanglewood because A went to Elk Lake Lodge) ∧
- ¬ (B went to Tanglewood because A went to Block Island) ∧



The mechanism of focus alternative computation introduced by Rooth (1985, 1992b), often called “pointwise composition,” predicts this interpretation with mismatched alternatives. In this system, each instance of the same F-marked constituent will introduce alternatives independently and they will be crossed with one another, resulting in these mismatches. This issue is avoided in the semantic framework for focus introduced by Kratzer (1991) and adopted here.²⁹

First, note that VP ellipsis requires a semantically identical antecedent VP (Sag, 1976; Williams, 1977). This semantic identity extends to identity of indices on unbound variables.³⁰ This can be observed in (77) below, where the interpretation of the ellipsis site must make reference to the same individual that is referenced in the antecedent VP.

(77) Noam_i was in the office today. Yusuke saw him_i. Iain did Δ too.

- a. ✓ Δ = see him_i
- b. * Δ = see him_j

²⁹But see also Beaver and Clark (2008), who discuss other possibilities for correctly interpreting Tanglewood-type examples without the use of focus indices.

³⁰For variables which are bound in the antecedent, VP ellipsis famously allows for different interpretations for the variables in the ellipsis site. See Takahashi and Fox (2005) and references therein for discussion.

Next, recall that under our semantic framework for focus, focused constituents bear a F-index, which ellipsis identity will be sensitive to. Let's say the F-marked constituent in (74B), which is part of the antecedent VP, has F-index F1. The interpretation of the ellipsis site will then necessarily be with a "Tanglewood" with the same index, F1:

(78) LF for (74B):

I_A PAST *only* $\underbrace{\text{go to [Tanglewood]}_{F1}}_{\text{antecedent VP}}$ because you_B did $\underbrace{\text{go to [Tanglewood]}_{F1}}_{\text{elided VP}}$.

The two "Tanglewood"s in the scope of *only* carry the same focus-index and thus will covary, resulting in the correct interpretation above.

Kratzer's (1991) focus index system for the interpretation of focus was explicitly designed to yield covarying alternatives when there are multiple instances of the same F-marked constituent in the syntactic representation. Ellipsis is one such construction which can lead to two LF-identical copies of F-marked constituents. Movement under the Copy Theory is another. The focus index system straightforwardly yields the correct alternatives in cases of copied focus, without having to filter out mismatched alternatives in some way. I have therefore adopted the Kratzer (1991) system here.

2.4 The requirements of the lower copy: evidence from ellipsis

Finally, I will conclude this framework chapter with discussion of the interpretation of the lower copy of movement. Under the Copy Theory of movement and Trace Conversion, the lower copy of movement has a semantic contribution, but in many simple cases, it is difficult to pinpoint its effect. Here I will present an observation from Merchant (1999, 2001) involving ellipsis, which elucidates the contribution of the lower copy. The requirements that are introduced by the lower copy of movement will play an important part in explaining the inability of *only* to associate backwards, which I will present in chapter 4.

I begin with the most straightforward case, where the two copies in the chain are identical modulo Trace Conversion. Consider (79) below, which gives the LF derivation for a simple declarative sentence with a quantifier in object position. The quantifier "every boy" will covertly move (QR) due to a type mismatch, which I implement as copying. At this point, both the higher and lower copies of movement have the same domain, "boy." The lower copy of the movement chain is then modified into a definite description bound variable, informally "the boy ①" below. (Recall that ① is the variable with index 1: given an ordinary assignment function g , it will denote $g(1)$. The index 1 is bound above by the λ -binder.)

(79) QR and Trace Conversion at LF:

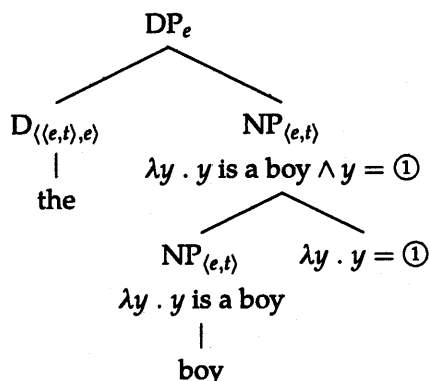
"John met every boy."

a. QR: $[\text{DP every boy}]$ John met $[\text{DP every boy}]$ illustrated in (63)

b. Trace Conversion: $[\text{DP every boy}] \lambda 1$ John met $[\text{DP the boy } \textcircled{1}]$ illustrated in (64)

"The boy $\textcircled{1}$ " is a definite description with the structure in (80), with some nodes annotated with their extensional type and denotation. This DP's restrictor is formally the intersection of those individuals that satisfy being a boy and those individuals that are equal to $\textcircled{1}$. But note that the set of individuals that are equal to $\textcircled{1}$ is simply the singleton set $\{\textcircled{1}\}$. Therefore the lower copy, "the boy $\textcircled{1}$," will necessarily denote $\textcircled{1}$.

(80) Structure of a Trace Converted lower copy:



There is, however, an important semantic difference between "the boy $\textcircled{1}$ " and the corresponding simple bound variable $\textcircled{1}$, which would be the representation of the trace position if the Copy Theory is not used. This comes from the fact that "the boy $\textcircled{1}$ " is a *definite description*. Definite descriptions introduce presuppositions of existence and uniqueness. In the case of a lower copy of movement such as "the boy $\textcircled{1}$ " this translates to a single presupposition, that $\textcircled{1}$ satisfy "boy." The presuppositions are illustrated below for the more general case of a lower copy "the NP \textcircled{i} ":

(81) Presuppositions of a Trace Converted lower copy:

"the NP \textcircled{i} "

a. Existence: there is an individual which satisfies the predicate $[\text{NP}]$ and is equal to \textcircled{i}

$\iff \textcircled{i}$ satisfies the predicate $[\text{NP}]$

b. Uniqueness: there is a unique individual which satisfies Existence above.

(Vacuous, as there is only one individual—namely \textcircled{i} —that is equal to \textcircled{i} .)

The result of this discussion is the denotation for "the $[\text{NP}] \textcircled{i}$ " given in (82), with superintensional type $\langle a, \langle s, e \rangle \rangle$. For the time being we can think of the presupposition as a definedness

condition, which I encode using the colon-dot notation of Heim and Kratzer (1998, p. 34–34, 41): given assignment function g and world w , “the NP ①” will be undefined if $g(i)$ does not satisfy $\llbracket \text{NP} \rrbracket$ in w under g . If it is defined, it will return $g(i)$.

(82) Denotation of a Trace Converted lower copy:

$$\llbracket [\text{DP the } [\text{NP } \textcircled{1}]] \rrbracket = \lambda g_a . \lambda w_s : \underbrace{\llbracket \text{NP} \rrbracket (g)(w)(g(i))}_{\text{definedness condition}} . g(i) \quad (\text{type } \langle a, \langle s, e \rangle \rangle)$$

Let’s return to example (79) above, the LF derivation for “John met every boy.” At LF, the quantifier “every boy” will take the nuclear scope predicate “ $\lambda 1$ John met the boy ①,” whose denotation is given in (83) below. We can think of this predicate as a partial function that is undefined if its individual argument is not a boy. If it is defined, its truth conditions will be the same as the predicate “ $\lambda x .$ John met x .”

(83) Denotation for the nuclear scope predicate in (79):

$$\llbracket \lambda 1 \text{ John met } [\text{the boy } \textcircled{1}] \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e : \underbrace{x \text{ is a boy in } w}_{\text{definedness condition}} . \underbrace{\text{John met } x \text{ in } w}_{\text{truth condition}}$$

To verify the truth of (79), “John met every boy,” one must step through every individual in the set of boys and check that it satisfies this predicate in (83). Luckily for each boy that we test, the boy satisfies the predicate “boy,”³¹ so the function in (83) is always defined.³²

Now let’s turn to a slightly more complex example, involving restriction of the higher copy of movement via *late merger*. Lebeaux (1988) argues that adjuncts such as relative clauses can be late-merged to a nominal after it has moved, in its higher position. This will not affect the structure of the lower copy. Here for demonstration purposes I consider a derivation for (84) where late merger modifies the domain of “every boys” after it QRs.³³

³¹Just like they say: boys will be boys.

³²It has long been noted that natural language determiners are *conservative*: for a determiner Q with domain A and nuclear scope B , $Q(A)(B) \iff Q(A)(A \cap B)$ (Barwise and Cooper, 1981; Higginbotham and May, 1981; Van Benthem, 1983; Keenan and Stavi, 1986, a.o.). Fox (2002, fn. 8) suggests that the semantics of lower copies of movement may help to ensure conservativity, crediting an unpublished Gennaro Chierchia talk from 1995 for a similar idea. Suppose $[\text{DP } Q \text{ A}]$ moves to a scope-taking position. Under the Copy Theory, its complement will be of the form $B = “\lambda x \dots [\text{the } A \text{ } x] \dots”$ As we have discussed here, B will only be defined for arguments which are members of A . In this case, verifying the truth of the $Q(A)(B)$ can only depend on verifying the membership of elements of A in B —we cannot verify the truth or falsity of membership in B for individuals which are not members of A . Therefore $Q(A)(B) \iff Q(A)(A \cap B)$ will hold. See also Romoli (2009) for a recent presentation and precisification of this explanation for conservativity.

³³The structures here illustrate the logical representations and their scopes. In order to derive the correct word order of the relative clause, we require QR to be a *rightward* movement. See Fox and Nissenbaum (1999); Fox (2002) for discussion.

As noted by Fox (2002, p. 75), there is reason to believe that the derivation I illustrate here may be strongly dispreferred compared to a derivation with no late merger involved. This derivation is presented for demonstration purposes only.

(84) **QR, late merger, and Trace Conversion:**

"John met every boy who was at the party."

a. QR: [_{DP} every boy] John met [_{DP} every boy]

b. Late merger: [_{DP} every [boy [*late-merged* who was at the party]]]; John met [_{DP} every boy];

c. Trace Conversion: [_{DP} every [boy [who was at the party]]] λ1 John met [_{DP} the boy ①]

At LF, the quantifier "every" has the restriction "boy who was at the party," with a nuclear scope predicate that is defined only for individual arguments that are boys and otherwise behaves as " λx John met x ." Evaluating the truth of the sentence involves checking individuals in "boy who was at the party" against the nuclear scope predicate. Because boys who were at the party are necessarily boys, the nuclear scope predicate will always be defined.

In both examples above, the presupposition introduced by the lower copy of movement was satisfied. In (79), the restrictors in both copies were identical; in (84), the higher copy was restricted due to late merger, and as a result was a subset of the restrictor in the lower copy. In general, if we are looking at movement chains with and without late merger on the higher copy, the requirements of the lower copy will always be satisfied in this way. It is therefore difficult to observe the effects of the presupposition hypothesized to result from the Trace Converted lower copy (81).

The effects of the presupposition of the lower copy can be verified, however, in certain complex ellipsis constructions. Informally, ellipsis allows us to swap out just the lower copy of a movement chain, testing bound variables with restrictors which might truly differ from the domain of quantification above. Here I will review a discussion of this form from Merchant (1999, 2001). We will see that the presuppositions introduced by a Trace Converted lower copy do contribute to the interpretation of sentences. Furthermore, we will see that the presupposition introduced by a Trace Converted lower copy projects universally over the bound variable's domain of quantification.

Merchant (1999, 2001) observes an interesting pattern of restrictions on the so-called "re-binding" interpretation of quantificational variables under ellipsis. As is well known, bound variables under ellipsis can be interpreted using strict or sloppy identity (85). Under strict identity, the variable in the ellipsis site is interpreted with the same referent as in the antecedent; under sloppy identity, the variable is interpreted with a different antecedent, based on the syntactic and semantic context surrounding the ellipsis site itself.

(85) **Strict and sloppy readings of bound variables under ellipsis:**

John₁ likes his mother and Bill₂ does Δ , too.

a. Δ = [like his₁ (=John's) mother]

strict identity

b. Δ = [like his₂ (=Bill's) mother]

sloppy identity

Merchant (1999, 2001) presents a more complex set of possibilities in examples where the antecedent contains a quantifier.³⁴ Example (86) below has three possible interpretations of the ellipsis site. First, there is a strict reading where the ellipsis is interpreted with an independent instance of the quantifier “many suspects” (86a). Second, there is a reading where a few suspects claimed that I hadn’t met the set of people who I did indeed meet (86b).³⁵ Third, there is a reading where the object in the ellipsis site is being interpreted as bound by the quantifier “a few (suspects)” which c-commands the ellipsis site, but not the antecedent VP. Merchant refers to this as the “re-binding” reading and I follow his term here.

(86) Readings of quantifiers under ellipsis (based on Merchant, 2001, p. 214):

I met with [many suspects]₁, though [a few (suspects)]₂ claimed I hadn’t Δ .

- a. $\Delta =$ [met with many suspects] strict identity with quantifier
- b. $\Delta =$ [met with them₁] strict identity with bound variable
- c. $\Delta =$ [met with them₂] sloppy “re-bound” variable

Here we will focus on this re-binding reading in (86c). Merchant (2001) suggests that the object “many suspects” in the antecedent QRs using the Copy Theory of movement, leaving a lower copy which is Trace Converted into the definite description “the suspect x ,” illustrated in (87) below. The subject quantifier “a few” in the ellipsis clause moves to Spec,TP, introducing a λ -binder. Under the re-binding reading, this λ -binder is used to bind not only the base position of the subject, but also the variable in the object of the ellipsis site.

(87) Deriving the re-binding reading (86c):

- a. Antecedent clause LF after QR and TC:

[many suspects] λx . I [_{antecedent} met with [the suspect x]]

- b. Ellipsis clause LF:

[a few suspects] λx . [the suspect x] claimed [I hadn’t [_{ellipsis} met with [the suspect x]]]

bound variable

Through the identity conditions on ellipsis, in (87b) the VP contains a bound variable identical to the Trace Converted lower copy of movement from the antecedent clause. Let’s think about

³⁴Merchant (2001) page numbers are given here. See footnote 23 on pp. 291–292 of Merchant (1999).

³⁵One approach to deriving this reading would be to have “many suspects” QR to a position where both the trace position of the QR-ed object in the antecedent as well as the object in the ellipsis site are in the quantifier’s scope, with the quantifier then binding both variables. Merchant (2001) instead proposes that the object position in the ellipsis site is interpreted as an E-type (donkey) pronoun. I will not discuss Merchant’s arguments here as the derivation of this reading is not important for our purposes.

the structure of (87b). The nuclear scope of the quantificational subject has two instances of "the suspect x ," which both introduce the presupposition that x must satisfy the predicate "suspect." Thinking of this presupposition as a definedness condition, this means that the nuclear scope predicate " λx . the suspect x claimed I hadn't met with the suspect x " is defined only for individual arguments that satisfy the predicate "suspect." Because the subject quantifier's domain is also "suspects," the nuclear scope predicate will always be defined; the presupposition of the lower copies will necessarily hold.

Merchant (2001) then discusses variants of the structure in (86) where the quantifier c-commanding the ellipsis site and used for re-binding (here, "a few") is given a restriction *distinct from the one used in the antecedent quantifier*. He first observes that having a disjoint restriction makes the re-binding reading impossible (88). Here we assume that "suspects" and "cops" are disjoint.

(88) **Re-binding is impossible with disjoint restrictions (based on Merchant, 2001, p. 214):**

I met with [many suspects]₁, though [a few cops]₂ claimed I hadn't Δ .

- | | | |
|----|--|--------------------|
| a. | $\Delta =$ [met with many suspects] | sloppy: quantifier |
| b. | $\Delta =$ [met with them ₁] | strict: bound |
| c. | * $\Delta =$ [met with them ₂] | sloppy: re-bound |

(89) **Intended derivation for re-binding reading (88c):**

a. Antecedent clause LF after QR and TC:

[many suspects] λx . I [_{antecedent} met with [the suspect x]]

b. Ellipsis clause LF:

[a few cops] λx . [the cop x] claimed [I hadn't [_{ellipsis} met with [the suspect x]]] 
bound variable

We can easily see why the re-binding reading is unavailable: in the intended LF illustrated above, the quantifier "a few cops" binds the variable "the suspect x ." Assuming that "cops" and "suspects" are disjoint sets, the presupposition introduced by the lower copy "the suspect x " will never be satisfied, for any choice of cop x . It is impossible to satisfy both of these restrictions simultaneously. In definedness terms, the nuclear scope predicate of "a few cops" will always be undefined, because its individual argument cannot be both a cop and a suspect. As a result, what we learn for the general case is that *the presupposition introduced by a bound variable definite description (81) must hold of the higher domain that is quantified over.*

This intuition is supported by the following asymmetry observed by Merchant (1999, 2001) in examples with contrasting NP restrictions in a subset relation. We assume here that “lifers” \subsetneq “inmates”:^{36,37}

(90) An asymmetry in re-binding with set-subset restrictions (based on Merchant, 2001, p. 215):

- a. I met with [many inmates]₁, though [a few lifers]₂ claimed I hadn't Δ .
- i. $\Delta =$ [met with many inmates] sloppy: quantifier
 - ii. $\Delta =$ [met with them₁] strict: bound
 - iii. $\Delta =$ [met with them₂] sloppy: re-bound
- b. I met with [many lifers]₁, though [a few inmates]₂ claimed I hadn't Δ .
- i. $\Delta =$ [met with many lifers] sloppy: quantifier
 - ii. $\Delta =$ [met with them₁] strict: bound
 - iii. * $\Delta =$ [met with them₂] sloppy: re-bound

Note that Merchant presents the contrast in (90) with three combinations of quantifiers: “every” \sim “many,” “every” \sim “most,” and “most” \sim “many,” all of which pattern with the contrast in (90). My example (90) here demonstrates the contrast for “many” \sim “a few.” I have also reproduced the same contrast with “many” \sim “a,” which I will return to below.

As per Merchant's (2001) Copy Theoretic analysis of the re-binding reading (see 87 above), for the re-binding reading of (90a), we have the quantifier “a few lifers” binding a definite description bound variable “the inmate x ” in the ellipsis site. This is illustrated below in (91). In contrast, in the minimally contrasting example (90b), “a few inmates” binds the variable “the lifer x ” in the ellipsis site, illustrated in (92).

³⁶Merchant uses “lifers” and “inmates,” but for some speakers, these examples were easier to judge with other pairs of predicates in a subset relation, such as “syntacticians” \subsetneq “linguists.”

³⁷Merchant (2001, pp. 214–215) notes that this same pattern also holds of non-ellipsis variants with unstressed pronouns:

- (i) a. I met with [many inmates]₁, though [a few lifers]₂ claimed I hadn't met with them_{1/2}.
 b. I met with [many lifers]₁, though [a few inmates]₂ claimed I hadn't met with them_{1/*2}.

Merchant suggests that the bound variable pronouns in (i) should therefore be analyzed as minimal spell-outs of rich definite descriptions; i.e. “them” is the spell-out of a DP akin to “that inmate” in (ia) and “that lifer” in (ib). See also Elbourne (2002); Schlenker (2005).

(91) Ellipsis clause LF for (90aiii):

[a few lifers] λx . [the lifer x] claimed [I hadn't [ellipsis met with [the inmate x]]] ○
↑ bound variable

(92) Ellipsis clause LF for (90biii):

[a few inmates] λx . [the inmate x] claimed [I hadn't [ellipsis met with [the lifer x]]] ×
↑ bound variable

The availability of the re-binding reading in (90aiii/91) but not (90biii/92) shows a clear sensitivity to the logical relationship between the predicates involved. Because (90aiii/91) is grammatical, it cannot be the case that the domain of quantification is required to match the NP domain of lower bound variable definite descriptions. Instead, what we observe is a requirement that the domain of higher quantification be a *subset* of the restrictor of lower definite description bound variables: the contrast corresponds to the fact that lifers are all inmates, but by assumption not all inmates are lifers.³⁸ This is precisely what is predicted by the view above, that the NP restrictor of the lower bound variable is projected as a presupposition and must be satisfied by the domain of quantification, but not vice versa.³⁹

More formally, in terms of presupposition projection—to be discussed in more detail in section 3.2—we learn that the existence presupposition of definites project universally. For Trace Converted lower copies of movement, this existence presupposition manifests itself as a requirement that the bound variable be in a particular restrictor, as in (81). The re-binding configuration constructed by Merchant (2001) uses ellipsis to introduce a Trace Converted lower copy of movement from one movement chain (the QR step in the antecedent clause) into the scope of a different quantifier as a bound variable.

For the purposes of later discussion, it is important to note here that Merchant's contrast extends to cases where the quantifier that "re-binds" the bound variable in the ellipsis site is an *indefinite*:

³⁸ Although note that requiring the higher and lower restrictors to be identical was already a non-starter, given the availability of late merging to the restrictor of the higher copy, as in example (84) above.

³⁹ This requirement that the presuppositions of lower copies must be satisfied by the higher domain of quantification predicts that late merger cannot target the *lower* copy of a movement chain, without also affecting the higher copy. Doing so will mean that the lower copy of the movement chain will have a more restricted domain, whose presupposition will not hold of the domain of the higher copy. Therefore, we predict that a derivation that involves restriction of the lower copy alone is banned for semantic reasons.

(93) **The re-binding asymmetry with an indefinite:**

- a. I met with [many inmates]₁, though [a lifer]₂ claimed I hadn't Δ .
 $\checkmark \Delta = [\text{met with him}_2]$
- b. I met with [many lifers]₁, though [an inmate]₂ claimed I hadn't Δ .
 $* \Delta = [\text{met with him}_2]$

This contrast in (93)—parallel to that in (90) and the other combinations presented in Merchant (1999, 2001)—shows that the existential presupposition of the definite description bound variable must hold of the entire higher domain of quantification when the quantifier is indefinite “a,” just as it does with other quantifiers.

2.5 Summary

In this chapter I lay the groundwork and present the syntactic and semantic assumptions that I will make in this dissertation. In section 2.1 I presented background on the Roothian Alternative Semantics approach to the interpretation of in-situ focus. I presented my formalism for interpreting focus based on Kratzer (1991) and my denotation for *only*. In section 2.2 I presented the Copy Theory of movement and discussed the Takahashi and Hulsey (2009) approach to the A/\bar{A} distinction. In section 2.3 I showed how the Copy Theory of movement can combine with the semantic framework I developed in section 2.1. I discussed configurations where F-marking (focus indices) are copied and their similarities to ellipsis configurations considered by Kratzer (1991). Finally, in section 2.4 I discussed the semantic contribution of the lower copy of movement, which will play an important role in my proposal in chapter 4.

Chapter 3

Refining the semantics of *even*

In chapter 1 I introduced the problem of *backwards association*: focus-sensitive operators vary in their ability to associate with F-marked material which has moved out of their scope. This configuration is schematized in (94).

(94) The backwards association schema:

$\alpha_F \dots [Op [\dots \underline{\quad} \dots]]$ (with α interpreted as the associate of the operator)

Before we tackle this problem of backwards association, however, there is a more foundational question that must be addressed. What happens when there is variable binding into the scope of a focus-sensitive operator? In particular, one common way in which this would happen is through *movement*, as is schematized using the movement of a quantificational DP in (95). The configuration in question is not one of backwards association; *even* is associating with an F-marked constituent in its scope.

(95) Moving quantificational material out of the scope of *even*:

$[DP \text{ Quantifier NP}] \lambda x \dots [even [\dots (\alpha_F) \dots x \dots (\alpha_F) \dots]]$

Notice that the scope of *even* includes a variable x . This variable is bound by the moved DP's λ -binder above when computing the assertion. However, there is a question as to how this variable will be interpreted in the evaluation of *even*. In this chapter I will establish the semantic contribution of *even* in such configurations. The result will be a refinement to the semantics of *even* and its projection behavior, which will inform my discussion of backwards association in chapter 4.

I will relate the behavior of *even* here to other, better studied cases of presupposition projection through quantifiers. For many of these established cases, the interpretation of variables in a

presupposition's content varies depending on the quantifier involved. Importantly, I show that the so-called "bound" reading, famously available with many other presupposition triggers, is not available with *even*. Instead, the inference of *even* projects universally through all quantifiers, modulo a small number of exceptions. I will formalize this as an instance of *generic quantification*.

Note that the material in this chapter is in some sense a detour from the "main event," my analysis for patterns of backwards association. It is, however, a necessary step in establishing a detailed working semantics for *even*, which I will make use of in my proposal. What I show in this chapter is that the full content of the inference of *even* is computed in its interpreted position, and is not dependent on the syntactic material above it. In particular, unbound variables in the scope of *even* will be closed under generic quantification. While the argumentation in this chapter is important to establish this view, readers who wish to continue with the main proposal in chapter 4 could do so at this point, keeping this short summary in mind.

3.1 The established semantics of *even*

I will begin in this section by presenting the established semantics of *even* by way of example. Imagine that we travel to Taiwan with our friend Donna. Donna generally doesn't like trying new foods. But tonight when we went to the night market, she was feeling more adventurous. In this context, I could utter example (96) and my use of *even* would be felicitous.

(96) Donna *even* ate the [stinky tofu]_F.

Let's consider the contribution of *even* in this example. Since Horn (1969), it is widely agreed upon that *even* has no effect on the truth conditions of an utterance. Therefore the asserted content of (96) is the same as the utterance without *even*, i.e. the *prejacent proposition* (97a). The contribution of *even* is instead in the form of a projective, non-assertive meaning, referred to as its *scalar inference* below (97b).⁴⁰

(97) An approximate semantics for (96):

"Donna *even* ate the [stinky tofu]_F."

- a. assertion: Donna ate the stinky tofu.
- b. \sim scalar inference: It is less likely for Donna to eat the stinky tofu than for her to eat other things.

⁴⁰Some authors propose that *even* also projects an *additive* inference—for (96), that Donna ate something else besides stinky tofu—but this has been controversial. Here I will concentrate on the scalar inference of *even*. See Wagner (2013) for recent discussion and a review of arguments for and against the existence of the additive inference.

Here I use the term “inference” to refer to a projective, non-assertive component of meaning, without taking a stand on precisely what form of projective meaning is involved. The scalar inference of *even* has been described as a conventional implicature by some and as a presupposition by others, although much of this literature has been affected by an early terminological problem. Karttunen and Peters (1979), who first formalized the scalar part of *even*, refers to the scalar inference of *even* as a “conventional implicature,” but does so together with many other constructions that would now be called presupposition triggers in modern parlance. See Potts (2005) section 2.2 for a brief history of these terms.

A compositional semantics for the adverb *even* can be defined straightforwardly as follows, parallel to my entry for *only* in (42) in section 2.1. Here I encode the content of the scalar inference as a definedness condition:⁴¹

(98) Semantics for *even* (first version):

$$\begin{aligned}
 \llbracket \textit{even} \rrbracket_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle \rangle} &\equiv \\
 \lambda g_a \cdot \lambda w_s \cdot \lambda P_{\langle a, \langle s, t \rangle \rangle} : \forall h \in H(h \neq h_0 \rightarrow &\underbrace{P(h_0 \parallel g)}_{\text{prejacent}} <_{\text{likely}} \underbrace{P(h \parallel g)}_{\text{alternative}}) \cdot \underbrace{P(h_0 \parallel g)(w)}_{\text{truth condition}} \text{ is true} \\
 &\underbrace{\hspace{15em}}_{\text{definedness condition (scalar inference)}}
 \end{aligned}$$

I assume that the adverb *even*, just like the adverb *only*, adjoins to a projection of propositional type (e.g. *vP*) and is interpreted in its surface position. An alternative proposal, where the scope of *even* differs from its surface position, will be discussed and dismissed in section 6.2.

Now I will compute the semantics of our example sentence in (96) using this definition for *even*. Here for presentational purposes I present the semantic computation with the subject “Donna” within the scope of *even* at LF (99). We can think of this as an application of syntactic reconstruction. This step may seem like a harmless simplification, but we will see later in this chapter that reconstruction of the subject in this way will not generally yield the correct inferences of *even*.⁴² However, this simplification will suffice for present purposes.

(99) LF for (96):

even [_{vP} Donna ate the [stinky tofu]_{F1}]

An additional point of debate surrounds the choice of scale in the scalar inference. Here I assume that the scalar inference of *even* can be expressed in terms of “likelihood,” with the prejacent being “less likely” (<likely) than alternative propositions. See Kay (1990) for arguments that the scalar inference is that the prejacent is “more informative,” or “more noteworthy” in Herburger’s (2000) terms. None of the following discussion will hinge on this issue.

⁴¹Recall that the notation $h \parallel g$, read as “*h* or else *g*,” represents function combination, as defined in (31).

⁴²Reconstruction will be discussed in much further detail in chapter 7.

I also assume that the relevant alternatives to “stinky tofu” with F-index F1 are other foods found in night markets. Each of these alternatives, including the prejacent value “stinky tofu,” corresponds to a different focus assignment function in H :

(100) **Focus assignment functions in H :**

- a. $h_0(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is stinky tofu in w
- b. $h_1(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a meat bun in w
- c. $h_2(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is fried chicken in w

The semantic value for vP in (99) can be computed straightforwardly using our definition for $[\cdot]$ (30). Here is the denotation for vP given an ordinary assignment function g and a focus assignment function h :

(101) **Denotation for vP :**

$$[[vP]](h \parallel g) = \lambda w . \text{Donna ate the } h(\text{F1})(g)(w) \text{ in } w$$

This denotation includes a focus-assignment-dependent denotation for the node with F-index F1: $h(\text{F1})(g)(w)$. Note that this is always a predicate of extensional type $\langle e, t \rangle$, providing a domain for the definite determiner. We can explicitly evaluate this denotation against each focus assignment function $h \in H$ to yield the following alternative propositions:

(102) **Alternative propositions for vP :**

- a. $[[vP]](h_0 \parallel g) = \lambda w . \text{Donna ate the stinky tofu in } w$
- b. $[[vP]](h_1 \parallel g) = \lambda w . \text{Donna ate the meat buns in } w$
- c. $[[vP]](h_2 \parallel g) = \lambda w . \text{Donna ate the fried chicken in } w$

Using the definition of *even* in (98) above, we yield the following as the interpretation of (96):

(103) **Computed interpretation of (96):**

a. Assertion of (96):

$$[[vP]](h_0 \parallel g)(w_0) \text{ is true} \iff \text{Donna ate the stinky tofu in } w_0$$

b. Inference of (96): $\forall h \in H (h \neq h_0 \rightarrow [[vP]](h_0 \parallel g) <_{\text{likely}} [[vP]](h \parallel g))$

$$\iff [[vP]](h_0 \parallel g) <_{\text{likely}} [[vP]](h_1 \parallel g) \wedge [[vP]](h_0 \parallel g) <_{\text{likely}} [[vP]](h_2 \parallel g)$$

$$\iff ((\lambda w . \text{Donna ate the stinky tofu in } w) <_{\text{likely}} (\lambda w . \text{Donna ate the meat buns in } w)) \wedge ((\lambda w . \text{Donna ate the stinky tofu in } w) <_{\text{likely}} (\lambda w . \text{Donna ate the fried chicken in } w))$$

One property of the definition used here is that the prejacent proposition—in (96) above, that “Donna eat the stinky tofu”—is required to be the absolute lowest on the scale ordered by $<_{\text{likely}}$. In

terms of our formulation from (98) above, this is due to our use of the universal quantifier \forall over all focus assignment functions in H . Intuitively, this may feel like too strong a condition on the felicitous use of *even*. This point is made by Kay (1990), who gives the following examples, which demonstrate that the stated associate of *even* need not be the absolute least likely alternative.

- (104) Not only did Mary win her first round match, she even made it to the semi-finals. (Kay, 1990, ex. 82)
- (105) The administration was so bewildered that they even had lieutenant colonels making major policy decisions. (Kay, 1990, ex. 83)

Kay (1990) points out, the acceptability of (104) does not require an unusual context in which reaching the semi-finals, as opposed to winning the competition, is considered to be the least likely achievement. Likewise, in (105), lieutenant colonels are not the least likely to make major policy decisions. More extreme assertions can be made if majors, captains, or sergeants would make major policy decisions.

Lycan (1991), discussing Kay's examples, sums this up nicely:

"But universal quantifiers have the habit of being universal... there can be no exceptions. Yet *even* does seem to admit exceptions." (Lycan, 1991, p. 140)

A solution would be to quantify *almost* universally over the set of alternatives. I will use *generic quantification* here, with the operator GEN, which we can think of as universal quantification but allowing some exceptions.⁴³ A formalization for *even* which quantifies generically over the set of alternatives is given in (106):

- (106) Semantics for *even* (second version, with generic):

$$\llbracket \textit{even} \rrbracket_{(a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle)} \equiv \lambda g_a \cdot \lambda w_s \cdot \lambda P_{\langle a, \langle s, t \rangle \rangle} : \underbrace{\text{GEN } h \in H (h \neq h_0 \rightarrow P(h_0 \parallel g) <_{\textit{likely}} P(h \parallel g))}_{\text{scalar inference}} \cdot \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\text{truth condition}}$$

With this tolerance for exceptions added, a scalar inference computed as in (99) accords with our intuitions regarding the scalar part of (96) and its felicity conditions. However, in order to compute this meaning of *even*, in (99) I chose to interpret an LF where no material has moved out of the scope *even*. What would have happened if we evaluated the subject, "Donna," in its surface position, outside of the scope of *even*? Or what if the subject, evaluated outside of the scope of

⁴³See Carlson and Pelletier (1995, ch. 1) and references therein for previous work on generic quantification and different formulations thereof. For the universal-with-exceptions characterization of GEN that I use here, the number of exceptions tolerated will be contextually determined.

even, were quantificational? In order to answer these questions, I will briefly review work on the projection of presuppositions through quantification in the next section, and then introduce new facts on the interpretation of *even* under quantification.

3.2 Projection through quantification

The so-called *projection problem* concerns how the presupposition of a complex sentence is derived from the presuppositions of its parts. In particular, here we are interested in how the presupposition triggered by a *presupposition trigger* manifests itself when a quantifier above the trigger binds a variable in the trigger's scope. This configuration is schematized here:

(107) **Quantifying into the scope of a presupposition trigger:**

$$\underbrace{[\text{DP Quantifier Domain}] \lambda x \dots [\text{presupposition trigger} \underbrace{[\dots x \dots]}_{\text{scope of the trigger}}]}_{\text{scope of the quantifier}}$$

The content of presuppositions generally does not compose with material above it, projecting instead through logical operators such as negation. When the trigger computes its presupposition based on the material in its complement, this complement will be assignment-dependent, reflecting the free variable x in the scope of the trigger in (107).

The descriptive facts regarding the presuppositions inferred from quantified sentences of the form in (107) are both complex and hotly contested.⁴⁴ Here I will highlight two approaches to interpreting this variable, corresponding to two types of readings which have been claimed to be possible. This discussion will be based on the work of Sudo (2012). Suppose the content of the presupposition, abstracted over the variable (x in (107)), is the predicate $P_{\langle e,t \rangle}$. There are two approaches to the interpretation of this variable:

(108) **Two approaches to interpreting the unbound variable:**

- a. The "bound" reading: P holds for the set of individuals that verifies the use of *Quantifier* in the assertion, i.e. the *witness set* of the quantifier. (See also (113b) and fn. 47 below.)
- b. Universal projection: P holds for all individuals in the domain of the quantifier, i.e. $\sim \forall x \in [\text{Domain}] . P(x)$

Let us consider each of these proposals in turn. Karttunen and Peters (1979) and van der Sandt (1992) discuss examples where quantifiers appear to be able to bind into the content of presuppo-

⁴⁴See Sudo (2012) for a recent review of the literature, and in particular his section 3.3 for a review of experimental evidence investigating this question.

sitions in tandem with quantifying in the assertion.⁴⁵ For example, the use of *again* in (109) below introduces a presupposition that one of the students—specifically, that student who just became pregnant—was pregnant previously.

(109) A student became pregnant again.⁴⁶

asserts: A student became pregnant.

~ “bound” presupposition: She_{the student who became pregnant} was pregnant before.

Crucially, this presupposition makes reference to those individuals who are referenced by the assertion, namely the witness set of the existential quantifier. It does not project using a presupposition using its own existential quantifier, that some of the students (who may or may not be pregnant now) were pregnant before (110a). Nor does it project a universal inference, that all of the students were pregnant before (110b).

(110) **Incorrect presuppositions of (109):**

a. Existential presupposition: Some student was pregnant before.

b. Universal presupposition: All of the students were pregnant before.

Sudo (2012, §4.1) discusses the quantifier “exactly one” extensively and argues that it projects a bound presupposition, similar to (109) above.

(111) *Exactly one* of the students became pregnant again.

asserts: Exactly one of the students became pregnant.

~ “bound” presupposition: She_{the student who became pregnant} was pregnant before.

In contrast, presuppositions have been claimed to project universally through negative quantifiers (Cooper, 1983; Heim, 1983). This is reflected clearly by the presupposition of (112) below.

(112) *No* student became pregnant again.

asserts: No student became pregnant.

~ universal presupposition: All of the students were pregnant before.

I note in passing that it has been suggested that cases of so-called universal projection, particularly with lexical presupposition triggers, sometimes allow for some exceptions. A solution would

⁴⁵Here the presupposition trigger *again* is used instead of *manage* as in the classic example from Karttunen and Peters (1979), *Someone managed to succeed George V on the throne of England.*, as intuitions regarding the inference of *manage* seem to be less clear. See also Sudo (2012, fn. 54) for discussion.

⁴⁶The example has an additional reading where *again* takes scope over the whole description “a student became pregnant.” This is not the reading we are interested in here.

be to model the universal presupposition using *generic* quantification rather than a strict universal. (See Sudo (2012, §4.1.4) for discussion.) I will return to this point later when discussing *even*.

Sudo (2012) relates the projection behavior of presuppositions under different quantifiers to the availability of different types of *discourse anaphora*. I define the following terms:

(113) **Definitions of some sets in the context of quantification:**

Suppose “Quantifier(Domain_(e,t)) (Predicate_(e,t))” is true. Then define:

- a. *domain* ≡ [[Domain]]
- b. *witness* ≡ the set of individuals in Domain whose satisfaction of Predicate allowed for the truthful use of “Quantifier(Domain)(Predicate).”
- c. *complement* ≡ [[Domain]] \ [[Predicate]]

(114) **Example:** *Two students played basketball.* is true because Isaac and Junya played basketball.

- a. *domain* = [[students]]
- b. *witness* = {Isaac, Junya}
- c. *complement* = [[students]] \ {Isaac, Junya}

These *domain*, *witness*, and *complement* sets correspond to different types of discourse anaphora. Each of these anaphors can be used cross-sententially under certain situations, as illustrated through the following examples modified from Sudo (2012, §5.3):⁴⁷

(115) *A student criticized John.*

- a. ✓ *He/she*_{the student who criticized John} *was very aggressive.* *witness*
- b. * *They*_{the students} *are all very sharp.* *domain*

(116) *No student criticized John.*

- a. * *They/he/she*_{the students that criticized John = ∅} *were/was very aggressive.* *witness*
- b. ✓ *They*_{the students} *are a good group.* *domain*

(117) *Few students showed up today. They*_{the students who *didn't* show up} *were all sick.* *complement*

⁴⁷Note that there are contexts where could be more than one set that verifies the use of the quantifier, so the witness set may not be unique. See for example Kadmon (1990, pp. 281–282) for discussion of cross-sentential anaphora in such situations.

My reference to a “witness set” anaphor in examples such as (115) differs slightly from the presentation in Sudo (2012, 2014), which I am otherwise following. The corresponding variety of cross-sentential anaphora is described by Sudo (2012, p. 89) as a “nuclear scope anaphor” and by Sudo (2014) as “refset anaphora.” Yasutada Sudo (p.c.) states that in situations where there is no unique witness set, his “refset anaphora” is predicted to behave as an indefinite that ranges over witness sets.

Sudo (2012) notes in particular that different quantifiers license different cross-sentential anaphora, and that this corresponds to the projection behavior of presuppositions under different quantifiers. For example, indefinites with *a* license witness anaphora but not domain anaphora (115); this corresponds to the fact that the example (109) with *a* introduces a “bound” presupposition rather than a universal one. In contrast, the negative determiner *no* in (116) licenses domain anaphora but not witness anaphora; this corresponds to the fact that the example (112) with *no* projects universally, and does not have a “bound” reading. Sudo shows that presuppositions do not make reference to complement anaphora (Sudo, 2012, §5.4.1), even though it can be used cross-sententially.

In addition, pronouns can always refer to a previous definite antecedent. This explains the availability of cross-sentential reference to the domain of students in the following example with a partitive construction, again taken from Sudo (2012, §5.3), in contrast to its non-partitive counterpart above. This effect contributes to the availability of universal presupposition readings with partitive quantifiers.

- (118) [One of [the students]] criticized John. (cf 115)
- a. ✓ He/she_{the student that criticized John} was very aggressive. *witness*
- b. ✓ They_{the students} are all very sharp. *definite antecedent*

To summarize, Sudo’s (2012) proposal is that unbound variables in the content of presuppositions are interpreted using *cross-dimensional anaphora*. The so-called “bound” reading of presuppositions involves witness anaphora, whereas cases of universal projection involves domain anaphora. These two options for projection are summarized in (119) below. Sudo’s viewpoint will be helpful when we turn to the projection of *even* in the following section.⁴⁸

(119) Options for presupposition projection through quantification:

$$[\text{DP Quantifier Domain}] \lambda x \dots [\text{presupposition trigger } \underbrace{[\dots x \dots]}_{\text{scope of the trigger}}]$$

scope of the quantifier

where *presupposition trigger* $\rightsquigarrow f(x)$

- a. “bound” inference: $\rightsquigarrow f(x)$, where *x* is the *witness* of the higher Quantifier
- b. universal inference: $\rightsquigarrow \forall x \in \text{domain} . f(x)$

⁴⁸Sudo (2012) presents a dynamic semantics unifying these cross-sentential and cross-dimensional uses. This formalism will not be adopted here.

3.3 The projection of *even* through quantification

We are now in a position to investigate how the scalar inference of *even* projects through quantification. The configuration I will study here is repeated below in (120). Based on the discussion in the previous section, we can imagine the unbound variable in the scope of *even* being interpreted in one of two ways: a “bound” reading, where the presuppositional content must hold of the individuals which verify the quantifier in the assertion (120a), and a universal reading, where the presuppositional content must hold of all individuals in the domain of quantification (120b). In terms of Sudo’s (2012) theory of cross-dimensional anaphora, we can think of these as utilizing a witness anaphor or a domain anaphor.

(120) **Options for the projection of *even* through quantification:**

[_{DP} Quantifier Domain] $\lambda x \dots [\textit{even} [\dots (\alpha_F) \dots x \dots (\alpha_F) \dots]]$
scope of *even*

- a. “bound” inference: Let $x = \textit{witness}$; the prejacent proposition with x is less likely than the alternative propositions with x
- b. universal inference: $\forall x \in \textit{domain}$. the prejacent proposition with x is less likely than the alternative propositions with x

I will show that the inference of *even* lacks the “bound” projection option and instead projects universally through all different quantifiers. This contrasts with the general behavior of presupposition triggers reviewed in the section above, where the projection behavior varied with the choice of higher quantifier. As noted above, the universal projection option actually allows for a small number of exceptions, and I will therefore characterize it as *generic* quantification using the operator GEN.

I begin with example (121), which shows that a non-binding, generic inference reading is possible. (The part of the context critical for the evaluation of the scalar inference of *even* is italicized.)

(121) **Cleaning for the party:**

Context: Ruth, Chris, and Anthony are hosting a party and have decided to clean their apartment. *They’re generally happy to clean any room, with the exception of one particular room, where people do have strong opinions: the bathroom. Ruth and Chris hate cleaning the bathroom, but Anthony actually really enjoys it.*

Ruth cleaned the kitchen. Chris cleaned the living room. Anthony cleaned the bathroom. Right before the party they confer to make sure that all the cleaning is done. Ruth comments:

✓ “I think we’re ready for the party. Someone *even* cleaned the [bathroom]_F.”

The use of *even* is felicitous in this context. Importantly, the discourse participants in this context all know that it is in fact Anthony who has cleaned the bathroom.⁴⁹ This context shows that the inference introduced by *even* here must not be referring specifically to the actual individual who did the cleaning (122a). This inference (122a) would be that it is less likely for Anthony to have cleaned the bathroom than other rooms, which is not supported by the context. I propose that the correct inference of *even* is instead in (122b): in general, people are relatively unlikely to clean the bathroom, as compared to other rooms.

(122) **The interpretation of (121):**

asserts: Someone x cleaned the bathroom. (verified true by $x = \text{Anthony}$)

a. "bound" inference: infelicitous in context

$(\lambda w . x \text{ cleaned the bathroom in } w) <_{\text{likely}} (\lambda w . x \text{ cleaned } <\dots\text{alternative}\dots > \text{ in } w) \dots$ ⁵⁰

b. available inference with generic quantification: felicitous in context

GEN y (y is a person $\rightarrow (\lambda w . y \text{ cleaned the bathroom in } w) <_{\text{likely}}$
 $(\lambda w . y \text{ cleaned } <\dots\text{alternative}\dots > \text{ in } w) \dots$)

Example (121) shows that a non-bound reading—where the content of the scalar inference is not tied to the actual individual who cleaned the kitchen—is possible.

One might imagine that the inference of *even* here is derived through reconstruction of the subject existential "someone" into the scope of *even*. For example (121) above, this approach yields the scalar inference in (123b), which also holds in this context.

(123) **Alternative computation for (121) using reconstruction:**

a. LF: *even* [someone cleaned the [bathroom]_F]

b. predicted inference from reconstruction: felicitous in context

$(\lambda w . \exists y . y \text{ cleaned the bathroom in } w) <_{\text{likely}} (\lambda w . \exists z . z \text{ cleaned } <\dots\text{alternative}\dots > \text{ in } w)$

...

⁴⁹This point can be made even more sharply using a variant of (121) where it is Anthony himself that utters the target sentence, "Someone *even* cleaned the [bathroom]_F." This too is judged as felicitous.

I've attempted to keep the target sentence in (121) simple. Another, very natural variant would be where the Dean shows up to the party, and then Anthony says to Ruth and Chris, "Aren't you glad someone *even* cleaned the [bathroom]_F this time?" This utterance is judged as very natural in this context and makes the same point as what we learn from (121).

⁵⁰The notation " $<\dots\text{alternative}\dots >$ " is used to stand in for an alternative denotation determined by the context.

The truth conditions of the inference in (122b), using wide-scope generic quantification, and (123b), with narrow-scope existential quantification, are often difficult to tease apart.⁵¹ The relationship between these formulations is discussed in the appendix to this chapter.

There are two reasons to believe that it is not the case that the correct inference of *even* in such cases is always derived through reconstruction of the quantifier. That is, there exists another possible strategy for interpreting the inference of *even* with quantificational material which has moved out of its scope, schematized in (120) above.⁵² The first reason comes from constructions where reconstruction is not an option, for example with subject control. I follow the widely-adopted view that subjects of control embeddings are a PRO bound by the matrix subject; the matrix subject does not originate within the scope of the control verb. This corresponds to the fact that the matrix subject is unable to take scope below the control verb—(124a) is an attempt at such a reading. The sentence also lacks a reading with the quantifier interpreted as both the matrix and embedded subject (124b), as discussed by Partee (1975) as an early problem for equi-NP analyses of control.⁵³

- (124) Everyone_i wants [PRO_i to clean the bathroom].
- a. ≠ It is wanted [for everyone to clean the bathroom].
- b. ≠ Everyone wants [everyone to clean the bathroom].

⁵¹The following example is potentially an argument for the wide-scope generic quantification over the narrow-scope existential quantification:

- (i) Different people have different jobs. Some people are firefighters. Some people are police officers. Someone is *even* the [President of the United States]_F. (Evan Jenkins, p.c.)

This use of *even* in (i) is felicitous. The existential inference in (ii) below is false, given the world knowledge that a President exists; in contrast the generic inference in (iii) is appropriate.

- (iii) **Predicted existential inference:** infelicitous
It is less likely that there exists someone who is the President, than that there exists someone who is <...*alt*...>.
- (ii) **Predicted generic inference:** felicitous
Given a generic person *x*, it is less likely that *x* is the President than that *x* is <...*alternative*...>.

However, a concern about this example is that it seems to be in some children's book register, which may be essential for the example. In particular, suppose that in this mode, the speaker does not assume that the listener already knows that there exists a person who is the President. If that is indeed the case, both predicted inferences could be felicitous in the conversation's common ground, and we cannot use this example as an argument for the generic inference.

⁵²Note that this discussion is independent of whether reconstruction is possible in a particular configuration, and how such reconstruction interacts with focus interpretation. This will be the subject of chapter 7.

⁵³See Partee (2004, chapter 1) for historical context.

Now consider example (125). This example projects the same scalar inference as (122b).⁵⁴ Because reconstruction of the subject into the scope of *even* is not possible, I conclude that the observed scalar inference of *even* does not rely on reconstruction of the subject into its base generated position in the scope of *even* illustrated in (123) above.

- (125) Someone wants to *even* clean the [bathroom]_F.
- a. LF: someone_i wants [*even* PRO_i clean the [bathroom]_F]
- b. inference: GEN x (x is a person \rightarrow ($\lambda w . x$ clean the bathroom in w) \langle_{likely}
 $(\lambda w . x$ clean $\langle \dots \text{alternative} \dots \rangle$ in w) ...)

The second reason that reconstruction is not the general solution for computing the correct inference of *even* is that in some cases reconstruction predicts an incorrect inference. Consider example (126) below, which involves the quantifier “exactly one” in place of the existentials that were used above.

- (126) **Hide and Seek:** (based on Irene Heim, p.c.)
- Context: Some students brought their kids to the party at Ruth, Chris, and Anthony’s place. The kids decided to start playing Hide and Seek. *There are many places to hide at the party, but some spaces are large enough for multiple kids to hide, while others, such as the broom closet, can only fit one child.*
- After the kids took their places, Snejana went to find them. Afterwards, she told us where she found the kids: “Exactly one kid was hiding in the garden. Exactly one kid was hiding in the basement. Exactly one kid was hiding in the living room.
- ✓“Exactly one kid was *even* hiding in the [broom closet]_F.”

Reconstruction of the subject within the scope of *even* predicts the inference in (127b) below. This inference is predicted to be infelicitous in this case, as it is relatively *more* likely for exactly one kid to hide in the broom closet, as opposed to other locations, due to its size. The felicity of (126) shows us, again, that the scalar inference of *even* is not derived through reconstruction of the subject into the scope of *even*.

⁵⁴It is difficult to intuit whether the scalar inference from (125) is really a claim regarding the relative likelihoods of *cleaning*, as in (125b), or is a claim regarding the relative likelihoods of *wanting to clean*. Here I assume, as I have thus far, that adverbial *even* takes scope in its surface position. In section 6.2, I will discuss and argue against views where *even* takes scope higher than its surface position.

(127) **The interpretation of (126):**

asserts: Exactly one kid x was hiding in the broom closet.

- a. “bound” inference: may or may not hold in context
 $(\lambda w . x \text{ was hiding in the broom closet in } w) <_{\text{likely}} (\lambda w . x \text{ was hiding in } <\dots\text{alt}\dots> \text{ in } w) \dots$
- b. unavailable inference using reconstruction: infelicitous in context
 $(\lambda w . \text{ exactly one kid was hiding in the broom closet in } w) <_{\text{likely}}$
 $(\lambda w . \text{ exactly one kid was hiding in } <\dots\text{alternative}\dots> \text{ in } w) \dots$
- c. inference with generic quantification: felicitous in context
 GEN y (y is a child $\rightarrow (\lambda w . y \text{ was hiding in the broom closet in } w) <_{\text{likely}}$
 $(\lambda w . y \text{ was hiding in } <\dots\text{alternative}\dots> \text{ in } w) \dots$)

I believe the correct inference of (126) is that in (127c) which claims that, in general, kids in this scenario are less likely to hide in the broom closet than to be elsewhere. However, this particular example and context does not serve to rule out the “bound” reading in (127a), that the specific individual who was hiding in the broom closet was less likely to hide there as opposed to in other rooms.

The following example (128) helps to distinguish between the generic and bound readings, showing that a “bound” reading for *even* is in fact impossible.

(128) **Buying beer:**

Context: Everyone who was invited to the party at Ruth, Chris, and Anthony’s apartment was asked to bring something to drink. *Generally, people bring beer or wine to such parties. But mitcho doesn’t drink alcohol and always brings something non-alcoholic to parties.*

When mitcho went to the store, they were all out of pop, so he reluctantly bought some beer. This turned out to be a fortuitous decision, as none of the other guests chose to bring beer. After mitcho arrives, Ruth looks over the collection of drinks people brought and says:

“This will be a good party. Many people brought wine. Some people brought whiskey.

“Someone *even* brought some [beer]_F.”

The use of *even* is infelicitous in this context.⁵⁵ An inference tied to the individual who actually brought the beer (129a) is predicted to be true in this context. As established by the context, it is more likely for mitcho to bring some non-alcoholic beverage than to bring beer. Hence, this context shows that this bound reading is unavailable. An inference which is based instead on the relative likelihood that beer will be brought by someone, as opposed to alternative beverages, (129b)—

which I take to be a possible interpretation based on (122) above—is predicted to be infelicitous here, as it is generally more likely for people to bring beer as opposed to other beverages.

(129) **The interpretation of (128):**

asserts: Someone x brought some beer. (verified true by $x = \text{mitcho}$)

a. unavailable “bound” inference: felicitous in context

$(\lambda w . x \text{ brought some beer in } w) \langle_{\text{likely}} (\lambda w . x \text{ brought } \langle \dots \text{alternative} \dots \rangle \text{ in } w) \dots$

b. inference with generic quantification: infelicitous in context

GEN y (y is a person $\rightarrow (\lambda w . y \text{ brought some beer in } w) \langle_{\text{likely}}$
 $(\lambda w . y \text{ brought } \langle \dots \text{alternative} \dots \rangle \text{ in } w) \dots$)

This behavior contrasts markedly with the behavior of the presupposition of *again* with *someone*, which crucially has a bound reading:

(130) *Someone* became pregnant again. (\approx 109)

asserts: Someone became pregnant.

\sim “bound” presupposition: She_{the person who became pregnant} was pregnant before.

The examples here show that a bound reading of the scalar inference of *even* is unavailable, regardless of the quantifier involved. This clearly contrasts with the behavior of other presuppositions, discussed in the previous section. Moreover, the inference computed is not derived by reconstruction of the subject into the scope of *even*. Instead, the inference of *even* always projects over the domain of quantification using generic quantification.

In the following sections I will investigate how the domain of generic quantification is determined, and present my formulation for the semantics of *even*.

3.4 Determining the domain of quantification

In the previous section I showed that when the complement of *even* contains a variable which is free in the scope of *even*, the inference of *even* quantifies generically over this variable. This generic quantification, of course, must range over a particular domain. In this section I will identify two mechanisms which contribute to the determination of the domain: first, the restrictor in the lower copy for variables introduced by movement, and second, the use of cross-dimensional domain anaphora.

⁵⁵One potential concern is that (128) may be infelicitous because the (controversial) additive inference of *even* is not satisfied. The context can be modified to show that this is not the problem, for example by ensuring that everyone was required to bring wine and whiskey to the party. Beer is still relatively *likely* to be brought to parties, but *unlikely* to be brought by mitcho. The use of “someone *even* brought some [beer]_F” is still infelicitous in this modified context.

3.4.1 Domain restriction by the lower copy and wholesale late merger

I begin by considering the LF for example (131) predicted using the syntactic assumptions I laid out in section 2.2. Here I assume that the subject was generated with a restrictor in the vP -internal subject position, instead of utilizing wholesale late merger. (The wholesale late merger option will be discussed later.) After Trace Conversion, the lower copy of the subject “a professor” will be represented by the DP “the professor ①,” bound by the associated λ -binder above.

(131) **The afterparty:**

Context: The faculty have been much more social this semester. For example, think back to the last colloquium:

✓A professor *even* came to the [afterparty]_{F1}.

a. Narrow syntax: A professor *even* [_{vP} a professor came to the [afterparty]_{F1}]

b. LF: [A professor] $\lambda 1$ *even* [_{vP} [the professor ①] came to the [afterparty]_{F1}]

The LF in (131) has the property that there is a copy of the predicate “professor” associated with the variable x in the scope of *even*. As we saw in section 2.4, the Trace Converted lower copy will introduce a presupposition, in this case that the individual assigned to index 1 is a professor. I will discuss this presupposition as a definedness condition on the denotation in (132), using colon-dot notation.

(132) $[[vP]](h || g) = \lambda w_s : g(1) \text{ is a professor in } w . g(1) \text{ came to the } h(F1)(g)(w) \text{ in } w$

The complement of *even*, vP , is ordinary-assignment-dependent: specifically, it is sensitive to the individual assigned to index 1. Following the discussion in the previous section, the projected inference of *even* will quantify generically over different assignments for this index 1.

When doing so, the default behavior could be to quantify over all values in the domain of individuals D_e . However, due to the definite determiner introduced through Trace Conversion, $[[vP]](h || g)$ is only defined if $g(1)$ is a professor. I propose that in such cases, *domain restriction* occurs to narrow the domain of generic quantification to those values for which the complement vP is defined:

(133) **Domain restriction to match definedness of the nuclear scope:**

(131) \rightsquigarrow GEN $(x \in D_e)((\lambda w . \text{the professor } x \text{ came to the afterparty in } w) <_{\text{likely}} (\lambda w . \text{the professor } x \text{ came to the } <\dots\text{alternative}\dots> \text{ in } w) \dots)$

↓ Domain restriction

GEN $(x \text{ professor})(\lambda w . x \text{ came to the afterparty in } w) <_{\text{likely}} (\lambda w . x \text{ came to the } <\dots\text{alternative}\dots> \text{ in } w) \dots)$

In this way, the predicate in the lower copy of movement restricts the domain that the generic quantification of *even* operates over.

There is reason to believe that the lower copy of movement does not always restrict the domain of quantification in this way. Consider the following example, from Irene Heim (p.c.):

(134) ? Someone who happens to love that kind of work *even* cleaned the [bathroom]_F.

Many speakers—though not all—find the use of *even* in (134) felicitous given the world knowledge that, in general, people are more likely to clean other rooms than to clean the bathroom. Let's first consider a naïve Copy Theory derivation for (134). Here I assume that "someone" is decomposable into the determiner "some" and an implicit head noun "person."

(135) Naïve Copy Theory derivation for (134):

a. Narrow syntax: Some person who happens to love that kind of work *even*

↑
[_{vP} some person who happens to love that kind of work cleaned the [bathroom]_{F1}]

b. LF: [Some person who happens to love that kind of work] λ1 *even*

[_{vP} [the person who happens to love that kind of work ①] cleaned the [bathroom]_{F1}]

c. ~ predicted inference: GEN (*x* is a person who likes that kind of work)

((λ*w* . *x* cleans the bathroom in *w*) <_{likely} (λ*w* . *x* cleans <...alternative...> in *w*)...)

The predicted inference is most likely false, predicting *even* to be infelicitous in (134).⁵⁶

This puzzle can be solved by invoking late merger of the adjunct "who happens to love that kind of work" (Lebeaux, 1988; Fox, 2002). Under this derivation, the DP "some person" is generated in the *vP*-internal base position, and the relative clause is later adjoined to further restrict this domain of quantification in the subject's surface position, outside of the scope of *even*. This derivation is illustrated in (136) below.

(136) Derivation for (134) with late merger of the relative clause:

a. Narrow syntax: [Some person] *even* [_{vP} [some person] cleaned the [bathroom]_{F1}]

↓ late merger of relative clause

[Some [person [who happens to love that kind of work]]]_i

even [_{vP} [some person]_i] cleaned the [bathroom]_{F1}]

⁵⁶The interpretation of this restriction in (135c) of course depends on the correct interpretation of "that kind of work," but I believe it can be resolved as intended to mean "cleaning the bathroom."

- b. LF: [Some person who happens to love that kind of work]
 $\lambda 1 \text{ even } [_{vP} \text{ [the person } \textcircled{1}] \text{ cleaned the [bathroom]}_{F1}]$
- c. \sim predicted inference: GEN (x is a person)
 $((\lambda w . x \text{ cleans the bathroom in } w) <_{\text{likely}} (\lambda w . x \text{ cleans } <\dots \text{alternative}\dots > \text{ in } w) \dots)$

By not including the relative clause in the content of the lower copy, inside the scope of *even*, it allows us to quantify generically over “people,” leading to a scalar inference that accords with our intuition for the example and makes the use of *even* felicitous in (134).

We can even take this one step further. Because this is A-movement, following Takahashi and Hulsey (2009), it is also possible for the theta-position to have been occupied by the determiner alone, with wholesale late merger of the head noun as well as late merger of the adjunct.

(137) **Derivation for (134) with wholesale late merger of the restrictor:**

- a. Narrow syntax: [Some] *even* [_{vP} [some] cleaned the [bathroom]_{F1}]
 \uparrow
 \Downarrow Wholesale late merger of the restrictor
 [Some [person who happens to love that kind of work]]_i
 $\text{even } [_{vP} \text{ [some]}_i \text{ cleaned the [bathroom]}_{F1}]$
- b. LF: [Some person who happens to love that kind of work]
 $\lambda 1 \text{ even } [_{vP} \text{ [the } \textcircled{1}] \text{ cleaned the [bathroom]}_{F1}]$
- c. \sim predicted inference: GEN ($x \in D_e$)
 $((\lambda w . x \text{ cleans the bathroom in } w) <_{\text{likely}} (\lambda w . x \text{ cleans } <\dots \text{alternative}\dots > \text{ in } w) \dots)$

In this derivation using wholesale late merger, the domain used for generic quantification is not restricted by the restrictor of a lower copy, unlike in previous examples. In (137c) I simply give the domain of individuals D_e , but it is reasonable to suspect that the verb “clean” itself requires the agent to be animate, if not human, or that the domain will be determined through context. (More on the effect of context later.) Either way, though, we predict the derivation in (137) to yield a plausible inference, similar to (136).

The logic that we are developing is as follows: if the base position of movement, inside the scope of *even*, includes a restrictor, that restrictor must be used to restrict the domain of quantification used for *even*’s generic projection. Under Takahashi and Hulsey’s (2009) wholesale late merger proposal, this means that a DP’s head noun will necessarily be a part of the scalar inference for \bar{A} -movement but does not have to be a part of the scalar inference for A-movement. We can explicitly test this by comparing comparable A- and \bar{A} -movement configurations.

I begin with the case of A-movement, using passivization, in (138) below. Again, the portion of the context establishing relevant world knowledge for the evaluation of the scalar inference is in italics.

(138) **Karaoke party: passivization**

Context: It's the NELS karaoke party. Sabine generally does not sing; she is in fact the least likely person to sing at the party. However, at the same time, it is well known that Sabine loves the Guns N' Roses song "November Rain." She always makes Iain play it when he is at her house with a ukelele. Now it's late in the party and "November Rain" is up next. A large group of people go up to the stage, and it looks like we're in for a treat.

✓ (It looks like) "November Rain" will *even* be sung by [Sabine]_F tonight.

The use of *even* in (138) is felicitous. Let's consider the derivation of this example. Because "November Rain" has been A-moved, wholesale late merger can apply, allowing there to be no copy of "November Rain" in the complement of *even*. Here I assume that proper names are covert definite descriptions, headed by a covert definite determiner and the pronounced material is a predicate (see Geurts, 1997; Elbourne, 2002; Matushansky, 2006, and references therein).⁵⁷ I illustrate this option here:

(139) **Derivation for (138) using wholesale late merger:**

a. Narrow syntax: \emptyset_{the} will *even* [_{vP} be sung \emptyset_{the} by [Sabine]_{F1}]

↓ Wholesale late merger of the restrictor

[\emptyset_{the} "November Rain"]₁ will *even* [_{vP} be sung $\emptyset_{\text{the},1}$ by [Sabine]_{F1}]

b. LF: The "NR" $\lambda 1$ will *even* [_{vP} be sung [the ①] by [Sabine]_{F1}]

c. ~ predicted inference:

true in context

GEN ($x \in D_e$)((λw . Sabine sings x in w) <*likely*(λw . <...alternative...> sings x in w)...)

The inference derived in (139c) is one which requires that, for a generic song, it is less likely for Sabine to sing it than for others to sing it. In the context given, this inference will be true, correctly predicting the use of *even* to be felicitous in this example.

If, in contrast, wholesale late merger were not used, the scalar inference of *even* is predicted to specifically be about the song "November Rain." This derivation is illustrated below in (140).

⁵⁷This proposal will be discussed further and defended in section 5.3.

(140) Derivation for (138) *without* wholesale late merger:

- a. Narrow syntax: [\emptyset_{the} "NR"] will *even* [_{VP} be sung [\emptyset_{the} "NR"] by [Sabine]_{F1}]
- b. LF: The "NR" $\lambda 1$ will *even* [_{VP} be sung [the "NR" ①] by [Sabine]_{F1}]
- c. \sim predicted inference: *false in context*
 $\text{GEN } (x \text{ is "NR"})((\lambda w . \text{Sabine sings } x \text{ in } w) <_{\text{likely}} (\lambda w . <\dots\text{alternative}\dots> \text{sings } x \text{ in } w)\dots)$
 $\iff (\lambda w . \text{Sabine sings "NR" in } w) <_{\text{likely}} (\lambda w . <\dots\text{alt}\dots> \text{sings "NR" in } w)\dots$

The scalar inference of *even* predicted in this case is one which is specifically about the song "November Rain." Given the world knowledge that Sabine loves this song and is therefore *more* likely to sing it than other songs, not *less* likely, this inference will be false in the context.

The felicity of *even* in (138) tells us that the wholesale late merger derivation in (139) must be used. Following Takahashi and Hulsey (2009), wholesale late merger is possible in (138) because the relevant movement chain is an A-movement step. If, in contrast, an \bar{A} -movement construction is used, there will be no wholesale late merger option, resulting in an inference similar to that in (140). The use of *even* is then predicted to be infelicitous. This prediction is borne out in the following topicalization example:

(141) Karaoke party: topicalization

Context: It's the NELS karaoke party. Sabine generally does not sing; she is in fact the least likely person to sing at the party. However, at the same time, it is well known that Sabine loves the Guns N' Roses song "November Rain." She always makes Iain play it when he is at her house with a ukelele. The organizers have been disappointed by the level of participation. They've queued up "November Rain" by Guns N' Roses, which they believe will be a hit and many will want to sing.

(They're confident because,) "November Rain," *even* [Sabine]_F will sing.

In example (141), "November Rain" has been topicalized out of the scope of *even*. Note that this example uses a pre-subject *even* instead of the preverbal adverb *even* which I am generally using here. I assume, again, that *even* takes scope in its surface position, in this case over TP. The \bar{A} -movement of "November Rain" targets a position outside of the scope of *even*.

The use of *even* in (141) is judged as infelicitous. This is intuitively because the scalar inference of (141) makes reference to "November Rain," requiring Sabine to be *less* likely than others to sing the renowned power ballad, contrary to fact. The topicalization in (141) may itself be slightly marked, but by constructing a minimal pair with a song which is not contextually tied to Sabine's preferences, we see that the construction itself is fine:

(142) **The national anthem:**

Sabine generally does not sing; you might say that she is the least likely person to sing.

A new law requires that everyone sing when the national anthem and “America the Beautiful” are played. Violations have been frequently prosecuted, with stiffer sentences and fines for not singing the national anthem. Since the law was passed, most people will now sing “America the Beautiful.”

✓ (And) the national anthem, *even* [Sabine]_F will sing.

Even those speakers who feel queasy about the topicalization in (141) report a contrast with (142), where the use of *even* is felicitous in (142) but not in (141).

The derivation for example (141) and the predicted scalar inference are given in (143) below. Crucially, because topicalization is \bar{A} -movement, wholesale late merger of the domain cannot be employed as in (139) above. Therefore the lower copy of movement will include the domain “November Rain,” which acts to restrict the domain of quantification for the generic.⁵⁸

(143) **Derivation for (141):**

a. Narrow syntax: [\emptyset_{the} “NR”], *even* [_{TP} [Sabine]_{F1} will sing [\emptyset_{the} “NR”]]

b. LF: The “NR” $\lambda 1$ *even* [_{TP} [Sabine]_{F1} will sing [the “NR” ①]]

c. \sim predicted inference:

GEN (x is “NR”)((λw . Sabine will sing x in w) $<_{\text{likely}}(\lambda w$. $<\dots\text{alt}\dots>$ will sing x in w)...)

$\iff (\lambda w$. Sabine will sing “NR” in w) $<_{\text{likely}}(\lambda w$. $<\dots\text{alt}\dots>$ will sing “NR” in w)...

false in context

The discussion above motivates the view that restrictors on the lower copy of movement—if present within the scope of *even*—act to restrict the domain of *even*’s generic quantification. Furthermore, adopting the theory of wholesale late merger Takahashi and Hulseley (2009), we are able to derive a novel contrast between A- and \bar{A} -movement.

3.4.2 Cross-dimensional domain anaphora

In addition to domain restriction by lower copies, the domain of generic quantification is also affected by context. I now consider a case where the domain of quantification cannot be determined solely by syntactic material in the complement of *even*. Irene Heim (p.c.) notes the following example:

⁵⁸I do not illustrate the movement chain of the subject from Spec, v P to Spec,TP in (143). All copies of the F-marked constituent are interpreted in the scope of *even*. See section 2.3 for how the copied F-marking will be interpreted.

(144) I know someone who is 95 and still *even* [goes grocery shopping]_F.

The use of *even* in example (144) is judged as felicitous, licensed by the commonly held view that going grocery shopping may be a relatively unlikely activity for a 95-year-old. Given that going grocery shopping is not an unlikely activity for the general populous, however, we learn that the scalar inference from (144) quantifies generically over the set of 95-year-olds.

(145) **The scalar inference of (144):**

a. correct inference: GEN(*x* is 95 years old)

(($\lambda w . x$ goes grocery shopping in w) $<_{likely}(\lambda w . x$ does $<...alternative...>$ in w)...)

b. incorrect inference: GEN(*x* is a person)

(($\lambda w . x$ goes grocery shopping in w) $<_{likely}(\lambda w . x$ does $<...alternative...>$ in w)...)

Let's look at the syntactic structure of example (144). (144) involves a subject relative made up of a conjunction of two verb phrases: "is 95" and "even goes grocery shopping." The predicate "be 95 (years old)" did not originate in the scope of *even*, and therefore the complement of *even* makes no reference to 95-year-olds. The question, then, is how this scalar inference comes to be about the set of 95-year-olds.

Note, however, that the following continuation to (144) is possible:

(146) Not all of them are that healthy.

"them_{cross-sentential}" = (contextually relevant) 95-year-olds

In this case, the cross-sentential anaphor "them" is able to pick out the set of 95-year-olds. Identifying the precise mechanism by which this coreference is made possible or the conditions which govern this ability is not the goal here. Following Sudo's (2012) analysis of presupposition projection through cross-dimensional anaphora, reviewed in 3.2, I propose that such a domain anaphor is used implicitly in the interpretation of *even* in (144). (147) sketches the compositional derivation of the scalar inference of *even* in (144), based on the definition of *even* in (106). Here I illustrate the syntax of the relative clause as movement of a null operator with no contentful lower copy.⁵⁹

(147) **Computing the scalar inference of *even* in (144):**

a. Narrow syntax: ... Op_1 ... *even* [_{VP} t_1 [go grocery shopping]_{F1}]

b. LF: ... $\lambda 1$... *even* [_{VP} ① [go grocery shopping]_{F1}]

⁵⁹The example in (144) is particularly complicated, by virtue of the trigger *even* being *within* the restrictor of the quantifier whose domain is apparently being used to compute the scalar inference. For our purposes, it suffices to note that the cross-sentential anaphor is able to pick out the domain made salient by the left conjunct "is 95 (years old)" alone.

- c. $H = \{h_0, h_1, h_2, \dots\}$
- i. $h_0(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x . x \text{ goes grocery shopping in } w$
 - ii. $h_1(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x . x \text{ swims in } w$
 - iii. $h_2(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x . x \text{ plays bridge in } w$
- d. scalar inference, after generic closure:
- $$\begin{aligned} & \text{GEN } (x \in \llbracket \text{pro}_{\text{domain}} \rrbracket) (\text{scalar inference of } \llbracket \text{even } v\text{P} \rrbracket ([1 \rightarrow x] \parallel g)(w_0)) \\ \iff & \text{GEN } (x \in \llbracket \text{pro}_{\text{domain}} \rrbracket) \\ & \quad (\text{GEN } h \in H (h \neq h_0 \rightarrow \llbracket v\text{P} \rrbracket ([1 \rightarrow x] \parallel h_0 \parallel g) <_{\text{likely}} \llbracket v\text{P} \rrbracket ([1 \rightarrow x] \parallel h \parallel g))) \\ \iff & \text{GEN } (x \text{ is 95-years-old}) \\ & \quad ((\lambda w . x \text{ goes grocery shopping in } w) <_{\text{likely}} (\lambda w . x \text{ does } <\dots\text{alternative}\dots > \text{ in } w) \dots) \end{aligned}$$

To summarize, building on the previous section where I showed that the inference of *even* projects generically, in this section I investigated the domain of quantification that is used by this generic quantification. I showed that there are two ways in which the domain of generic quantification is determined: the restriction in the lower copy of movement and cross-dimensional anaphora (Sudo, 2012). For the former, I showed that the theory of wholesale late merger (Takahashi and Hulsey, 2009) accurately predicts when the head noun must be included in the content of the scalar inference. For the latter, the lack of “binding” readings with *even* suggests that the use of witness anaphora is not an option for *even*, unlike with many other presuppositions. In the next section I will present my proposal for *even*.

3.5 Proposal

The issue we have looked at in the past few sections is how the inference of *even* projects through quantification. As the discussion in section 3.2 shows, this question is a concern for all presuppositions, and in fact for all projective material. There are two logical sources in the theory for a particular presupposition trigger’s projection behavior. The first is in the lexical item of the trigger: we give each presupposition trigger a semantics that handles the quantification over different assignments. The second is in some general system of projection: each presupposition trigger has a semantics which is written assuming assignment-invariant scope.

The second option, if possible, would be the preferred option. However, I have shown here that the scalar inference of *even* differs from many other presuppositions in not allowing a “bound” reading. To make this option work, we would need an independent reason why domain anaphora, but not witness anaphora, can be used in determining the domain of quantification for *even*. I hope that future work will offer such an explanation and therefore make this second option tenable. But

for the time being, I will take the first option, and build the generic quantification and its domain specification into the lexical entry of *even*.

Here I will use the example sentence from (121) for illustration:

(148) **Computing the scalar inference of *even* in (121):**

“Someone *even* cleaned the [bathroom]_F.”

a. LF: [Some person] $\lambda 1$ [*even* [_{vP} [the person ①] cleaned the [bathroom]_{F1}]]

b. $H = \{h_0, h_1, h_2\}$

i. $h_0(F1) = \lambda g_a . \lambda w_s . \lambda x . x$ is a bathroom in w

ii. $h_1(F1) = \lambda g_a . \lambda w_s . \lambda x . x$ is a living room in w

iii. $h_2(F1) = \lambda g_a . \lambda w_s . \lambda x . x$ is a kitchen in w

c. scalar inference, after generic closure:

$\text{GEN } (x \in \llbracket \text{pro}_{\text{domain}} \rrbracket)$ (scalar inference of $\llbracket [\text{even } vP] \rrbracket ([1 \rightarrow x] \parallel g)(w_0)$)

$\iff \text{GEN } (x \in \llbracket \text{pro}_{\text{domain}} \rrbracket)$

$(\text{GEN } h \in H(h \neq h_0 \rightarrow \llbracket vP \rrbracket ([1 \rightarrow x] \parallel h_0 \parallel g) <_{\text{likely}} \llbracket vP \rrbracket ([1 \rightarrow x] \parallel h \parallel g)))$

$\iff \text{GEN } x$ (x is a (contextually-salient) person)

$((\lambda w . x$ cleaned the bathroom in $w) <_{\text{likely}} (\lambda w . x$ cleaned $\langle \dots \text{alternative} \dots \rangle$ in $w))$

Generic quantification can be thought of as akin to universal quantification, but allowing a certain degree of exceptions. In the case of example (148), we can say that the inference computed in (148c) will be true. For two out of the three people mentioned in the context—Ruth and Chris—the nuclear scope property holds: they are less likely to clean the bathroom than the kitchen or living room. Alternatively, we could interpret this inference against a wider domain of generic quantification—for example, over the entire domain of people—relying on the world knowledge that, in general, one is less likely or less willing to clean a bathroom than a kitchen or a living room. Either way, the inference computed is compatible with the fact that the use of *even* is felicitous in the context in (121).

Now I will present a general formulation for the scalar inference of *even*. For our purposes, I will represent the generic quantification of unbound variables, motivated here, as a generic quantification over different assignment functions. The set G of possible assignment functions g that this will range over will be contextually restricted, such that $g(i)$ is an individual in the domain anaphor associated with index i . A first formulation is given in (150).

(149) **The set of possible assignment functions G :**

$G = \left\{ g : \mathbb{N} \rightarrow D_e \mid \forall i . g(i) \in \llbracket \text{pro}_{\text{domain},i} \rrbracket \right\}$

(150) The semantics of *even* (updated with quantification over ordinary assignments):

$$\begin{aligned} \llbracket \textit{even} \rrbracket_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle \rangle} &\equiv \lambda g_a . \lambda w_s . \lambda P_{\langle a, \langle s, t \rangle \rangle} : \\ &\underbrace{\text{GEN } g' \in G \left(\underbrace{\text{GEN } h \in H (h \neq h_0 \rightarrow P(g' \parallel h_0) <_{\textit{likely}} P(g' \parallel h))}_{\textit{even from (106) evaluated with } g} \right)}_{\textit{scalar inference}} . \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\textit{truth condition}} \end{aligned}$$

The resulting scalar inference includes two instances of generic quantification: one over $g' \in G$, which is used to bind any free variables, and one over $h \in H$, which quantifies over different focus alternatives. I will collapse the two instances of generic quantification for my final, technical definition for *even*:

(151) Proposal for *even* (final version):

$$\begin{aligned} \llbracket \textit{even} \rrbracket_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle \rangle} &\equiv \lambda g_a . \lambda w_s . \lambda P_{\langle a, \langle s, t \rangle \rangle} : \\ &\underbrace{\text{GEN } (g' \in G, h \in H) (h \neq h_0 \rightarrow P(g' \parallel h_0) <_{\textit{likely}} P(g' \parallel h))}_{\textit{scalar inference}} . \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\textit{truth condition}} \end{aligned}$$

3.6 Summary

In this section I concentrated on establishing a detailed compositional semantics for the scalar inference of *even*. Of particular import for my proposal is the projection behavior of *even* through quantification. I showed that variables in the scope of *even* do not have a bound interpretation, unlike other presuppositions. Instead, variables in the scope of *even* are quantified over using generic quantification over a contextually determined domain. A consequence of this is that once *even* introduces a scalar inference, the content of that inference will not be affected by syntactic material outside of the scope of *even*. This feature of the semantics of *even* will be crucial for explaining the differences in association patterns between *even* and *only*.

Note that the proposal in this chapter, that the inference of *even* projects generically, is logically independent of any claim on the position where *even* takes scope at LF. Here I have assumed that the adverb *even* takes scope in its surface position. An alternative approach which involves *even* taking scope in a different, higher position at LF will be discussed in section 6.2.

Appendix: A similar scalar inference

In this chapter I gave a formulation for the scalar inference of *even* in cases where the complement of *even* includes a variable which is bound by a quantifier outside of *even*. For example, for the first example (121) above, the following was proposed as the correct inference of *even*:

(152) **An example of the proposed inference of *even*, using generic quantification:**

“Someone *even* cleaned the [bathroom]_F.”

GEN (*x* is a person)

$((\lambda w . x \text{ cleaned the bathroom in } w) <_{\text{likely}} (\lambda w . x \text{ cleaned } <\dots\text{alternative}\dots> \text{ in } w) \dots)$

The formulation I use here involves *generic quantification* (GEN) to bind the unbound variable in the scope of *even*. I assume that generic quantification is intuitively like universal quantification, but allowing for some relatively small number of exceptions (see Carlson and Pelletier, 1995, ch. 1 and references therein). The generic quantification in (152) takes scope over the $<_{\text{likely}}$ relation between the prejacent and its alternatives.

In the course of determining the correct scalar inference of *even*, we also considered another formulation which involves existential closure of the unbound variable, within each of the alternatives considered.

(153) **Another statement of the scalar inference, using existential closure:**

$(\lambda w . \exists y . y \text{ cleaned the bathroom in } w) <_{\text{likely}} (\lambda w . \exists z . z \text{ cleaned } <\dots\text{alternative}\dots> \text{ in } w) \dots$

I noted in the text that it is very difficult, if not impossible, to distinguish between a formulation in terms of wide-scope generic quantification (152) and one in terms of narrow-scope existential quantification (153). In fact, in previous versions of my work (Erlewine, 2014), I proposed this formulation involving existential closure of each alternative, within the scope of $<_{\text{likely}}$, as the formulation of the inference.

In this appendix I formally show that the apparent similarity of these two formulations is not accidental. I will show that, in general, the formulation in terms of wide-scope generic option is stronger than the narrow-scope existential option. However, the vague tolerance for exceptions built into generic quantification makes this comparison not crisp.⁶⁰

Here I will demonstrate this relationship using a case with two alternatives and one unbound variable. Let *prej* and *alt* be functions of type $\langle e, \langle s, t \rangle \rangle$ which correspond to the prejacent and alternative propositions in the scope of *even*, abstracting over the variable.⁶¹ *D* is the domain of quantification for the open variable. For example, for the example above:

⁶⁰I thank Evan Jenkins for discussion which clarified my thinking and led to the work in this section.

⁶¹The λ -abstraction of the unbound variable does not occur syntactically in the derivation of *even*. The λ -binder is used for illustration here, instead of quantifying over different choices of assignment functions as in the final definition I provide in (151).

(154) *prej, alt, D* in the Cleaning example:

- a. $prej = \lambda x . \lambda w . x$ cleaned the bathroom in w
- b. $alt = \lambda x . \lambda w . x$ cleaned the kitchen in w
- c. $D = \{x : x \text{ is a person}\}$

The following, then, are the two definitions of the scalar inference of *even*:

(155) Scalar inference of *even* using wide-scope generic quantification ($SI_{G<}$):

$$GEN(x \in D)(prej(x) <_{likely} alt(x))$$

(156) Scalar inference of *even* using narrow-scope existential quantification ($SI_{<\exists}$):

$$(\lambda w . \exists x \in D . prej(x)(w)) <_{likely} (\lambda w . \exists x \in D . alt(x)(w))$$

Here I will first consider an approximation of the generic quantification definition which uses a universal instead, which I call $SI_{V<}$. We will return to the generic quantification variant later.

(157) Scalar inference of *even* using wide-scope universal quantification ($SI_{V<}$):

$$\forall(x \in D)(prej(x) <_{likely} alt(x))$$

(158) $SI_{V<} \Rightarrow SI_{G<}$

I will formalize $<_{likely}$ using a probability function over the set of possible worlds considered (W), $P_{w \in W}$:⁶²

$$(159) \quad p_{(s,t)} <_{likely} q_{(s,t)} \iff P_{w \in W}(p) < P_{w \in W}(q)$$

Note that, for any $x \in D$, without loss of generality:

- (160) a. $prej(x) <_{likely} alt(x)$
- b. $\iff P_{w \in W}(prej(x)(w)) < P_{w \in W}(alt(x)(w))$
- c. $\iff P_{w \in W}(\neg prej(x)(w)) > P_{w \in W}(\neg alt(x)(w))$

Using (160), we can restate $SI_{V<}$ (157) as follows:

$$(161) \quad \text{for all } x \in D, P_{w \in W}(\neg prej(x)(w)) > P_{w \in W}(\neg alt(x)(w))$$

⁶²Kay (1990) cites Fauconnier (1976) with the idea to use *probability* to model the inference of the scalar, *even*-like use of French *même*. See Kay (1990) for criticisms of this approach. The argument presented here, however, can be reproduced for other measure functions returning non-negative values on a scale.

I then compute the products of the left hand side and the right hand side of the inequality in (161), multiplying out the probabilities for all values of x :

$$(162) \quad \prod_{x \in D} (P_{w \in W}(\neg \text{prej}(x)(w))) > \prod_{x \in D} (P_{w \in W}(\neg \text{alt}(x)(w)))$$

There is no complication multiplying these values and preserving the inequality, as probabilities in the range of $P_{w \in W}$ are all nonnegative, in $[0, 1]$. Therefore $\text{SI}_{\forall <}$, equivalent to (161), entails (162). Note that $\prod_{x \in D} (P_{w \in W}(\neg \text{prej}(x)(w)))$ is the probability, across all $w \in W$, that *prej* does not hold for any value of $x \in D$. Therefore (162) is equivalent to:

$$(163) \quad \begin{aligned} \text{a. } & P_{w \in W}(\nexists x \in D . \text{prej}(x)(w)) > P_{w \in W}(\nexists x \in D . \text{alt}(x)(w)) \\ \text{b. } & \iff 1 - P_{w \in W}(\exists x \in D . \text{prej}(x)(w)) > 1 - P_{w \in W}(\exists x \in D . \text{alt}(x)(w)) \\ \text{c. } & \iff P_{w \in W}(\exists x \in D . \text{prej}(x)(w)) < P_{w \in W}(\exists x \in D . \text{alt}(x)(w)) \\ \text{d. } & \iff \text{SI}_{< \exists} \end{aligned}$$

Therefore, $\text{SI}_{\forall <} \Rightarrow \text{SI}_{< \exists}$. The one step in this derivation which prevents us from deriving the other direction, $\text{SI}_{\forall <} \Leftarrow \text{SI}_{< \exists}$, happens in (162).

$$(164) \quad \text{SI}_{\forall <} \iff (161) \Rightarrow (162) \iff \text{SI}_{< \exists}$$

This is because an ordering relation between each corresponding entry of two non-negative sequences entails that the same ordering relation will hold across the products of each sequence, but not vice versa:⁶³

$$(165) \quad \text{for two sequences } \alpha_i, \beta_i \geq 0: \\ (\forall i (\alpha_i < \beta_i)) \not\Rightarrow \left(\prod \alpha_i < \prod \beta_i \right)$$

The reasoning above shows that $\text{SI}_{\forall <}$ is a stronger inference than $\text{SI}_{< \exists}$. However, my final proposal for the inference of *even* is that it projects through quantification generically, rather than strictly universally. The small number of exceptions that the universal allows for is enough to make $\text{SI}_{\forall <}$ not necessarily entail $\text{SI}_{< \exists}$. (In particular, the inequality between products in (162) does not necessarily hold, given a generic version of (161).) This illustration reflects what I have noted intuitively above, that it often seems difficult to distinguish between these two formulations.

⁶³For example, consider the sequences $\alpha = [1, 1, 0.5]$ and $\beta = [0.5, 0.5, 1]$. The product $\prod \alpha_i = 1 \times 1 \times 0.5 = 0.5$ is greater than the product $\prod \beta_i = 0.5 \times 0.5 \times 1 = 0.25$. However, it is not the case that for each corresponding entry (index i) between α and β , $\alpha_i > \beta_i$. Here $\alpha_3 = 0.5 < \beta_3 = 1$.

Chapter 4

Proposal

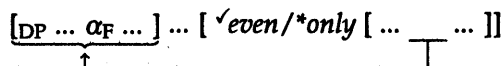
We now turn to the problem of backwards association, introduced in chapter 1. Some focus-sensitive operators—in particular, *even* but not *only*—are able to associate with a focused constituent which has moved out of the surface scope of the operator. This difference can be seen in simple minimal pairs such as the topicalization examples in (166), repeated from (4).

(166) **Backwards association with a topic:**

- a. ✓ [Mary]_F, John *even* saw ___ at the party.
- b. * [Mary]_F, John *only* saw ___ at the party.

This asymmetry was demonstrated through examples with English A and \bar{A} -movements of DPs. The pattern of the data presented thus far is schematized in (167) below.

(167) *Even* but not *only* can associate with material which has moved out:



In addition, in section 2 I presented the semantics for focus association that I adopt here, which is a version of the Roothian *Alternative Semantics* approach. Following Kratzer (1991), focused constituents are interpreted in-situ as variables with a focus-index. Focus assignment functions are used to evaluate these focused constituents within the scope of the focus-sensitive operator. For the English focus adverbs that I consider here, this is in effect a c-command requirement—a focus-sensitive operator must c-command its associate (Tancredi, 1990). Given this established approach to the semantics of focus association, the ability of *even* to associate backwards in (167) is striking.

In this chapter, I will present and motivate my proposal for this behavior. My solution involves the Copy Theory of movement, whereby the movement of an F-marked constituent leaves an F-marked “copy” in the lower, “trace” position. This is schematized in (168) below for the case where the entire restrictor of the DP is focused.

(168) Associating with F-marking in the lower copy:

- a. Narrow syntax: [DP Quant ...<...*prej*...>_F...] ... Op [... [DP Quant ...<...*prej*...>_F...] ...]
- b. LF: [DP Quantifier ...<...*prejacent*...>_F...] λx ... Op [... [DP the ...<...*prejacent*...>_F... x] ...]

In cases of backwards association with *even*, the operator *even* associates with the F-marked constituent in the *lower copy*, within the operator's scope. This is true even in cases of overt movement, where the lower copy of the movement chain is unpronounced, and therefore F-marking is not realized in that position at PF. The complement of *even* in this configuration includes an unbound variable (x above), which will be interpreted through generic quantification, using my semantics for *even* developed in chapter (3).

The same configuration with *only* will be shown to yield an unusable meaning, for one of two reasons, depending on the structure of alternatives involved. *Only* presupposes the truth of its prejacent and asserts the negation of the non-prejacent alternatives, schematized in (169) below. In one case, the variable being quantified over by the movement chain (here, x) will have incompatible requirements on it: the higher copy of DP requires that x be a ...<...*prejacent*...>..., while the lower copy of DP requires that x be a ...<...*alternative*...>... In the case that these requirements of the higher and lower copies can be satisfied simultaneously, we will then run into an issue where the utterance's truth conditions will either be a logical tautology (always true) or contradiction (always false), given the prejacent presupposition of *only*.

(169) Interpreting (168) with *Op = only*:

- [DP Quantifier ...<...*prejacent*...>_F...] λx ... *only* [α ... [DP the ...<...*prejacent*...>_F... x] ...]
- \iff [DP Quantifier ...<...*prejacent*...>...] λx ... NEG [α ... [DP the ...<...*alternative*...>... x] ...]
- \sim [α ... [DP the ...<...*prejacent*...>... x] ...] is true
- (for different values of x , depending on the projection of the presupposition)

In order to illustrate the proposal clearly, in this chapter I will be constraining the range of data considered, as well as our theoretical assumptions. I will limit attention to movement of DPs with F-marking in their restrictors. Association with the determiner and entire DP will be discussed later in chapter 5. I will also not consider the possibility of reconstruction feeding focus association, but this possibility will be discussed in detail in chapter 7.

I begin in section 4.1 by showing that the F-marked constituent originating within the scope of *even* is necessary for backwards association, thereby motivating the view that backwards association is made possible by the existence of the lower copy of the DP within the operator's scope. Evidence from backwards association will support the Takahashi and Hulsey (2009) wholesale late merger proposal regarding the difference between A- and \bar{A} -movement. This is followed by

a brief note regarding the possibility of syntactic reconstruction, in section 4.2. I will then explain the inability of *only* to associate backwards in section 4.3 and the ability of *even* to do so in section 4.4. In section 4.6, I then show that this behavior and proposal extends beyond overt movement to a type of covert movement, Quantifier Raising, supporting the view that quantifier scope shift involves a covert movement step which leaves copies, as overt movement does.

4.1 The importance of the lower copy

We can see that the lower, base position of movement is at least relevant for backwards association with *even* by varying the corresponding gap to be inside or outside of the surface scope of *even*. The sentences in (171) below are all based on the grammatical sentence in (170). Two variables are crossed: placement of *even* in the higher clause or the lower clause, and topicalization of an F-marked “the judges” from the higher clause or an F-marked “the Canadians” from the lower clause. Example (171d) is the only condition where the gap position of the topic is not within the surface scope of *even*, and it is ungrammatical with the intended association.

(170) Baseline: The report convinced the judges that we spied on the Canadians.

(171) *Even* associating with leftward topic requires base position in *even*'s scope:

- a. ✓ The [Canadians]_F, the report *even* convinced the judges that we spied on ___.
- b. ✓ The [Canadians]_F, the report convinced the judges that we *even* spied on ___.
- c. ✓ The [judges]_F, the report *even* convinced ___ that we spied on the Canadians.
- d. * The [judges]_F, the report convinced ___ that we *even* spied on the Canadians.

(Ungrammatical with the intended association of *judges* with *even*.)

A theory which attempts to explain the contrast in (166) by saying that *even*, but not *only*, is generally able to associate with an F-marked constituent which is outside of its scope, would incorrectly predict (171d) to be grammatical. An alternative generalization would be to require that the *linear* position of *even* precede the gap. However, such a generalization is counterexemplified by examples such as (172)

(172) *Even* linearly preceding the gap is not sufficient:

- * The [judges]_F, [for the report to *even* impress me] would annoy ___. (David Pesetsky, p.c.)

Instead, these contrasts show that backwards association requires that the lower, “gap” position of the movement chain be in a position which *even* can normally associate with, within the scope of *even*.

We observe the same pattern with A-movement as well. Here I will use the subject raising predicate *seem*. In the examples in (173) below, I assume the subject “a professor” originated in the embedded nonfinite clause. In (173a), *even* adjoined to the embedded verb phrase is able to associate with the leftward predicate “professor.” In (173b), *only* in the same position is not able to associate in this pattern.

(173) *Even* can associate with a leftward derived subject, but not *only*:

- a. ✓ A [professor]_F seems to *even* be at the party.
- b. * A [professor]_F seems to *only* be at the party.

Here too we can verify that the movement chain, originating within the scope of *even*, is crucial for this pattern of association. The grammatical raising example in (173a) is compared to a superficially similar control configuration below in (174). The *even* adjoined to the nonfinite embedding is unable to associate with the leftward subject in (174b).

(174) **Backwards association across raising vs control:**

- a. ✓ A [professor]_F seems to *even* [___ be at the party]. raising (=173a)
- b. * A [professor]_F wants to *even* [PRO be at the party]. control

I assume that control verbs such as “want” take a complement with a null PRO subject.⁶⁴ Therefore in (174b), “a professor” is base-generated in the matrix subject position as the embedding verb is a control verb, “want.” The ungrammaticality of (174b) then correlates with the F-marked constituent not originating within the scope of *even*. In contrast, in (174a) the DP containing the F-marked “professor” has raised out of the nonfinite embedding, where it originally was below the surface position of *even*, allowing for the intended focus association. This contrast in (174) shows us that *even* is not simply able to associate with material which is outside of its scope; instead, the associate must have originated within the scope of *even*.

The contrast between *even* and *only* in associating with material which has A-moved out is able to explain a long-standing puzzle in the literature on focus association: the ability of *even* but not *only* to associate with a leftward subject. This difference was first noted by Jackendoff (1972), whose original examples are given below in (175). An auxiliary is added in (176a) to show that the *even* in question is indeed a VP-*even* and not simply a post-nominal constituent *even*.⁶⁵

⁶⁴Alternatively, we could say that “control” also involves movement (Hornstein, 1999, a.o.) but is not allowed to leave a lower copy of the restriction. See discussion below on wholesale late merger, which shows that what backwards association actually requires is for the F-marked restrictor to be generated in the scope of *even*.

⁶⁵Example (175b) is given a * in Jackendoff (1972), but given a ? in Rooth (1985, p. 95). I believe Rooth’s (1985) judgment reflects the fact that (175b) could be interpreted as grammatical by speakers who freely permit a postpositional, nominal-

(175) VP-*even* but not VP-*only* can associate with a leftward subject:

- a. ✓ [John]_F *even* gave his daughter a new bicycle. (Jackendoff, 1972, p. 248)
 b. * [John]_F *only* gave his daughter a new bicycle. (*ibid*, p. 250)

(176) a. ✓ [John]_F will *even* give his daughter a new bicycle.

- b. * [John]_F will *only* give his daughter a new bicycle.

Under the VP-internal subject hypothesis,⁶⁶ the subject in (175a,176a) has originated within the scope of *even*. The derivation for (176) is schematized below in (177). The contrast between *even* and *only* here can then be reduced to the more general difference between *even* and *only* in associating with material which has moved out of their scope, schematized in (167) above.

(177) [John]_F will ✓ *even* / **only* [_{vP} _____] give his daughter a new bicycle]

The evidence presented in this section motivates the view that the ability of *even* to associate backwards is dependent upon the focused material originating within the scope of *even*. The existence of a movement chain somehow allows for this exceptional pattern of association.

We can go one step further to show that what *even* requires in such cases is an instance of the F-marked predicate in the scope of *even*. Under the Copy Theory of movement, this is what a movement chain originating within the scope of *even* provides. Evidence for this view comes from the theory of A-movement based on *wholesale late merger*, developed in Takahashi and Hulsey (2009).

Recall from section 2.2 that A- and \bar{A} -movements behave differently with regard to Binding Condition C. In particular, the trace positions of A-movement chains are apparently never active for

adjoined *only*. It is therefore important to test such examples with an intervening auxiliary or adverb, as in (176b), linearly separating the subject from *only*. (See Beaver and Clark (2008, p. 161 fn. 2) for a similar intuition.) The speakers I have consulted with have judged (176b) as robustly ungrammatical.

⁶⁶The VP-internal subject hypothesis is often credited to work of the mid-80's, such as Kitagawa (1986) and Kuroda (1988). See McCloskey (1997) for additional references and a summary of arguments from these works. David Pesetsky (p.c.) notes that the idea was also proposed earlier in Hale (1978).

McCawley (1970) may also be an early predecessor for the hypothesis. Interestingly, one argument offered by McCawley in support of the VP-internal subject hypothesis comes (not called that by him, of course) comes from focus association. McCawley notes that *only* can be in pre-verbal position and associate with the propositional content of the whole clause, which also includes the subject which is pronounced to the left.

- (i) The judge *only* sent you to prison; your wife didn't leave you too.

McCawley proposes that the subject originates in a VSO clause to the right of *only* and is then fronted through a process of "V-NP inversion." In modern terms, this example could be thought of as *only* associating with its complement *vP*, with the subject moved out of it. Here I will not consider such configurations where the moved constituent is a proper subpart of what is F-marked, and will leave the study of such configurations for future research.

Condition C, as illustrated through the examples in (178). As discussed in section 2.2, this contrasts from the behavior of \bar{A} -movements, where the moved DP (with the exception of adjoined material: Lebeaux 1988) will act as if it is present in the base position of movement for the evaluation of Condition C.

- (178) **A-movement bleeds Condition C:** (=48)
- a. [The claim that John_i was asleep] seems to him_i to be correct. (Chomsky, 1993, p. 37)
 - b. [John_i's mother] seems to him_i to be wonderful. (Lebeaux, 1998, p. 23)
 - c. [Every argument that John_i is a genius] seems to him_i to be flawless. (Fox, 1999, p. 192)

Takahashi and Hulseley (2009) propose that A-movement chains can take advantage of an operation of *wholesale late merger*. A determiner without a restrictor will first be base-generated and then A-moved, and the restrictor can be wholesale-late-merged into the DP in its higher position. This is illustrated for example (178c) in (179) below. Because "John" was late-merged in with the restrictor in a position above "him," we do not incur a Condition C violation.

- (179) **Copy Theory with wholesale late merger:**⁶⁷ (=51, Takahashi and Hulseley, 2009, p. 400)
- [Every] seems to him_i to be [[every] flawless]
- ↑
- ↓ wholesale late merger
- ✓ [Every [argument that John_i is a genius]]₁ seems to him_i to be [[every]₁ flawless]

Takahashi and Hulseley (2009) further argue that this option (base-generating a determiner without a restrictor and then wholesale-late-merging a restrictor above) is possible for A-movement but not \bar{A} -movement because a DP must have its restrictor when it receives Case. Because an A-moving DP receives Case in the higher position, its base position need not have a restrictor. In \bar{A} -chains, in contrast, the base position is a Case position, which must have a head noun. Crucially, the use of wholesale late merger in A-chains is an *option*: A-movement can have a restrictor in its base position or not, whereas \bar{A} -movement must have a restrictor in the base position.

The proposed theory of A-movement involving wholesale late merger can be used to show that backwards association with *even* requires an instance of the F-marked predicate in *even*'s scope at LF. This argument will be based on the examples in (178) above, involving the raising predicate "seem" with an intervening experiencer. I first show that, in such configurations, backwards association with *even* in the nonfinite clause is possible:

- (180) **Baseline: backwards association is possible with raising across an experiencer**
- Context: Recently on TV, many commentators have been presenting arguments for why their favorite politician is a genius. Among these arguments, the arguments that Sarah Palin is a genius have been the most outlandish and consistently refuted.

John believes everything he hears; specifically, he believes all the arguments that claim that some politician is a genius.

✓ [Every argument that [Palin]_F is a genius] seems to John to *even* be valid.

Now I will make the experiencer coreferential with the subject “Sarah Palin” in the subject’s embedded clause. Recall that this configuration, without backwards focus association, did not trigger a Condition C violation (178). However, with backwards association with *even*, we yield a Condition C effect: the sentence in (181) cannot be interpreted with the experiencer coindexed with “Sarah Palin”:

(181) **Backwards association makes Condition C effects reemerge with A-movement:**

* [Every argument that [Sarah Palin]_{F,i} is a genius] seems to her_i to *even* be valid.


(ungrammatical with intended association and coreference)

In the theory of Takahashi and Hulsey (2009), A-movement can normally invoke wholesale late merger, leading to a configuration which avoids the Condition C violation—see (51c) above. However, when *even* associates backwards with a part of the A-moved restrictor, a Condition C violation occurs, indicating that the R-expression “Sarah Palin” is now in a position that is commanded by the experiencer.⁶⁸ In other words, backwards association has forced the A-moved DP to be derived *without* wholesale late merger (182b). This motivates the view that backwards association requires (a copy of) the F-marked predicate in its scope. If wholesale late merger were used in (180/181), the scope of *even* would not contain a copy of the F-marked predicate, although we would then avoid the Condition C violation (182c).⁶⁹

(182) **Different derivations for (180/181):**

a. Without the Copy Theory:

[Every argument that [Palin]_{F,i} is a genius] seems to [John/her_i] to *even* be [t valid]

⇒ predicts no Condition C effect in (181); unclear how *even* associates with focus 

⁶⁸It is well known that the DP in certain PP arguments can behave for binding purposes as if it c-commands out of the PP. The following example (i) is a baseline for this configuration triggering Condition C. This issue is orthogonal to the argument here. See Bruening (2014) for recent discussion.

(i) * It seems to her_i that the argument that [Sarah Palin]_i is a genius is valid.

⁶⁹Note that in cases where the DP scope reconstructs into its A-chain, the restrictor of an A-moved DP must be generated low, without wholesale late merger (Takahashi and Hulsey, 2009, based on Fox 1999). One way to view the Condition C data presented here is to claim that backwards association with *even* triggers obligatory scope reconstruction. This, however, is not true. See section 4.2 and chapter 7 for discussion.

b. With Copy Theory, without wholesale late merger:

[Every argument that [Palin]_{F,i} is a genius] seems to {John/her_i} to *even* be
 ↑
 [[every argument that [Palin]_{F,i} is a genius] valid]

⇒ predicts Condition C effect in (181); the scope of *even* contains F-marking ☺

c. With Copy Theory and wholesale late merger:

[Every] seems to {John/her_i} to *even* be [[every] valid]

↓ wholesale late merger

[Every [arg. that [Palin]_{F,i} is a genius]]₁ seems to {J/her_i} to *even* be [[every]₁ valid]

⇒ predicts no Condition C effect in (181); unclear how *even* associates with focus ☹

In some sense, what we have done is take an A-movement chain (raising) and make it behave like an \bar{A} -movement chain. A-movement normally bleeds Condition C, but backwards association forces the restrictor to be generated in the base position of movement, and thus be visible for Condition C as in \bar{A} -movement chains. We can similarly use backwards association to trigger weak crossover, a constraint traditionally described as a characteristic of \bar{A} -movement.

(183) **Weak crossover:**a. A-movement:

(Bresnan, 1994, ex. 21)

✓ { John_i
 Everyone_i } was visited by his_i mother on the night of the murder.

b. \bar{A} -movement:

(Bresnan, 1994, ex. 22a)

* [Which suspect]_i did his_i mother visit on the night of the murder?

I will again use raising over an experiencer as my example of A-movement. Consider the following example:

(184) **Backwards association can make an A-chain trigger weak crossover:**

Context: We know that mothers have a very high opinion of their children. We also know that girls are generally better behaved than boys, so it may be unsurprising that [every girl]_i seems to her_i mother to be wonderful. However, according to the results of this surprising new study:

* [Every [boy]_F]_i seems to his_i mother to *even* be wonderful.

(ungrammatical with intended association and coreference)

The fact that the sentence “[every girl]_i seems to her_i mother to be wonderful” is possible in the context shows that variable binding is normally possible between the derived subject position

of the raising predicate and the possessor “her” in the experiencer. No Binding Conditions are violated in this A-movement structure. The test sentence, however, involves backwards association of *even* at the nonfinite clause edge with “boy” in the moved DP. The test sentence yields a weak crossover effect, even though the raising movement chain itself has not changed in terms of its trigger or where the DP receives Case. Therefore, the weak crossover effect must be attributed to the presence of *even* in the sentence. We have again used backwards association to give an A-chain properties characteristic of an \bar{A} -chain.

The contrast here furthermore teaches us something about the nature of weak crossover itself. The fact that backwards focus association triggers a weak crossover effect in (184) points towards an analysis of weak crossover that is *sensitive to the presence of the restrictor* in the base position of movement. That is, the following could be a characterization of weak crossover effects:

- (185) If a DP moves from position A to position B and the DP *has a restrictor in position A*, the DP cannot bind variables that are outside of the scope of A.

There are many proposals in the literature regarding the source of weak crossover effects. Here I will note that Sauerland (1998) and Ruys (2000) both suggest that the content of the base position of movement matters for weak crossover, making particularly promising the semantic approach to weak crossover that they each independently proposed. If backwards association with *even* requires there to be an instance of the F-marked constituent in the scope of *even* at LF, the movement chain will necessarily have its restriction merged in the base position of the chain, triggering Sauerland/Ruys’ weak crossover configuration.⁷⁰

In this section I took a close look at patterns of backwards association with *even* and argued that backwards association requires the focus associate to have originated within the scope of *even*. Furthermore, following Takahashi and Hulsey’s (2009) approach to the difference between A- and \bar{A} -movement, I showed that backwards association requires the *restrictor* of the moved DP—which

⁷⁰This is most explicit in the discussion in Sauerland (1998, §5.2).

Briefly, Sauerland (1998) and Ruys (2000) independently propose that when a DP includes its restrictor in its base position—as in \bar{A} -movement, according to Takahashi and Hulsey (2009)—the restrictor is interpreted only in the base position of movement, and the quantificational material—interpreted in the higher position—quantifies over different *choice functions* over the domain of the DP. Because choice functions are not possible antecedents for e-type variables, this predicts the weak crossover effect for \bar{A} -movement. If movement does not include the restrictor in the lower copy—as in A-movement, according to Takahashi and Hulsey (2009)—the movement quantifies over an e-type variable, which can also bind other variables.

The interpretation of movement chains in Sauerland and Ruys’ work differs non-trivially from the Copy Theory/Trace Conversion approach I explore here (§2.2 above). I will leave for future work fully reconciling this approach to weak crossover with my approach to backwards association.

I thank Juliet Stanton for bringing this to my attention and for discussion of these issues.

includes the F-marking—to be present in the scope of *even*. Before turning to my proposal for why backwards association is possible with *even* but not with *only*, I will now briefly discuss one possible analysis utilizing syntactic reconstruction.

4.2 A short note on syntactic reconstruction

Having established that the focus associate originating within the scope of *even* is crucial for this pattern of association, we might imagine that backwards association involves syntactic reconstruction of the moved constituent. That is, even though the F-marked constituent is not in the scope of the focus operator at PF, the moved DP is interpreted under reconstruction within the scope of the operator at LF.⁷¹ This possibility is schematized here:

(186) **One possible approach: syntactic reconstruction**

- a. NS/PF: [DP ... α_F ...] ... [Op [... ...]]
- b. LF: ... [Op [... [DP ... α_F ...] ...]]

Reconstruction has been proposed as a solution to certain non-surface scope readings. For example, the sentence in (187) below has an inverse scope reading, where negation takes scope over the subject universal quantifier. As argued by McCloskey (1997) and Sauerland (2003), I assume that this inverse scope reading is due to the subject originating in a lower, *vP*-internal position and then A-moving to its surface position, together with the ability of this movement chain to “reconstruct” into this base position.⁷²

(187) **Subjects can scope reconstruct into their *vP*-internal position:**

Every professor didn't come to the party.

$\checkmark \forall > \neg, \checkmark \neg > \forall$

Assumed LF for inverse scope: NEG [_{*vP*} every professor come to the party]

There are roughly two approaches to scope reconstruction, known in the literature as “syntactic” and “semantic” reconstruction. Here I will discuss syntactic reconstruction.⁷³ Under this view, each interface can independently determine which copy within a movement chain will be used—at LF for interpretation and at PF for pronunciation. Additional copies are ignored (at PF, and at LF for higher copies) or interpreted as bound variables (lower copies at LF). Such a syntactic

⁷¹Such an approach is suggested as a possibility in Kayne (1998, fn. 75).

⁷²See Johnson and Tomioka (1998) for a slightly different view.

⁷³For our current purposes, “PF movement” approaches which derive “reconstruction” behavior by moving constituents on the PF branch after narrow syntax (Sauerland, 1998; Sauerland and Elbourne, 2002), are the same as syntactic reconstruction.

approach to \bar{A} -reconstruction was suggested as part of the original proposal of the Copy Theory by Chomsky (1993), and has been extended to the reconstruction of A-chains (Hornstein, 1995). This Copy Theory-based syntactic reconstruction approach is demonstrated in (188) for the inverse scope reading of (187):

(188) **Syntactic reconstruction for (187):**

- a. Narrow syntax: [Every professor] didn't [every professor] come...]
- b. Pronounce higher copy at PF: "Every professor didn't come to the party."
- c. Ignore higher copy at LF: ~~[every professor]~~ NEG [every professor come...] $\rightarrow > \forall$

In the same way, we could imagine that when a focus-sensitive operator such as *even* apparently associates backwards, the moved constituent containing the F-marking syntactically reconstructs into the scope of the operator. This possibility is illustrated for an example of backwards association with *even*, in (189):

(189) **Syntactic reconstruction feeding backwards association with *even*:**

- ✓ Every [professor]_F will *even* come to the party.
- a. Narrow syntax: [Every [professor]_F] FUT *even* [every [professor]_F] come...]
 - b. Pronounce higher copy at PF: "Every PROFESSOR will *even* come to the party."
 - c. Ignore higher copy at LF: ~~[every [professor]_F]~~ FUT *even* [every [professor]_F come...] \Rightarrow *even* can then associate with "professor" possible

There are two problems with relying on syntactic reconstruction to explain backwards association. The first problem has to do with the scope of the moved DP that contains the F-marking. Using syntactic reconstruction as in (189) predicts that in cases of backwards association with *even*, the DP containing F-marking will take scope in its lower position within the scope of *even*. This prediction is not borne out. Instead, the choice of DP scope is independent of the need to associate with *even*. Consider the sentence in (190), where *even* associates backwards with the restrictor "student" of the subject. The sentence includes a sentential negation so that scope reconstruction of the universal subject will be clearly detectable as a difference in truth condition. As the contexts below show, example (190) is compatible with both surface scope and inverse scope between the universal subject and negation.

(190) **Subject association with *even* is compatible with different scopes for the subject:**⁷⁴Every [student]_F didn't *even* come to the party. $\checkmark \forall > \neg, \checkmark \neg > \forall$ a. $\checkmark \forall > \neg$: \Rightarrow No student came.


We planned a party with free food. We expected at least some students to come, because they love free food, but we didn't particularly expect professors to come. No one attended the party. We weren't surprised that every professor didn't come ($\forall > \neg$). What surprised us, though, is that every [student]_F didn't *even* come to the party.

b. $\checkmark \neg > \forall$: \Rightarrow Not every student came, but some may have.

We planned a party with free food. We expected all the students to come, because they love free food, but we didn't expect all the professors to come. Not all of the students came and not all of the professors came. We weren't surprised that every professor didn't come ($\neg > \forall$). What surprised us, though, is that every [student]_F didn't *even* come to the party.

The two scopes in (190) show that the possibility of *even* associating with "student" in the subject is independent of the scope of the quantificational subject. If backwards association required that the DP containing F-marking reconstruct, as in (189), we predict only the inverse scope reading to be possible. I will return to the two readings of this example in section 7.1 and present their derivations and semantic computations there.

The second problem has to do with the difference between *even* and *only* in allowing backwards association or not. The derivation presented in (189) which used syntactic reconstruction to feed association with *even* could also be used to feed association with *only*. The availability of this option would predict the backwards association with *only* in (191) to be grammatical, contrary to fact.⁷⁵

(191) **Syntactic reconstruction predicts backwards association with *only* to be possible:*** Every [student]_F will *only* come to the party.Intended: \approx *Only* every [student]_F will come to the party.a. Narrow syntax: $\underbrace{[\text{Every } [\text{student}]_F] \text{ FUT } \text{only } [\text{every } [\text{student}]_F] \text{ come...}]}$ b. Pronounce higher copy at PF: "Every STUDENT will *only* come to the party."c. Ignore higher copy at LF: $\{ \text{every } [\text{student}]_F \}$ FUT *only* [every [student]_F came...] \Rightarrow predicts association to be possible, contrary to fact 

⁷⁴The inference of *even* in these sentences is complicated by the fact that there is a downward-entailing operator, negation. The effect of *even* in downward-entailing environments is discussed in chapter 6.

The possibility of syntactic reconstruction predicts backwards association to be generally possible, regardless of the focus-sensitive operator involved. We are not able to explain the contrast between *even* and *only* unless this reconstruction step were independently blocked across *only*. As the following examples show, it is not the case that subjects cannot reconstruct within the scope of *only*. The contrast between *only* and *even* would be unexplained by this reconstruction-based approach.

(192) **A-moved subjects can generally reconstruct within the scope of *only*:**

- a. In order to win the Valentine's Day competition, we need someone to kiss every student in the class ($\forall > \exists$). Unfortunately, someone *only* kissed [John]_F. \checkmark *only* > \exists
- b. A Jamaican is *only* likely to win the [Bobsled]_F competition. \checkmark *only* > likely > (de dicto) \exists
- c. A general *only* died in [this]_F war. \checkmark *only* > (de dicto) \exists

We are left, then, with a puzzle. Moved DPs in English are able to scope-reconstruct, but this reconstruction cannot result from ignoring the higher copy, which would predict the availability of backwards association with *only*, as well as with *even*. For the rest of the dissertation until chapter 7, I will therefore assume that syntactic reconstruction is not possible in English as a way of feeding focus association. In chapter 7 I will then return to discuss reconstruction in detail, ultimately concluding that syntactic reconstruction *does* exist in English, but it does not work by simply ignoring the higher copy at LF, as commonly assumed.

⁷⁵One difference between (187) and (191) is the presence of the negation in (187). One could imagine, following Fox's (1995) Scope Economy, that reconstruction of the subject in (187) is licensed because doing so introduces a new scope relation ($\neg > \forall$), but perhaps the reconstruction of the subject is not licensed in (191) in the same way. However, this cannot be the case as reconstruction of the subject *does* introduce a truth-conditional difference, independent of the focus association it may or may not feed. This is observed through the ambiguity of the following example, where *only* associates with a constituent in its surface scope.

(i) **Scope reconstruction of the subject would make a truth-conditional difference:**

Every student has *only* read the [syllabus]_F.

- a. $\forall > \textit{only}$: Every student is such that they only read the syllabus. (None of them read anything else.)
- b. $\textit{only} > \forall$: The syllabus is the only thing that every student read. (But some of them may have read other things, too.)

Also note that adding sentential negation, which *every* is able to reconstruct under, does not allow for backwards association:

(ii) * Every [student]_F will *only* not come to the party.

Intended reading ($\textit{only} > \neg > \forall$): Only students, not professors, are such that λP . not all P come to the party.

4.3 Only

In this section I present my proposal for why *only* cannot associate with material which has moved out of its scope. The core idea of this proposal is the following. Under the framework I have laid out in section 2, when a constituent containing F-marking is moved via the Copy Theory, both copies retain their F-marking. I propose that when one copy of a chain is within the scope of *only* but another is not, this results in contradictory requirements on the variable being quantified over, resulting in an unusable meaning.

I will illustrate this using example (193), repeated from (191) above. As with the previous case with *even*, copying the subject will yield two instances of the F-marked predicate “student,” with one being in the scope of *only* (193a). Following Trace Conversion, we will have the LF in (193b). Here I will assume that the definite determiner introduced in the process of Trace Conversion is a presuppositional definite, as discussed above in section 3.4. The denotation for the lower copy of movement will then carry a presupposition based on its domain (193c).⁷⁶ This leads to the denotation for *vP* in (193d).

(193) An example of attempted backwards association with *only* (=191):

* Every [student]_F will *only* come to the party.

- a. Narrow syntax: [Every [student]_{F1}] FUT *only* [every [student]_{F1}] come to the party
- b. LF after TC: [Every [student]_{F1}] λ1 . FUT *only* [_{vP} [the [student]_{F1} ①] come to the party]
- c. [[the [student]_{F1} ①]] (h || g) = λw_s : g(1) is a h(F1)(g)(w) in w . g(1)
presupposition
- d. [_{vP}] (h || g) = λw_s : g(1) is a h(F1)(g)(w) in w . g(1) come to the party in w
presupposition

Notice that this denotation for *vP* has a particular shape: it has a presupposition introduced by the lower copy of movement, and the content of this presupposition is focus-assignment-sensitive; but the truth conditions of *vP* are not. This general shape is characteristic of backwards association configurations. In backwards association configurations, the F-marking in the scope of the operator is in the restrictor of the moved DP, and the restrictors of lower copies of movement contribute only a presupposition (section 2.4). Therefore, the denotations of the complement of focus-sensitive operators in backwards association configurations will only be focus-assignment-dependent in their presupposition.

⁷⁶This presupposition comes from the existence presupposition of the definite determiner: there must be an individual in the restrictor which is also the variable $g(1)$; therefore the variable $g(1)$ must be in the restrictor. See section 2.4 for detailed discussion of the content of this presupposition.

Suppose here that the salient alternative to “student” in the context is “professor”:

- (194) $H = \{h_0, h_1\}$
- a. $h_0(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a student in } w$
 - b. $h_1(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a professor in } w$

We now have the necessary ingredients to interpret the semantics of [*only vP*]. I will begin by computing the truth-conditional part of the meaning of *only*, using my definition from chapter 2. I will discuss the contribution of *only*'s prejacent presupposition below.

- (195) **Semantics for *only* without prejacent presupposition:**
- $$\llbracket \text{only} \rrbracket_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle, t \rangle \rangle \rangle} \equiv \lambda g_a . \lambda w_s . \lambda P_{\langle a, \langle s, t \rangle} . \forall h \in H (h \neq h_0 \rightarrow P(h \parallel g)(w) \text{ is false})$$

Composing *only* using Superintensional Functional Application (44), we yield the denotation for [*only vP*] in (196). The denotation of *vP*, computed above as (193d), includes the contribution of the Trace Converted lower copy, “the [professor]_{F1} ①”: a presupposition that, given an ordinary assignment function g and a focus assignment function h , $g(1)$ must satisfy the predicate $h(\text{F1})$. The entire denotation for [*only vP*] will presuppose that the bound variable, $g(1)$, satisfies each of the non-prejacent alternatives to “student”; here, this means that a presupposition that $g(1)$ is a *professor* is introduced.

- (196) $\llbracket \text{only vP} \rrbracket = \lambda g_a . \lambda w_s . \forall h \in H (h \neq h_0 \rightarrow \llbracket vP \rrbracket (h \parallel g)(w) \text{ is false})$
 $= \lambda g_a . \lambda w_s . \llbracket vP \rrbracket (h_1 \parallel g)(w) \text{ is false}$
 $= \lambda g_a . \lambda w_s : g(1) \text{ is a } h_1(\text{F1})(g)(w) \text{ in } w . g(1) \text{ does not come to the party in } w$
 $= \lambda g_a . \lambda w_s : g(1) \text{ is a professor in } w . g(1) \text{ does not come to the party in } w$

We then λ -abstract over the index 1, to derive the nuclear scope predicate for the moved quantifier “every student” in subject position:

- (197) $\llbracket \lambda 1 . \text{only vP} \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e : ([1 \mapsto x] \parallel g)(1) \text{ is a professor in } w .$
 $([1 \mapsto x] \parallel g)(1) \text{ does not come to the party in } w$
 $= \lambda g_a . \lambda w_s . \lambda x_e : x \text{ is a professor in } w . x \text{ does not come to the party in } w$

The question now is how this presupposition of (197)—that the individual argument of the nuclear scope predicate must be a *professor*—projects through the quantifier that binds it, “every student.” This question of how the existential presupposition of a Trace Converted lower copy of movement was covered in section 2.4. There I reviewed a set of examples involving *re-binding* described in Merchant (1999, 2001) which showed us that such presuppositions project *universally*

over the higher domain of quantification. In other words, in this example, we predict that the presupposition that x is a professor (in 197) holds for each x in the domain of quantification—that is, for every student:

(198) **The presupposition of (193), due to the Trace Converted lower copy of movement:**

$\sim \forall x (x \text{ is a student in } w_0 \rightarrow x \text{ is a professor in } w_0)$

The presupposition of the Trace Converted lower copy projects over the domain of quantification, “student.” Let’s imagine for current purposes that not all students are professors. The presupposition of (193) is not satisfied, leading to the unacceptability of the sentence.

Note furthermore that the universal projection of this presupposition is a general property of this type of presupposition, rather than a property of the quantifier involved in (193) being “every.” As discussed in chapter 3, presuppositions are known to project in different ways through higher quantification. In particular, existential quantifiers have been shown to generally allow “bound” projection of a presupposition, so that the presupposition is required to hold only of the individuals that verify the use of the existential quantifier (see section 3.2). However, the evidence presented in section 2.4 showed that this presupposition—the existence presupposition of a Trace Converted lower copy of movement—projects universally even through existential quantifiers. Thus for an example such as (199) below, we still predict the presupposition of being a “professor” to project over the entire domain, “student.”

(199) **The same presupposition predicted for an example with an existential quantifier:**

* A [student]_F will *only* come to the party.

$\sim \forall x (x \text{ is a student in } w_0 \rightarrow x \text{ is a professor in } w_0)$

This explanation for the unacceptability of (193) immediately raises two questions. The first regards the nature of the *badness* that speakers sense from this sentence. In many cases when someone says an utterance with a presupposition not shared in the common ground, the presupposition is accommodated by others or consciously rejected. Speakers do not react to sentences such as (193) in this way, for example by questioning whether it’s true that all students are professors, but instead judge the structure as ungrammatical with the intended reading.⁷⁷ The second question has to do with the content of the hypothesized presupposition. In the case of (198), the

⁷⁷ All examples of backwards association with *only* do have a grammatical parse, if we allow the speaker to reanalyze the position of *F*-marking with a corresponding change in interpretation. In my experience, this can occur when these target sentences are spoken as well; many speakers will reply that the sentence is grammatical, but when asked for a paraphrase of the truth conditions, it becomes clear that they have parsed the sentence with *only* associating with material in its scope, rather than the intended backwards associate. For this reason, it is difficult to say that these surface structures are ungrammatical *per se*; the relevant notion is *grammaticality with the intended association*.

presupposition may indeed not hold given standard world knowledge. But it seems possible to construct examples where the presupposition projected in this way *can* be satisfied. Does backwards association with *only* become grammatical in such cases?

I will first answer the second question. As we will see in a moment, cases where this hypothesized presupposition might hold do *not* allow for backwards association with *only*. When we think about the meaning of such cases, taking the contribution of the preadjacent presupposition of *only* into account, we see that its truth conditions will always be a logical tautology (always true) or contradiction (always false). Later in this section I will explain why backwards association with *only* always yields such a degenerate meaning.

This state of affairs, that the truth conditions of a sentence with backwards association with *only* always yield a degenerate meaning, provides an answer to the question regarding the grammaticality status of such examples. It is not simply the case that an example like (193) carries a presupposition that must hold or be accommodated; if the presupposition *does* hold and the truth conditions can be computed, it turns out that the truth conditions are not contingent. Utterances which are inherently tautological or contradictory by virtue of their logical structure are judged by speakers as ungrammatical (Gajewski, 2002), explaining the judgment of backwards association with *only* as ungrammatical. The presuppositions may hold or not, but the configurations are damned if they do and damned if they don't.

Let's begin now by looking at a concrete example where the presupposition projected by the lower copy of movement can be satisfied:

(200) **New degree requirements:**

Context: David wants to change the requirements for the syntax track of the Linguistics degree program. The department rules say that such changes must be passed by a unanimous vote of all students and faculty that it will affect. Initially people thought that meant that every syntactician must vote for the change for it to pass, but the Rules Committee determined that every linguist must vote for it.⁷⁸

David is worried that the proposal will not pass, because...

* Every [syntactician]_{F1} will *only* vote for the change.

Based on the context in (200), the F-marked constituent "syntactician" contrasts with "linguist," as reflected in the set of focus assignment functions *H* in (201) below. I assume here that "syntactician" $\not\subseteq$ "linguist." I present the computation of components of (200) below, following the computation for example (193) above.

⁷⁸Relevant world knowledge: this fictional department also has non-linguist members. It's a Department of Linguistics and Philosophy.

(201) $H = \{h_0, h_1\}$

- a. $h_0(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a syntactician in w
 b. $h_1(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x$ is a linguist in w

(202) **Computing (200):**

* Every [syntactician]_F will *only* vote for the change.

a. Narrow syntax:

$\underbrace{[\text{Every } [\text{syntactician}]_{\text{F1}}] \text{ FUT } \text{only } [\text{every } [\text{syntactician}]_{\text{F1}}]}_{\text{presupposition}}$ vote for the change

b. LF after TC:

$[\text{Every } [\text{syntactician}]_{\text{F1}}] \lambda 1 . \text{FUT } \text{only } [{}_{vP} [\text{the } [\text{syntactician}]_{\text{F1}} \textcircled{1}]]$ vote for the change

c. $[[\text{the } [\text{syntactician}]_{\text{F1}} \textcircled{1}]] (h \parallel g) = \lambda w_s : \underbrace{g(1) \text{ is a } h(\text{F1})(g)(w) \text{ in } w}_{\text{presupposition}} . g(1)$

d. $[{}_{vP}] (h \parallel g) = \lambda w_s : \underbrace{g(1) \text{ is a } h(\text{F1})(g)(w) \text{ in } w}_{\text{presupposition}} . g(1)$ vote for the change in w

e. $[\text{only } {}_{vP}] = \lambda g_a . \lambda w_s . \forall h \in H (h \neq h_0 \rightarrow [{}_{vP}] (h \parallel g)(w) \text{ is false})$

$= \lambda g_a . \lambda w_s . [{}_{vP}] (h_1 \parallel g)(w) \text{ is false}$

$= \lambda g_a . \lambda w_s : g(1) \text{ is a } h_1(\text{F1})(g)(w) \text{ in } w . g(1) \text{ does not vote for...in } w$

$= \lambda g_a . \lambda w_s : \underbrace{g(1) \text{ is a linguist in } w}_{\text{presupposition}} . g(1) \text{ does not vote for the change in } w$

f. $[[\lambda 1 . \text{only } {}_{vP}]] = \lambda g_a . \lambda w_s . \lambda x_e : x$ is a linguist in $w . x$ does not vote for the change in w

The denotation we yield for “ $\lambda 1 . \text{only } {}_{vP}$ ” includes the presupposition that its individual argument must be a linguist. This predicate forms the nuclear scope of the subject “every syntactician.” The presupposition will project through the quantification, by applying universally to each individual in the domain of quantification, “syntactician.” This yields the following presupposition for (200):

(203) **The presupposition of (200), due to the Trace Converted lower copy of movement:**

$\sim \forall x (x \text{ is a syntactician in } w_0 \rightarrow x \text{ is a linguist in } w_0)$

The presupposition in (203) accords with the context given and our world knowledge. Let's assume that it does indeed hold. We will now set aside this presupposition stemming from the lower copy of movement, and consider the rest of the interpretation of (200).⁷⁹ Recall that *only* presupposes the truth of its prejacent, in addition to asserting the negation of the non-prejacent alternatives. In this case, the presupposition of $[\text{only } {}_{vP}]$ is as in (204) below.

⁷⁹Because we are now limiting ourselves to the case where this presupposition holds, to simplify the presentation I will not continue to illustrate its projection below.

(204) **The prejacent presupposition of *only* (assignment-dependent):**

Given an ordinary assignment function g and world w ,

$$\llbracket \textit{only } vP \rrbracket (g)(w) \rightsquigarrow \llbracket vP \rrbracket (h_0 \parallel g)(w) \text{ is true} \iff g(1) \text{ votes for the change in } w$$

Because this presupposition is ordinary-assignment-dependent, we must ask how this presupposition will project. Following the literature on presupposition projection, reviewed in section 3.2 above, based on the higher quantifier binding this variable being “every,” this presupposition will project universally over that domain of quantification, “syntacticians.”⁸⁰ This gives us the following presupposition for the entire sentence (200):

(205) **The prejacent presupposition of *only* in (200):**

$$\rightsquigarrow \forall x (x \text{ is a syntactician in } w_0 \rightarrow x \text{ votes for the change in } w_0)$$

This prejacent presupposition seems to be exactly what we might expect from *only* in sentence (200), and is completely unproblematic by itself. But a problem occurs when we consider the truth conditions of the entire sentence in light of this prejacent. This is computed in (206) below, by combining the subject “every syntactician” with its nuclear scope, which is computed in (202f) above as “ $\lambda g_a . \lambda w_s . g(1)$ does not vote for the change in w .” The negation here was introduced by the assertion of *only*.

(206) **Truth conditions for (200):**

$$\forall x (x \text{ is a syntactician in } w_0 \rightarrow x \text{ does not vote for the change in } w_0)$$

Now consider the relationship between the prejacent presupposition and the assertive content that we have computed for (200). Given the truth of the prejacent presupposition in (205), *the assertion in (206) is necessarily false*. We could make the presupposition and the assertion compatible with one another by adding a high negation, which by assumption affects the assertion but not the presupposition, but this is also ungrammatical:

(207) * It is *not* the case that [every [syntactician]_F] will *only* vote for the change].

- a. presupposition: $\forall x (x \text{ is a syntactician in } w_0 \rightarrow x \text{ votes for the change in } w_0)$ (=205)
- b. assertion: $\neg \forall x (x \text{ is a syntactician in } w_0 \rightarrow x \text{ does not vote for the change in } w_0)$
 $\iff \exists x . ((x \text{ is a syntactician in } w_0) \wedge \neg(x \text{ does not vote for the change in } w_0))$
 $\iff \exists x . ((x \text{ is a syntactician in } w_0) \wedge (x \text{ votes for the change in } w_0))$

In this case, the assertion is not incompatible with the presupposition, but the assertion *will always be true*, given the truth of the prejacent presupposition.

⁸⁰The problem that I demonstrate here will also occur if the quantifier were existential and the prejacent presupposition projects a “bound” inference. This will be illustrated below for the general case.

I argue, following the work of Gajewski (2002), that these cases of backwards association with *only* result in meanings that are *always true* or *always false*, and therefore are not useful meanings to contribute to a conversation. This reasoning must of course be qualified as a general rule. As noted by Grice (1975), the sentences in (208) below seem to be tautologies, but are certainly grammatical sentences of English and can be used in a conversation when flaunting his maxim of Quantity.

(208) Examples of grammatical tautologies, from Grice (1975, p. 52):

- a. Women are women.
- b. War is war.

Gajewski (2002) develops a formal notion of triviality based on the logical structure of a sentence, which he calls “logical analyticity” or “L-analyticity.” Informally, we first identify the open-class items in a logical form, whose denotations are not invariant under permutations of the domain of individuals D_e , and replace them with variables of the appropriate type. If the resulting logical form is always true or always false, under every variable assignment, the structure is L-analytic.

Notice that Grice’s sentences above in (208) are *not* L-analytic: each instance of “women” in (208a) or each instance of “war” in (208b) could be replaced with other predicates of the same type and yield a contingent truth-condition. In contrast, Gajewski (2002) shows that logical forms which violate the Definiteness Restriction of *there*-constructions (Milsark, 1977; Barwise and Cooper, 1981) or combine a *but*-exceptive with an upward monotone quantifier (von Stechow, 1993), which are judged as ungrammatical, are L-analytic: that is, they do not improve under variable replacement. Gajewski (2002) proposes that logical forms that contain an L-analytic constituent are judged as ungrammatical.

I argue that configurations of backwards association with *only*, where the presuppositions introduced by lower copies of movement are satisfied, are L-analytic, when considering the truth condition of the logical form together with the prejacent presupposition of *only*.

In the remainder of this section, I will now consider the interpretation of backwards association with *only* in more general, schematic terms, to show that these problems always occur. I begin with the structural schema which I describe here, repeated from (168) above. A DP which contains F-marking in its restrictor is copied to a position outside of the scope of *only*. After Trace Conversion of the lower copy of movement, we yield the logical form in (209b). (The “<...prejacent...>” below stands in for the stated value for the F-marked constituent.)

(209) The structural schema for backwards association with *only*: (=168)

- a. NS: $[\text{DP Quant } [\alpha \dots \langle \dots \text{prej} \dots \rangle_{\text{F}} \dots]] [\gamma \dots \text{only } [\delta \dots [\text{DP Quantifier } [\alpha \dots \langle \dots \text{prej} \dots \rangle_{\text{F}} \dots]] \dots] \dots]$
- b. LF: $[\text{DP Quant } [\alpha \dots \langle \dots \text{prej} \dots \rangle_{\text{F}} \dots]] \lambda 1 [\gamma \dots \text{only } [\delta \dots [\text{DP the } [\alpha \dots \langle \dots \text{prej} \dots \rangle_{\text{F}} \dots] \textcircled{1} \dots] \dots] \dots]$

I will assume here that there is just one relevant alternative to the F-marked value and therefore $H = \{h_0, h_1\}$. The semantics of *only* presupposes the truth of its complement, evaluated using the prejaent focus assignment function h_0 , and asserts the negation of its complement, evaluated using the alternative focus assignment function h_1 .

(210) Interpreting backwards association with *only*: (=169)

[DP Quantifier ...<...prej...>F...] $\lambda 1$ [... *only* [... [DP the [...<...prej...>F...] ①] ...] ...]

a. asserts: [DP Quant [α ...<...prej...>...]] $\lambda 1$ [γ ... NEG [δ ... [DP the [β ...<...alt...>...] ①] ...] ...]

b. presupposes: [δ ... [DP the [α ...<...prejaent...>...] ①] ...] is true

In these schemas above, I use the following labels:

- (211) a. α is the NP restrictor of the DP computed using h_0 .
 b. β is the NP restrictor of the DP computed using h_1 .
 c. γ is the sister of the higher copy of movement.
 d. δ is the complement of *only*.

I will discuss how this structure is interpreted, in three different cases, depending on the relationship between the prejaent value of the NP restrictor, α , and the alternative value of the NP restrictor, β . Thinking of these denotations of extensional type $\langle e, t \rangle$ as sets—that is, the set of individuals for which the predicate holds—I will discuss cases (i) where α and β are disjoint, (ii) where α and β overlap but α is not a subset of β , and (iii) where α is a subset of β . Note that it cannot be the case that $\alpha = \beta$; if so, the complement of *only* will not be focus-assignment-sensitive, violating the Focus Assignment Dependency Requirement (43). I also assume that both α and β are non-empty. Therefore these three cases cover all possibilities.

Case i: $\alpha \cap \beta = \emptyset$

Assume α and β are disjoint. Let's look at the computation in (169) above. Here, due to the semantics of *only*, the Trace Converted lower copy of movement is interpreted using h_1 , yielding a denotation using the alternative value of the F-marked constituent: "[DP the [β ...<...alternative...>...] ①]." This definite description projects an existential presupposition, that ① satisfies the predicate β .

This assignment-dependent presupposition will project through the higher quantification. As established in section 2.4, the presupposition of Trace Converted lower copies project universally over the higher domain, so in this case we yield the presupposition that $\forall x (\alpha(x) \rightarrow \beta(x))$, or in set terms, $\alpha \subseteq \beta$. This presupposition will not hold given the assumed disjointness of α and β .⁸¹

⁸¹It has been hypothesized by Wagner (2005, *et seq*) that focus alternatives must always form a partition and therefore be pairwise disjoint. For our purposes, this would mean that α and β are always disjoint (case i), and other cases

Case ii: $\alpha \cap \beta \neq \emptyset, \alpha \not\subseteq \beta$

Assume that α and β overlap, but that α is not a subset of β . Here again, the Trace Converted lower copy of movement will introduce a presupposition that ① must satisfy the restrictor of the lower copy of movement, as interpreted using h_1, β . This presupposition will project universally over the higher domain of quantification, α , yielding the presupposition that, in set terms, $\alpha \subseteq \beta$. By assumption, α is not a subset of β , so the presupposition will not be satisfied.

Case iii: $\alpha \subsetneq \beta$

Assume that α is a proper subset of β . As discussed in the previous two cases, the lower copy of movement will introduce a presupposition that the bound variable satisfies β . This will project universally over the higher domain of quantification, requiring that $\alpha \subseteq \beta$. This will be satisfied by assumption.

With this presupposition satisfied, let's look at the structure in δ in (210). There are two things to note about δ . First, it is ordinary-assignment-dependent, including the variable ①. Second, in the structure in δ , only the restrictor of the lower copy of movement is focus-assignment-dependent. This restrictor of the lower copy of movement introduced the presupposition considered above. Since we have already shown that the presupposition of the lower copy of movement will be satisfied in this case, we can safely ignore the contribution of this lower copy's presupposition. As noted in section 2.4, with this presupposition satisfied, the lower copy DP is identical in denotation to the bound variable ①. This is true for both the lower copy DP with the restrictor β in (210a) and the lower copy DP with the restrictor α in (210b) above:

(212) Assuming that ① satisfies both α and β :

- a. $[[\text{DP the } [\beta \dots \langle \dots \textit{alternative} \dots \rangle \dots] \textcircled{1}]] = [[\textcircled{1}]]$
- b. $[[\text{DP the } [\alpha \dots \langle \dots \textit{prejacent} \dots \rangle \dots] \textcircled{1}]] = [[\textcircled{1}]]$

As a result, we can simplify the assertion and presupposition from (210) as follows:

(213) Interpreting backwards association with *only*, when $\alpha \subseteq \beta$: (based on 210)

$[\text{DP Quantifier } [\alpha \dots \langle \dots \textit{prejacent} \dots \rangle \dots]] \lambda 1 [\gamma \dots \textit{only} [\delta \dots \textcircled{1} \dots] \dots]$

- a. asserts: $[\text{DP Quantifier } [\alpha \dots \langle \dots \textit{prejacent} \dots \rangle \dots]] \lambda 1 [\gamma \dots \text{NEG} [\delta \dots \textcircled{1} \dots] \dots]$
- b. presupposes: $[\delta \dots \textcircled{1} \dots]$ is true

need not be considered. However, recent work by Katzir (2013, 2014) has called Wagner's disjointness condition into question. Here I show that the configuration is ungrammatical in other cases as well, and therefore my explanation for the ungrammaticality of backwards association with *only* does not rely on this hypothesized disjointness condition.

Now note that with this substitution, it is clear that the denotation of the constituent δ in (213a) is truth-conditionally equivalent to that of the constituent δ in (213b). This is an assignment-dependent presupposition, so there is a question of how it will project through the higher quantification, i.e. for what values of ① the presupposition should hold. As discussed in section 3.2, there are two possible ways that presuppositions project through quantification, depending on the choice of higher Quantifier. These two strategies are summarized here:

(214) Options for presupposition projection through quantification: (=119)

$$[\text{DP Quantifier Domain}] \lambda x \dots [\text{presupposition trigger } \underbrace{[\dots x \dots]}_{\text{scope of the trigger}}]$$

scope of the quantifier

where *presupposition trigger* $\rightsquigarrow f(x)$

- a. “bound” inference: $\rightsquigarrow f(x)$, where x is the *witness* of the higher Quantifier
- b. universal inference: $\rightsquigarrow \forall x \in \text{domain} . f(x)$

Let’s consider what happens under each one of these projection options. Consider first the “bound” inference. If this projection strategy is available for the Quantifier in question, there must be a particular individual—call it W —that the projected presupposition holds for. The projected presupposition will then be that δ with index 1 mapped to the witness W will be true. Because δ with ① evaluated as W is presupposed to be true (213b), we know that the constituent δ in the assertion (213a)—which uses W to verify the quantifier—will also necessarily be true. All of this is the case regardless of the choice of open-class lexical items involved, which I have shown here by abstracting away from them, making δ in this logical form L-analytic. Because the logical form in (213) includes this L-analytic constituent, it is judged as ungrammatical (Gajewski, 2002). Note that additional operators above, for example negation, cannot reverse L-analytic status of this configuration; it will necessarily be either always true or always false.

Now consider the universal inference. In this case we presuppose that δ holds for all individuals in α as the value of index 1. Similar to the case of the “bound” inference above, this means that the constituent δ in the assertion (213a) will necessarily be true, as every possible value for ① here will be in the domain of quantification α . This makes δ in (213a) L-analytic. The logical form in (213) includes this L-analytic constituent, making it be interpreted as ungrammatical (Gajewski, 2002).

Summary

With the discussion of these three cases here, we have covered the logical space of relationships between the predicates α and β , the denotations of the restrictors of the quantificational DP under h_0 and h_1 , respectively. When $\alpha \not\subseteq \beta$, the presupposition introduced by the Trace Converted lower

copy of movement, a definite description, will not hold when projected (see section 2.4). When $\alpha \subseteq \beta$ does hold, we are now able to compute the assertion and preajacent presupposition for the utterance, and we see that its truth conditions will be a logical tautology (always true) or a contradiction (always false), given the preajacent presupposition. I showed that this is true due to the logical structure of the backwards association configuration, rather than due to any property of the particular lexical items chosen in the examples, making the structure L-analytic and thus ungrammatical (Gajewski, 2002). As a result, regardless of whether the projected presupposition from the lower copies of movement is satisfied or not, the meaning that this configuration introduces is intuitively not a semantic object which can productively contribute to a discourse.⁸² This explains the impossibility of backwards association with *only*.

4.4 Even

I now turn to the second half of my proposal: why *even* is able to associate with material which has moved out of its scope. In cases where an operator seemingly associates with F-marked material which has moved out of its scope, I propose that the operator is in fact associating with the F-marked predicate in the *lower copy* of the movement chain, within the operator's scope. In this section I will demonstrate how this yields a useable meaning when the operator is *even*.

I will demonstrate this approach using example (215), which is interpreted with *even* associating with the predicate "professor" in the subject. Because "a professor" was generated in Spec,vP position and then moved to the surface Spec,TP position, there is a lower copy of the subject within the scope of *even*. After the lower copy undergoes Trace Conversion (TC), we yield the LF representation in (215b). I will show that *even* in this configuration is able to associate with the F-marked "professor" in the lower copy.⁸³

(215) An example of backwards association with *even*:

✓ A [professor]_F will *even* come to the party.

a. Narrow syntax: [A [professor]_F] FUT *even* [a [professor]_F] came to the party

b. LF after TC: [A [professor]_F] λ1 . FUT *even* [_{vP} [the [professor]_F ①] come to the party]

In order to interpret *even*, we must compute the denotation of vP and also identify the set of focus assignment functions. Here I take "student" to be the contextually relevant alternative to the

⁸²"Though my problems are meaningless, that don't make them go away." — Neil Young, "On The Beach"

⁸³Here I continue to assume that *even* takes scope in its surface position. There is an idea in the literature—the so-called "scope theory" of *even*—whereby adverbial *even* takes scope in a position higher than its pronounced position. This approach will be discussed in chapter 6.

F-marked predicate “professor.” ($\llbracket vP \rrbracket$ and H here are the same as those used in the computation of the *only* example in the previous section, except that “student” and “professor” have been flipped.)

$$(216) \quad \llbracket vP \rrbracket (h \parallel g) = \lambda w_s : \underbrace{g(1) \text{ is a } h(F1)(g)(w) \text{ in } w}_{\text{presuppositions}} . g(1) \text{ come to the party in } w$$

$$(217) \quad H = \{h_0, h_1\}$$

$$a. \quad h_0(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a professor in } w$$

$$b. \quad h_1(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a student in } w$$

There are two details here worth note. First, the denotation for vP (216) includes the presupposition of the lower copy definite description. Furthermore, as is always the case in cases of backwards association, the content of this presupposition of $\llbracket vP \rrbracket$ is dependent on the choice of focus assignment (h), but the truth conditions of $\llbracket vP \rrbracket$ are not. I propose that the presupposition in $\llbracket vP \rrbracket$ is *locally accommodated* for the purposes of computing the scalar inference of *even*, i.e. that it is incorporated into the content of the truth conditions. For concreteness I think of these presuppositions as a definedness condition, and will represent local accommodation using the operator LA, defined below in (218).⁸⁴ The locally accommodated variant of $\llbracket vP \rrbracket$ is as follows:

$$(218) \quad \text{Local accommodation: LA}$$

$$LA(\varphi_t) \equiv \begin{cases} \text{true} & \text{if } \varphi \text{ is defined and true} \\ \text{false} & \text{otherwise} \end{cases}$$

$$(219) \quad LA(\llbracket vP \rrbracket (h \parallel g)(w)) \text{ is true} \iff g(1) \text{ is a } h(F1)(g)(w) \text{ in } w \wedge g(1) \text{ comes to the party in } w$$

Two alternatives to this local accommodation step are discussed below.

Second, the denotation for vP is dependent on the ordinary assignment (g) given. As discussed in chapter 3, this is handled by *even* quantifying generically over different ordinary assignment functions, which correspond to different referents for the variable with index 1. It is important for the resulting scalar inference of *even*, which I will present in a moment, that the variable with index 1 range over both professors and students. Here I will assume that the set of relevant assignments G are the assignments which map the index 1 to different elements in the set of “persons,” and map other indices to the value determined by the global assignment g_0 .

(220) The set of contextually relevant ordinary assignments G is such that:

⁸⁴The LA operator could be thought of as part of the definition of the semantics of *even*, similar to my choice of including GEN in the definition of *even*, quantifying over the set of relevant ordinary assignment functions in G . Note that the truth conditions of $\llbracket vP \rrbracket$ in (216), and in backwards association configurations more generally, is not focus-assignment-dependent. If *even* does not consider the presuppositional part of its scope, its complement would violate the Focus Assignment Dependency Requirement (43).

- a. for any $g \in G, i \neq 0, g(i) = g_0(i)$
 b. $\bigcup_{g \in G} g(1) = \llbracket \text{person} \rrbracket$

We are now ready to compute scalar inference of *even*. The definition I use for the semantics of *even* is repeated here from section 3.5.

(221) **Proposal for *even* (final version): (=151)**

$$\llbracket \text{even} \rrbracket_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle, t \rangle \rangle \rangle} \equiv \lambda g_a . \lambda w_s . \lambda P_{\langle a, \langle s, t \rangle \rangle} :$$

$$\underbrace{\text{GEN } (g' \in G, h \in H) (h \neq h_0 \rightarrow P(g' \parallel h_0) <_{\text{likely}} P(g' \parallel h))}_{\text{scalar inference}} \cdot \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\text{truth condition}}$$

(222) **The scalar inference of *even* in (215):**

$$\llbracket \text{even } vP \rrbracket (g)(w_0) \rightsquigarrow$$

$$\text{GEN } (g' \in G, h \in H) (h \neq h_0 \rightarrow (\lambda w . \text{LA} (\llbracket vP \rrbracket (g' \parallel h_0)(w))) <_{\text{likely}} (\lambda w . \text{LA} (\llbracket vP \rrbracket (g' \parallel h)(w))))$$

Note $H = \{h_0, h_1\}$, so simplify:

$$\iff \text{GEN } (g' \in G, h = h_1) (\lambda w . \text{LA} (\llbracket vP \rrbracket (g' \parallel h_0)(w))) <_{\text{likely}} (\lambda w . \text{LA} (\llbracket vP \rrbracket (g' \parallel h)(w)))$$

$$\iff \text{GEN } (g' \in G) \left(\begin{array}{l} (\lambda w . g'(1) \text{ is a } h_0(\text{F1})(g')(w) \text{ in } w \wedge g'(1) \text{ comes to the party in } w) <_{\text{likely}} \\ (\lambda w . g'(1) \text{ is a } h_1(\text{F1})(g')(w) \text{ in } w \wedge g'(1) \text{ comes to the party in } w) \end{array} \right)$$

$$\iff \text{GEN } (g' \in G) \left(\begin{array}{l} (\lambda w . g'(1) \text{ is a professor in } w \wedge g'(1) \text{ comes to the party in } w) <_{\text{likely}} \\ (\lambda w . g'(1) \text{ is a student in } w \wedge g'(1) \text{ comes to the party in } w) \end{array} \right)$$

$x = g'(1)$ ranges over people, across all $g' \in G$:

$$\iff \text{GEN } (x \text{ is a person}) \left(\begin{array}{l} (\lambda w . x \text{ is a professor and comes to the party in } w) <_{\text{likely}} \\ (\lambda w . x \text{ is a student and comes to the party in } w) \end{array} \right)$$

This yields the expected scalar inference of (215), which is satisfied in a context where it is more likely for a student than a professor to come to the party.

Note that the inference in (215) is also satisfied in a context where the likelihood of coming to the party conditioned on being a student is equal to that conditioned on being a professor, but where it is simply less likely to be a professor than to be a student, due to the relative rarity of professors. I believe this prediction is correct, as verified by the felicity of *even* in (223) below.

(223) **Context:** God said to Noah, “And of every living thing, of all flesh, you shall bring two of every kind into the ark, to keep them alive with you; they shall be male and female.” (Genesis 6:19) Noah did this.

✓ Two [unicorns]_F were *even* on the ark.

Given this context, for any category X of living things in the time of Noah, the probability that two X s are on Noah's ark is 1.⁸⁵ However the use of *even* is supported in (223). This then reflects the fact that, given a generic living thing in the Noah's time, it being a unicorn is less likely or more noteworthy than it being another type of living thing.

At this point, we are now done computing the scalar inference of *even*, and the truth conditions of vP have been left untouched by *even*. This is the crucial difference between *even* and *only*. *Only* uses the alternatives in its complement to compute its truth conditions, which will then compose with material above it. As illustrated in the previous section, having one copy of the F-marking inside the scope of *only* and one outside puts conflicting requirements on the variables being quantified over.

In contrast, *even* uses the alternatives to produce a projective scalar inference. This scalar inference then projects without composing with additional material above it. As we saw in chapter 3, not even the choice of higher quantifiers binding into the scope of *even* affect the inference computed by *even*. Furthermore, *even* does not modify the truth conditions of its complement. For example, in (215), the truth conditions of vP will be unchanged by *even* and will continue to compose with material above it, including the higher copy of the quantificational subject "a professor."

Before closing this section, I will briefly discuss what would happen if local accommodation were not used above. The denotation of the complement of *even*, vP , is a partial function with an assignment-dependent domain condition. We can think of a couple other strategies for working with this denotation, but neither of them produce the desired result.

One option is to ignore the definedness conditions for the purposes of computing *even*. Because, in this example, only the definedness condition of vP is focus-assignment-dependent, $[[vP]](h \parallel g)$ would no longer vary with different values of h , and therefore the scalar ranking $<_{likely}$ would never be satisfied.

A second option is to project these presuppositions in some way. However, when we unpack the meaning of *even*'s scalar inference, which ranges over different focus assignments $h \in H$, we see that different alternative propositions compared by $<_{likely}$ will each introduce conflicting requirements on the same variable. In the case of example (215) above, the variable with index 1 would be required to be a professor as well as a student.⁸⁶

⁸⁵I treat the numeral "two" as part of the content of the complement of *even* at LF, rather than part of the quantificational determiner which is replaced in the process of Trace Conversion. Evidence for this choice is presented later, in section 5.1. However, note that the argument here does not hinge on this treatment of the numeral.

⁸⁶There is an alternative to invoking an application of a "local accommodation" operation, as I present it here, whereby $<_{likely}$ is defined so that it takes the definedness condition of each comparand into account. However, the result will be the same as the derivation here involving LA operating on the complement of *even*.

In this section I demonstrated how *even* is able to apparently associate with material which has moved out of its scope, by associating with F-marking in the lower copy of the movement chain. In contrast, the same configuration with *only* is illicit, due to independent differences in the semantic contribution of *even* and *only*. As a special case of this configuration, the proposal also explains the ability of VP-*even* but not VP-*only* to associate with a leftward subject, first observed in Jackendoff (1972).

4.5 Associating with the higher copy of focus

My analysis for backwards association with *even*, presented above, relies on focus being a syntactic annotation which is copied in the process of movement. *Even* introduces a projective meaning based solely on the copy (or copies) of the F-marking in its scope. There is, however, also an instance of F-marking in the higher copy of movement—after all, in the overt movement cases that we have looked at thus far, it is this higher copy of focus which will be realized prosodically. What happens to this higher copy of focus? Can another focus-sensitive operator associate with it? This configuration is schematized here:

(224) **Associating with the higher copy of focus:**

$$Op_1 [\dots [DP \dots \alpha_F \dots] \dots [Op_2 [\dots [DP \dots \alpha_F \dots] \dots]]]$$

with both Op_1 and Op_2 using alternatives that vary based on α_F

My system for focus evaluation predicts that higher focus operators would indeed be able to associate with the higher copy of focus.⁸⁷ The fact that the same focus-index has already contributed to the meaning of a lower operator should not be a concern; the denotation of the complement of the higher operator, Op_1 , will still be focus-assignment-sensitive. The utterance is then predicted to carry the meaning of both focus-sensitive operators, associating with the same focused constituent. In this section I will present evidence that supports this prediction.

I begin by noting that the alternatives introduced by focus contribute to the interpretation of a sentence even when there is no overt focus-sensitive operator. This is often called “free focus.” The placement of free focus is relevant for a variety of phenomena, including discourse congruence. The following examples in (225–226) show that the position of F-marking in backwards association constructions does contribute to the evaluation of question-answer congruence. Note in particular

⁸⁷Technically, the higher operator would not be associating with just the F-marking in the higher copy, but instead would quantify over alternatives derived by covarying all the instances of the F-marked predicate in the operator’s scope at LF. See section 2.3 above for a demonstration of this process with *only*.

that the last sentence in (225B) and the sentence in (226Bii) are the same, with the same backwards association of *even* with the restrictor of the subject, "professors."

(225) A: I heard the big homecoming football game is this coming Saturday. *Who will come to the game?*

B: All sorts of people. Students will come to the game. Alumni will come to the game.

✓ Some [professors]_F will *even* come to the game.

(226) A: There's a lot of things going on on campus on Saturday. Many students will be at the game. *What will the professors be doing?*

B: Some professors will be at home with their families. Some professors will go to the orchestra concert.

i. ✓ Some professors will *even* [come to the game]_F.

ii. # Some [professors]_F will *even* come to the game.

The utterance is felicitous in (225) when the question under discussion (in italics) is about different people coming to the game, but is infelicitous in (226) when the question under discussion is about the activities of the professors.⁸⁸ Note that the context otherwise supports the felicitous use of *even* associating with "professor" in (226Bii): (a) "professors" have a natural and previously mentioned alternative, "students"; (b) professors are less likely to come to the game than students are,⁸⁹ satisfying the scalar inference of *even*; and (c) it has been established that some others (students) will come to the game, satisfying the additive inference of *even*. The infelicity of (226Bii) therefore must be due to the failure of F-marking on "professors" to satisfy discourse congruence.

One common approach to modeling the effects of free focus is to posit a covert focus-sensitive operator which evaluates the focus structure of the utterance against the current discourse context.⁹⁰ Under this view, the effect above shows that the F-marking on "professor" is being interpreted by an additional focus-sensitive operator outside of the scope of *even*.

Now let us turn to examples with an overt focus-sensitive operator above, associating with the higher copy of focus. Because the semantics of both the higher and lower operators must be satis-

⁸⁸The sentences in question here constitute one part of what could be described as a list or conjunction of partial answers to the question under discussion. It is known that free focus placement is also evaluated for congruence between items in a list or conjunction (Rooth, 1992b). The contrast here between (225) and (226) could therefore be due to the fact that this list-internal contrast congruence holds between the individual sentences in B's reply in (225) but not (226), instead of or in addition to the issue of question-answer congruence.

⁸⁹assumed world knowledge

⁹⁰This idea that free focus is actually associating with a covert illocutionary operator is credited by von Stechow (1991a) and Krifka (1995) to the work of Joachim Jacobs (e.g. Jacobs, 1984). The same idea is implemented in Rooth (1992b) by adjoining his squiggle (~) operator to the root of the utterance.

fied simultaneously, constructing felicitous examples of this form can be challenging. Nonetheless, I believe association in such configurations is in general possible.⁹¹ Here are some examples with contexts; the target sentence comes at the end of each story.

(227) **High *only* with backwards-associating *also* below:**⁹²

Everyone knew that the President would speak at the press conference, but wasn't sure who else would. We expect journalists like John to know who will speak at these events. It turned out that multiple people spoke in addition to the president. But John *only* knew that [Biden]_F would *also* speak.

(228) **High *also* with backwards-associating *even* below, and vice versa:**⁹³

The fortune teller predicted lots of surprising results about the Olympics, which turned out to be true. She correctly predicted that an [Estonian]_F would win a gold medal in skiing.

a. She *also* correctly predicted that a [Fijian]_F would *even* win a gold medal (in skiing).

b. She *even* correctly predicted that a [Fijian]_F would *also* win a gold medal (in skiing).

It is important in such examples to ensure that the higher focus-sensitive operator is indeed associating with the same F-marking, rather than some larger constituent containing the narrow F-marking used by the lower operator. This is clearest in example (227), where the higher operator is *only*, which uses the alternatives to construct its truth-conditional meaning. We might be concerned, for example, that *only* actually associates with the entire complement of “know” in (227), the proposition that Biden would also speak. However, this would entail that John does not know

⁹¹Using *even*, *only*, and *also*, there are six possible combinations with the structure in (224). Of these, *Op*₂ cannot be *only*, as it cannot associate backwards, ruling out two possibilities. Furthermore, I have found that examples with a higher *only* with a lower *even* associating backwards are difficult to judge. The scale reversal behavior of *even*, to be discussed in chapter 6, may also be relevant to understanding why such examples are difficult. This leaves three possible combinations, which are all attested here. It's worth noting that judgments of these types of examples can vary by speaker. I thank Hadas Kotek for discussion and for help in constructing these examples.

⁹²The lower operator here is *also*, associating backwards with the subject, Vice President Joe Biden. Like *even*, [*also*]_F can *also* associate with material which has moved out of its scope. This will be discussed in the Appendix to chapter 5.

⁹³The inference of *also* in (a) seems to be that “there is an alternative X, such that she predicted that an X would win a gold medal,” satisfied by X = “Estonian.” Note that this inference of *also* does not include its own instance of *even*; if it were included, the use of *also* would not be felicitous here, as there was no *even* associating with “Estonian” in the preceding sentence. Similarly, in the (b) variant, the inference of *even* expresses the unlikelihood of predicting that a Fijian would win a gold medal—without the contribution of the *also*—rather than the unlikelihood of a Fijian winning a gold medal in addition to someone else.

Michael Wagner (p.c.) notes that this may be explained by the theory of Sauerland (2013), which argues that “purely presuppositional” meanings, which do not modify the truth-conditions of the assertion, can be included or not in the content of focus alternatives. The meanings of *even* and *also* satisfy Sauerland's criteria for being purely presuppositional.

anything else, which is not true in (227). We can therefore be fairly certain that we are testing the right configuration, where both the higher and lower focus-sensitive operators are associating with the same position; i.e. quantifying over alternative propositions which vary in the position of the given F-marking.

4.6 Focus association and covert movement

All of the examples that we've looked at thus far have involved *overt* movement of a DP containing F-marking in its restrictor out of the scope of a focus-sensitive operator. I've shown that in such configurations, *even* is able to associate backwards but *only* cannot, and have presented a proposal which derives this difference from the operators' compositional semantics.

The explanation I presented in the last two sections for this difference between *even* and *only* is based entirely on the LF representation. I therefore predict that *covert* movements will also be sensitive to this pattern of backwards association, being possible with *even* but not *only*. We can think of covert movement in terms of its output interface representations, schematized in (229): a constituent is interpreted at LF in a different (normally higher) position than its pronounced position at PF.⁹⁴

(229) Covert movement:

- a. PF: ... [Op [... [DP ... α_F ...] ...]]
 b. LF: [DP ... α_F ...] ... [Op [...] ...]
-

In this section I show that this prediction is correct, in particular looking at the interaction of focus association and covert scope shift. We will see that association of a quantifier's restrictor with *only* but not *even* will restrict the possible scopes of the quantifier, so that the quantifier must be interpreted with scope below the focus-sensitive operator. This behavior can be subsumed under our generalization that *even* but not *only* can associate with F-marked material which has moved out of the operator's scope, provided that quantifier scope shift involves a (covert) movement relation. This parallel between overt movement and quantifier scope shift therefore acts as an argument for the covert movement approach to quantifier scope shift.

Consider first a baseline case of scope ambiguity presented in (230).

⁹⁴Under the Roothian model of focus interpretation I adopt here, focus association is sensitive only to the position of F-marking in the LF_F representation (cf Reinhart, 2006). Therefore the data from focus association that I discuss here will not be able to distinguish between covert movements which occur in the narrow syntax but then pronounce their tail position at PF and covert movements which occur strictly in the LF branch.

- (230) Someone wants to meet every boy in the room. $\checkmark \exists > \forall, \checkmark \forall > \exists$ ⁹⁵
- a. $\exists > \text{want} > \forall$: Mary is told that there are boys in the room and is looking forward to meeting each of them when she enters the room. She doesn't know how many boys there are, or who they are.
- b. $\forall > \exists > \text{want}$: Three boys, Mark, Kyle, and Ian, are alone in the room. Three girls each want to meet one of the boys: Mary wants to meet Mark, Kerri wants to meet Kyle, and Isabel wants to meet Ian.

One common approach to quantifier scope shift analyzes readings with different quantifier scopes correspond to different LFs generated by different choices of Quantifier Raising (QR), a covert movement operation (May, 1977, a.o.). In (230) the inverse scope readings ($\forall > \exists$) are available due to QR of "every boy in the room" to a position above "someone" at LF, as seen in (230'). As "want" is a control verb, I take lowering of "someone" to not be a viable operation. Thus, with the scope of "someone > want" fixed, there are three predicted readings, which are all attested (see footnote 95), two of which are illustrated with contexts in (230) above.

- (230') a. LF for $\exists > \text{want} > \forall$: someone wants [[every boy...] PRO meet]
- b. LF for $\forall > \exists > \text{want}$: [every boy...] someone wants [PRO meet]

If the different scopes of *every boy* in (230) indeed involve covert movement of the DP, without affecting the PF output, we would expect this movement to bleed association with *only* but not with *even*. This is indeed what we observe. In example (231) below, I introduce *only* within the nonfinite embedding, associating with "boy," and we see that only the surface scope interpretation is available.⁹⁶

⁹⁵An additional reading with the scope relation " $\exists > \forall > \text{want}$ " is also possible. However, it is not discussed here as this intermediate scope reading is difficult to verify in later examples, whereas the availability of the " $\forall > \exists > \text{want}$ " reading is easier to judge.

⁹⁶There does seem to be a *de re* reading of "every boy" with regard to "want." Consider the following context: Bill wants to meet Mark, Kyle, and Ian when he gets in the room. As far as Bill knows, there may be girls in the room as well as other boys, but he doesn't want to meet them. In reality, there are three boys and three girls in the room: Mark, Kyle, and Ian and Mary, Kerri, and Isabel. So Bill wants to only meet every boy in the room.

In this reading the *de dicto* statement of Bill's desires is not to meet every boy in the room, but rather to meet the particular boys Mark, Kyle, and Ian, which we know to correspond to "every boy in the room" in this context.

Under one theory of *de re* interpretation, the quantifier would have to be interpreted outside of the scope of the intensional operator (here, "want"). However, another option may be that the *de re* "every boy in the room" is interpreted *in situ* as a type *e* expression. If so, this is not a counterexample to the generalization that *only* cannot associate with material in DPs which have moved out of *only*'s scope. See Keshet (2008) for different approaches to *de re* interpretation.

- (231) **Association with *only* restricts the scope of “every boy”:** (based on Aoun and Li, 1993)⁹⁷
 Someone wants to *only* meet [every [boy]_F in the room]. $\checkmark \exists > \forall, * \forall > \exists$
 $\exists > \text{want} > \text{only} > \forall$: To pass the “socialization” exam, Bill must go into a room and talk to either every boy in the room or every girl in the room. Bill is not an overachiever, so he wants to do the bare minimum to pass. Moreover, as he prefers to talk to other boys, he wants to only meet every [boy]_F in the room.

However, such an effect of focus association limiting the scope of “every boy” is not observed with *even*:

- (232) ***Even* does not restrict the scope of “every boy”:**
 Someone wants to *even* meet [DP every [boy]_F in the room]. $\checkmark \exists > \forall, \checkmark \forall > \exists$
- a. $\exists > \text{want} > \text{even} > \forall$: There are 10 girls in the room and 100 boys, making it much more feasible for someone to meet all the girls in the room than for someone to meet all the boys in the room. Mary wants to meet every girl in the room. Furthermore, she wants to even meet every boy in the room.
- b. $\forall > \exists > \text{want} > \text{even}$: Two boys, Mark and Kyle, and two girls, Mary and Kerri, are alone in the room. The boys in the room are generally less sociable than the girls in the room, and therefore it is more likely that someone would meet one of the girls in the room than for someone to meet one of the boys in the room.
 Alex, Billy, Chris, and Dana are not in the room and don’t know where anybody else is. Alex wants to meet Mark, Billy wants to meet Kyle, Chris wants to meet Mary, and Dana wants to meet Kerri. Therefore, we can say that someone wants to meet every boy in the room.

⁹⁷ Aoun and Li (1993) discuss an example similar to (231), “someone only loves every [boy]_F in the room,” with the judgment $\checkmark \exists > \forall, * \forall > \exists$, and thereby claim that association with *only* restricts quantifier scope. However, inverse scope readings for this sentence are in fact available given an appropriate context, and therefore their claim is not supported by the data:

- (i) In order to win the Valentine’s Day contest, everyone in your school (regardless of gender) must be loved by someone ($\forall > \exists$). However, our school lost because *someone only loves every [boy]_F in the school* ($\forall > \exists$).

Note, however, that this reading is compatible with an LF with “every [boy]_F” taking scope below *only*. This reading corresponds to reconstruction of the “someone” to its base vP-internal position below *only*, with QR of “every [boy]_F” above “someone” but still below the VP-*only*.

- (ii) LF for *only* > $\forall > \exists$: *only* [_{vP} [every boy...] someone loves ____]

The control predicate “want” is added in the examples here in order to avoid this confound. I thank Maziar Toorsarvandani for discussion of these examples.

QRed to a position at least as high as the edge of the intended antecedent VP. Therefore the QR is required to be higher in (234b), to create the antecedent VP “want to read,” than in (234a), to create the antecedent VP “read.”

(234) LFs for each reading of (233):

- a. ✓ $\Delta = \text{“read”}$:
 John wants to [[every book that Mary did Δ] [antecedent read]]
- b. ✓ $\Delta = \text{“want to read”}$:
 John [[every book that Mary did Δ] [antecedent wants to read]]
-

Judgments regarding ACD resolution can be made clearer by using predicates which require different auxiliaries at the ellipsis site.

(235) Different auxiliaries forcing larger and smaller ACD:

- a. John is willing to read [every book that Mary *did* Δ]. ✓ “read,” * “willing to read”
- b. John is willing to read [every book that Mary *is* Δ]. * “read,” ✓ “willing to read”

Now consider example (236), where *only* has been adjoined to the nonfinite embedding, above “read,” associating with the head noun “book.”⁹⁹ The ellipsis is now only possible as “read” with the auxiliary “did,” and not with the larger ellipsis with auxiliary “is.”¹⁰⁰ This contrast is explained by the fact, presented here, that *only* can restrict the height of QR. The ellipsis resolution of “read” is possible because the DP—containing both the intended F-marked associate of *only* and the VP ellipsis gap—is able to QR to a position below *only* but high enough to create an antecedent VP “read” which does not include the ellipsis gap (236a). In contrast, an antecedent VP corresponding to “willing to (*only*) read” cannot be constructed. To do so would require QR of the DP to a position above “willing” (236b), and therefore the F-marked “book” will be copied to a position outside of the scope of *only*, leading to an illicit structure.

⁹⁹David Pesetsky (p.c.) notes that if (236) is modified so that *only* is adjoined to the DP “every book...,” the larger ACD becomes possible. This is expected because the DP-adjoined *only* is able to take scope in higher clauses, in contrast to adverbial *only* which always takes scope in its surface position (Taglicht, 1984).

¹⁰⁰Large ACD is similarly ungrammatical if *only* associates with material in the relative clause itself (ii), although the explanation will be different. See the discussion of (239) below.

- (i) ✓ John is willing to *only* read [every book that [Mary]_F *did* Δ]. $\Delta = \text{“read”}$
- (ii) * John is willing to *only* read [every book that [Mary]_F *is* Δ]. $\Delta = \text{“willing to (only) read”}$

(236) The height of *only* restricts ellipsis resolution:

- a. ✓ John is willing to *only* read [every [book]_F that Mary *did* Δ]. Δ = “read”
 LF: John wants to *only* [[every [book]_F that Mary *did* Δ] [antecedent read]]
- b. * John is willing to *only* read [every [book]_F that Mary *is* Δ]. = “willing to (*only*) read”
 LF: John [[every [book]_F that Mary *did* Δ] [antecedent wants to *only* read]]

Now consider the case of association with *even*. Here both the larger and smaller ACD resolutions are possible, corresponding to the fact that association with *even* does not restrict the height of covert movement. This is presented here with informal LFs:

(237) The height of *even* does not restrict ellipsis resolution:¹⁰¹

- a. ✓ John is willing to *even* read [every [book]_F that Mary *did* Δ]. Δ = “read”
 LF: John wants to *even* [[every [book]_F that Mary *did* Δ] [antecedent read]]
- b. ✓ John is willing to *even* read [every [book]_F that Mary *is* Δ]. = “willing to (*even*) read”
 John [[every [book]_F that Mary *did* Δ] [antecedent wants to *even* read]]

Let’s look more carefully at the derivation of example (237b). Following Fox (2002), the relative clause in ACD is base-generated high, in a position outside of the antecedent VP; only the head noun restrictor is generated in the lower copy. This will include the F-marked “book,” in the scope of *even*. *Even* can associate with this F-marked “book” in the lower copy of movement:

(238) Detailed derivation for (237b):

John is [[every [book]_F] [antecedent willing to *even* read [every [book]_F]]]

↓ Late merger of relative clause (adjunct)

John is [[every [book]_F] [that Mary is Δ]] [antecedent willing to *even* read [every [book]_F]]]

↓ Trace Conversion

John is [[every [book]_F] [that Mary is Δ]] λx [antecedent willing to *even* read [the [book]_F x]]]
 ⇒ the scope of *even* includes the associate, “Mary” ☺

¹⁰¹It is difficult to intuit whether the ellipsis site in (237b) includes its own instance of *even* or not, although I suspect that it does not. Michael Wagner (p.c.) notes that this may be related to the theory of Sauerland (2013), which argues that “purely presuppositional” meanings such as *even*, which do not modify the truth-conditions of the assertion, can be included or not in the content of focus alternatives. Under the view that ellipsis identity conditions are based on focus alternative values (Rooth, 1992a,b), Sauerland’s theory explains why *even* need not be included in the ellipsis here.

Note, however, that under this derivation for the large ACD, the content of the relative clause is never in the scope of *even*. This predicts that this large ACD will be impossible if *even* associates with material in the relative clause, rather than with the head noun as in (237b) above. This prediction is borne out:

(239) *Even* associating with material in the relative clause *does* restrict large ACD:

- a. ✓ John is willing to *even* read [every book that [Mary]_F did Δ]. Δ = "read"
- b. * John is willing to *even* read [every book that [Mary]_F is Δ]. = "willing to (*even*) read"

4.7 Detecting lower copies

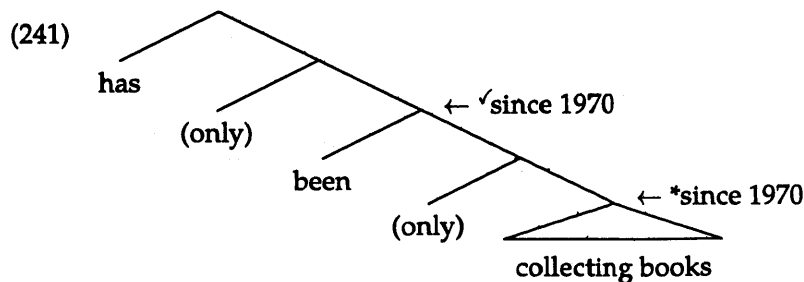
Given that the focus-sensitive operator *only* must associate with an F-marked associate in its scope, association with *only* has been used as a diagnostic of structural c-command relations (Aoun and Li, 1993; McCawley, 1995; Sportiche et al., 2014).

For example, McCawley (1995) uses *only* to determine the height of the temporal adverbial "since 1970" in (240a) relative to the verbal heads in the clause. In example (240b), we see that *only* between "has" and "been" is able to associate with "1970." This same pattern of association is impossible when *only* is placed further downstream, between "has been" and "collecting" in example (240c). Example (240d) shows that this position, between "has been" and "collecting," is a valid position for the adverb *only* to occupy while associating with VP-internal material.

(240) *Only* as a c-command diagnostic (McCawley, 1995, p. 175):

- a. Baseline: John has been collecting books since 1970.
- b. John has *only* been collecting books since [1970]_F.
- c. * John has been *only* collecting books since [1970]_F.
- d. John has been *only* collecting [books]_F since 1970.

Based on this contrast, McCawley (1995) proposes that "since 1970" must adjoin to a position above "been." The tree below illustrates the possible attachment height of "since 1970" according to this diagnostic:



In this way, association with *only* has been used as a diagnostic of structural c-command. If *only* is able to associate with an F-marked constituent α , *only* must c-command α . In this chapter I have presented an explanation for why this c-command restriction holds, based on the compositional semantics of *only*.¹⁰² Therefore, we can give the following description, in copy theoretic terms:

(242) **Association with *only* as a structural diagnostic:**

If *only* associates with α and there are multiple copies of α in the representation, *only* must c-command the *highest* copy of α .¹⁰³

At the same time, I highlighted the difference between *only* and *even* in their patterns of association. I showed that *even* but not *only* is able to associate with material which has moved out of its scope. Association with *even*, then, has the potential to be a new kind of c-command diagnostic:¹⁰⁴

(243) **Association with *even* as a structural diagnostic:**

If *even* associates with α and there are multiple copies of α in the representation, *even* must c-command *at least one* copy of α .

Note that in this chapter I have focused on F-marking in the restrictor of a moved DP. Therefore these structural diagnostics in (242–243) may only hold if α is *part* of the moved constituent, but not if α is the entire moved constituent. In chapter 5 I will show that backwards association with the entire moved constituent is impossible for both *even* and *only*.

In this section, I will briefly explore some implications of this new diagnostic. Association with *even* allows us to detect whether a constituent actually originated lower or was base-generated in its high position.

¹⁰²We might wonder whether movement out of the scope of *only*, and the consequential bleeding of focus association, can be “repaired” via reconstruction. This will be the subject of chapter 7.

¹⁰³Put another way, association of *only* with α entails that *only* c-commands *all* copies of α . It is possible that in cases of so-called *sideward movement* (Nunes, 2001), these two characterizations can be teased apart. I will not discuss such cases here.

¹⁰⁴Sportiche et al. (2014) suggests the use of association with *also* as a diagnostic of lower positions of movement, similar to what I advocate for here with *even*. Rullmann’s (2003) proposal that *also* is able to associate with traces may explain why this holds. See also the discussion in the appendix to chapter 5.

4.7.1 *It...that* clauses

I begin by examining a little-studied alternation in English, brought to my attention by Kenyon Branán, which I will call "*it...that* clauses."¹⁰⁵ English finite CP complements can appear with an "it" apparently resuming the clause:

- (244) The *it...that* alternation: (Kenyon Branán, p.c.)
- a. Sue believed [_{CP} that the universe emerged from a cosmic egg].
 - b. Sue believed *it* [_{CP} that the universe emerged from a cosmic egg].

When the *it...that* variant is used, the finite clause is interpreted as factive (Kiparsky and Kiparsky, 1970). In addition, the embedded clause seems to occupy a structurally higher position, as diagnosed by Condition C. In example (245a) below, we see that the indirect object of "mention" cannot be interpreted as coreferential with "Mary" below. This Condition C effect shows that the embedded clause is structurally lower than the indirect object.¹⁰⁶ When the *it...that* variant is used, in (245b), coreference becomes possible, indicating that the indirect object no longer c-commands "Mary" in the complement clause.

- (245) *It...that* clauses are higher: (Kenyon Branán, p.c.)
- a. * John mentioned to her_i [_{CP} that Mary_i won the contest].
 - b. ✓ John mentioned *it* to her_i [_{CP} that Mary_i won the contest].

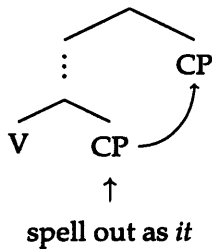
What is the structure of these *it...that* constructions? We can imagine two possibilities: either the CP originated in the normal position for complement clauses and then moved up, with the base position being spelled out as "it," or "it" is generated as the complement of verb, and then the corresponding clause is base-generated high. These two hypotheses are schematized below:

¹⁰⁵The discussion in this section comes from discussions with Kenyon Branán. I thank him for bringing this construction to my attention. Some discussion of these and other "it"s apparently resuming clauses are in Postal and Pullum (1988, §3.1.1) and references therein.

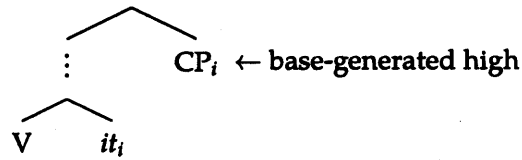
¹⁰⁶It is well known that the DP in certain PP arguments can behave for binding purposes as if it c-commands out of the PP. See also footnote 68.

(246) Two hypotheses for the *it...that* construction:

a. The movement hypothesis:



b. The base-generation hypothesis:



Here, patterns of association with *only* and *even* show that the movement hypothesis is superior. Example (247a) is a baseline showing that both *even* and *only* adjoined to the CP-selecting verb phrase can associate with an F-marked constituent inside the complement clause. When the *it...that* variant is used, in (247b), *only* is no longer able to associate with the F-marking in the clause. *Only*'s inability to associate with material in the *it...that* clause in (247b) is explained if the clause is outside of the scope of *only*. This is compatible with both the movement and base-generation hypotheses.

(247) *Even*, but not *only*, can associate into an *it...that* clause:

- a. I { \checkmark only } knew [CP that [John]_F was a spy].
 { \checkmark even }
- b. I { *only } knew it [CP that [John]_F was a spy].
 { \checkmark even }

In contrast, *even* continues to be able to associate with the constituent in the *it...that* variant. This is unexplained by the base-generation analysis, where the clause containing F-marking is base-generated high. It is however expected under the movement hypothesis. The lower copy of the moved CP will contain an instance of the F-marked constituent. This lower copy of the F-marking is sufficient to feed association with *even*. In other words, although this is not clear from the surface word order, this is a backwards association configuration.

4.7.2 Raising to object and object control

One of the original arguments I gave for backwards association with *even* requiring a lower copy of the F-marking in its scope involved a contrast between raising to subject and subject control. The relevant example, (174), is repeated below. We see that *even* is able to associate backwards with F-marking in the restrictor of the matrix subject in the raising example with "seem" (248a) but not in the control example with "want" (248b).

(248) **Backwards association and raising to subject vs subject control: (=174)**

- a. ✓ [DP A [professor]_F] seems to *even* be at the party. *raising to subject*
- b. * [DP A [professor]_F] wants to *even* be at the party. *subject control*

Backwards association with *even* thus distinguishes between these two superficially similar constructions. The contrast is explained by the common view that the control construction does not involve a movement chain, and therefore there is no lower copy of the F-marking in the scope of *even*.¹⁰⁷

In this section I will consider the superficially similar constructions of raising to object (also called Exceptional Case Marking or ECM) and object control. I begin with raising to object. Example (249) shows that *even* in the embedded nonfinite clause is able to associate with the matrix object/embedded subject ("DP") in a raising to object construction. This indicates that the argument originated within the scope of *even* and moved to its surface position.

(249) **Backwards association across raising to object/ECM:**

- ✓ David expected/believed/considered [DP the [laziest]_F students] to *even* be able to pass the course.

However, by this same diagnostic, it seems that the same is true of object control embeddings:

(250) **Backwards association across object control:**

- ✓ David persuaded/instructed/ordered [DP the [laziest]_F students] to *even* come take the exam.

The availability of backwards association in the object control construction (250) teaches us that the DP containing F-marking must have originated within the scope of *even*, within the nonfinite embedding. This could be explained by adopting a movement derivation for object control constructions (Hornstein, 1999, a.o.). Note, however, that this behavior in (250) contrasts from that of subject control in (248b), where backwards association was not possible. Evidence from focus association provides support for the movement-based theory of control for object control constructions, but not for subject control.

4.7.3 *Tough* constructions

In this section I will look at the English *tough* construction, exemplified in (251). The adjective (here "hard") takes a nonfinite clause which includes a gap. This gap is intuitively interpreted as the matrix subject, thus making the expletive example (252) below a possible paraphrase for (251).

¹⁰⁷As noted in footnote 64, the contrast could potentially also be explained under the movement theory of control (Hornstein, 1999), if the movement chain in subject "control" does not leave a rich lower copy.

(251) **A *tough* construction:**

The reviewers are tough to please ____.

(252) **It's tough to please the reviewers.**

Chomsky (1977) showed that *tough* constructions involve an \bar{A} -chain relating the gap position to the edge of the embedded nonfinite clause, based on evidence from island-sensitivity. Engdahl (1983) later supported this view with the observation that *tough* constructions license parasitic gaps. However, the relationship between this \bar{A} -chain and the subject of the *tough* construction has remained controversial. There are broadly two approaches. Brody (1993); Hornstein (2000) proposes that the subject originates in the gap position (cf 252), undergoes \bar{A} -movement, and then A-moves to its surface position. Chomsky (1977, 1981) proposes that there is \bar{A} -movement of a null operator and the subject is base-generated high. These two views are schematized here:

(253) **Two hypotheses for the subject of *tough* constructions:**a. The movement hypothesis:

[The reviewers] are tough [____ to please ____]

b. The base-generation hypothesis:

The reviewers are tough [Op to please ____]

Association with *even* provides new evidence in favor of the base-generation approach. Consider the following two pairs of expletive constructions and corresponding *tough* constructions. In the expletive constructions (a), *even* can adjoin to the adjective "hard" or to the embedded verb and successfully associate down with the F-marked constituent. When the F-marking is instead in the subject of a corresponding *tough* constructions (b), *even* in the nonfinite embedding is no longer able to associate with the F-marked constituent.

(254) Context: Translation is difficult.

a. It's (*even*) hard to (*even*) translate [_{DP} a [children's]_F book].

b. [_{DP} A [children's]_F book] is (*even*) hard to (**even*) translate ____ . (David Pesetsky, p.c.)

(255) Context: This town is terrible.

a. It's (*even*) hard to (*even*) find [_{DP} a good [sandwich]_F].

b. [_{DP} A good [sandwich]_F] is (*even*) hard to (**even*) find ____ .

Under the movement hypothesis, the pattern of backwards association here is unexpected. Furthermore, because the subject's first movement chain is an \bar{A} -movement step, under Takahashi and

Hulsey's (2009) theory of Wholesale Late Merger, the noun phrases "children's book" and "good sandwich" must originate low within the scope of *even*. The inability for *even* in the nonfinite embedding to associate backwards in the *tough* constructions in (b) is therefore unexpected under the movement view.

In contrast, under the base-generation hypothesis, the lack of backwards association with *even* in the nonfinite clause is easily explained. The DP "the reviewers" and "a good sandwich" did not originate within the scope of *even*, and therefore there will be no F-marked constituent in the scope of these lower *evens* at LF. This new evidence from focus association therefore supports the base-generation hypothesis for the subjects of *tough* constructions.¹⁰⁸

4.8 Summary

The semantic contribution of focus-sensitive operators such as *even* and *only* makes reference to a particular constituent which I have here called the *associate*. The structural relationship between these operators and their associates is in principle unbounded but not completely free—in particular, the associate must be in the scope of the operator. Jackendoff (1972) showed that this requirement is strictly true in the case of *VP-only* and almost always true with *VP-even*. This structural requirement is also what is expected based on the widely-adopted compositional semantics of focus following Rooth (1985, 1992b).

In this chapter I discussed the behavior of sentences where the F-marked associate has moved out of the scope of the focus-sensitive operator, schematized in (167) repeated here as (256) below. I showed that in such cases, we observe a contrast between *even* and *only*: *even* is able to associate "backwards" with F-marking in the restrictor of a DP moved out of *even*'s scope, but *only* is not able to. Furthermore, I showed that the lower copy of the DP's restrictor which includes the F-marking is necessary for this backwards association. It is not simply the case that *even* is somehow able to associate with material which was never in its scope.

(256) *Even but not only can associate with material which has moved out:* (=167)

$$\underbrace{[\text{DP } \dots \alpha_F \dots] \dots [\check{\text{even}}/*\text{only} [\dots \text{---} \dots]]}$$

The pattern in (256) was observed with standard cases of A-movement, including movement of subjects under the VP-internal subject hypothesis and subject raising constructions, as well as common cases of \bar{A} -movement, including topicalization and *wh*-movement, as well as the covert movement operation of QR. Most of these empirical observations are new, although the difference

¹⁰⁸To explain the availability of backwards association with the higher *evens*, a short movement step of the subject, for example from a matrix predicate-internal subject position, must be posited.

between *even* and *only* in associating with a leftward subject was noted in Jackendoff (1972) and has remained a puzzle since then. As such, the proposal here acts as the first principled proposal for Jackendoff's observation.

My proposal adopts the Copy Theory of movement (Chomsky, 1993) and associated work on the interpretation of movement chains (Sauerland, 1998; Fox, 1999). I proposed that in cases of apparent backwards association, the operator is associating with the F-marking in the lower copy of the moved DP, within its scope. While the grammar is able to generate this configuration, different focus-sensitive operators vary in their ability to associate in this way. This is due to the different semantic contributions of the operators rather than a lexical property that they inherently possess.

In this chapter I showed that this backwards association configuration is possible with *even* because *even* uses focus-alternatives to introduce a non-assertive inference, without affecting the truth conditions of the utterance. On the other hand, *only* uses focus alternatives to introduce new truth conditions. In the backwards association configuration, the moved DP will be interpreted both outside and inside of the scope of *only*, with the focused constituent substituted with alternatives in the lower copy but not in the higher one. This leads to a semantically ill-formed object; in particular, either the presuppositions of the lower copy of movement will not be satisfied, or the configuration will necessarily be always true or always false, leading to ungrammaticality (Gajewski, 2002).

This proposal makes a more general prediction regarding the behavior of focus-sensitive operators in backwards association configurations: the grammar is able to generate configurations with copies of focus, as we have seen with *even* and *only*. Operators will vary in their ability to interpret backwards association structures depending on their semantic contribution: if the meaning generated by the focus-sensitive operator using the focus alternatives composes above with the higher copy of movement, we predict this configuration to be uninterpretable and therefore that the operator will not be able to associate backwards. If, on the other hand, the operator uses the focus alternatives to introduce a projective meaning, it is predicted to be able to associate backwards.

With this basic proposal in place, I will next move to the consideration of backwards association configurations with F-marking on different parts of the DP, for example on the determiner and on the whole DP. These configurations will pattern differently from the configuration studied in this chapter, leading to a more nuanced set of considerations for determining whether a particular focus-sensitive operator can associate backwards or not.

Chapter 5

Varying the associate

In my investigation thus far, I have considered configurations where F-marking is on part of the *restrictor* of a DP which moves out of the scope of a focus-sensitive operator. In these cases, *even* but not *only* is able to associate backwards with the F-marked constituent. In this chapter I will extend the discussion to other configurations: F-marking on the determiner of the moved DP (257a) and F-marking on the entire moved DP (257b). We will see that in both of these cases, backwards association is impossible with both *even* and *only*. I will show that both of these results are predicted by my proposal.

(257) Both *even* and *only* cannot associate backwards with determiners and entire DPs:

- a. * $\underbrace{[\text{DP } D_F \text{ NP}] \dots \text{ even/only } [\dots \text{ } _ \text{ } \dots]}_{\text{scope}}$
- b. * $\underbrace{\text{DP}_F \dots \text{ even/only } [\dots \text{ } _ \text{ } \dots]}_{\text{scope}}$

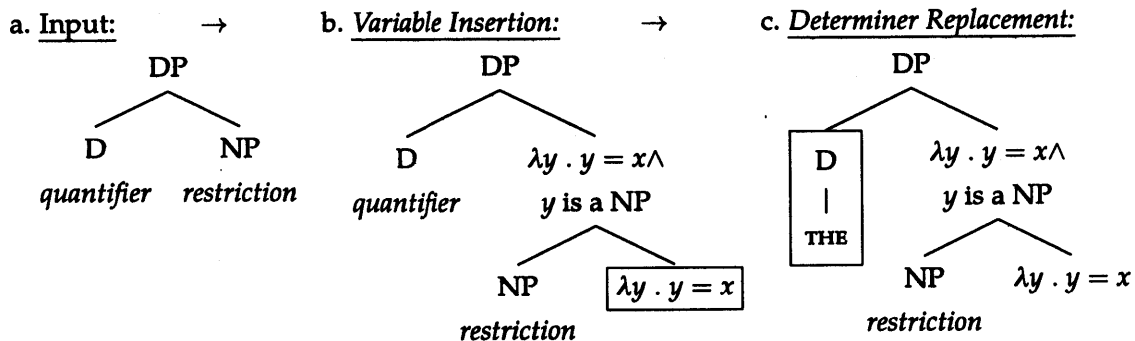
I discuss backwards association with determiners in section 5.1 and then turn to backwards association with the entire DP in section 5.2. Afterwards, in section 5.3 I briefly discuss implications for the analysis of different kinds of DPs, including proper names and those with numerals and demonstratives. Finally in section 5.4 I will discuss backwards focus association with *wh*-phrases.

5.1 Association with quantifiers

In chapter 4, I proposed that when *even* associates “backwards” with F-marked material which has moved out of its scope, it is actually associating with a lower copy of the F-marked constituent within its scope. This was made possible by adopting the Copy Theory of movement, the view that movement leaves full copies in the base positions of movement. At LF, these lower copies are

converted into bound definite descriptions through the process of Trace Conversion (Rullmann and Beck, 1998; Fox, 2002). The individual steps involved in modifying the lower copy using Trace Conversion are repeated here from (53) above.¹⁰⁹

(258) Trace Conversion of the lower copy: (=53)



An important detail of this Trace Conversion process is that the quantificational material in the lower copy is replaced in the Determiner Replacement step (258c). Consider an example where this determiner bears an F-index. Suppose that in the process of replacing the determiner, its indices and featural content will also be replaced, destroying the F-index. We then make the prediction that, if this DP moved out of the scope of a focus-sensitive operator, backwards association with this determiner will not be possible, regardless of the choice of focus-sensitive operator.¹¹⁰ The following contrast, using *even*, appears to bear out this prediction:

(259) *Even* cannot associate backwards with a determiner:

- a. Of course we arrested some protesters. ✓We *even* arrested [every]_F protester.
 b. Of course some protesters were arrested. *[?][Every]_F protester was *even* arrested.¹¹¹

In example (259a) the F-marked determiner “every” is in the surface scope of *even*, making this association possible. Example (259b) has been minimally modified by passivization, making “every protester” move to Spec,TP, out of the surface scope of *even*. *Even* associating with “every” in this variant is judged as degraded. We can confirm that it is not generally the case that *even*

¹⁰⁹As noted in footnote 22, the order of the two subprocesses, Variable Insertion and Determiner Replacement, is to my knowledge not strict. That is, Determiner Replacement could just as well apply before Variable Insertion.

¹¹⁰The same would hold under the more common Jackendoff (1972) view of F-marking as a syntactic feature [F] assigned to nodes in the narrow syntax, instead of F-marking being an index as it is here.

¹¹¹There seems to be some variation in how marked (259b) is with the intended interpretation, although all speakers consulted report a clear degradation in (259b) as compared to (259a).

Counterexamples to this generalization can be constructed by adding another scope-bearing element, such as a modal. Such cases will be discussed in section 7.2.

cannot associate with F-marked determiners of derived subjects, through the following example. Example (260b) has an adverb *even* in a higher clause, successfully associating with the “every” below. What makes association possible in (260), but not (259), is that “every” is in the scope of *even* in (260), but not in (259).

(260) It *even* is the case that [every]_F protester was arrested.

I also note in passing that the same contrast holds of association with F-marked simplex quantificational DPs, such as “everyone,” in (261). However, the reason why this backwards association is not possible is different, and will be presented in section 5.2.

(261) Of course some protesters were arrested. *[Everyone]_F was *even* arrested.

Let’s consider the derivation of example (259b). The determiner “every” in the subject DP bears the F-index F1. This DP is copied in the narrow syntax (262a), resulting in a representation with two nodes with index F1, one inside the scope of *even* and one outside. When Determiner Replacement replaces this node at LF, the F-marking on the lower copy of the DP is lost with it (262b), and therefore the denotation of *vP*, the complement of *even*, will not be focus-assignment-dependent (262c). Note that here, following the discussion in section 4.2, I do not consider the option of syntactic reconstruction; I will discuss examples of this form and the possibility of syntactic reconstruction later in section 7.2.

(262) Hypothetical derivation of (259b):¹¹²

- a. Narrow syntax: [[Every]_{F1} protester]_{F1} PAST *even* [[every]_{F1} protester]_{F1} BE arrested
- b. LF after TC: [[Every]_F protester] λ1 PAST *even* [_{vP} [the protester ①] BE arrested]
- c. [_{vP}] (*h* || *g*) = λ*w*_s : *g*(1) is a protester in *w* . *g*(1) is arrested in *w*

For concreteness, consider the set of relevant focus assignment functions *H*. I assume that the relevant alternative to “every” is the determiner “some,” which is of the same semantic type. The definition for focus assignments in section 2.1 gives us the following set of focus assignments:

(263) $H = \{h_0, h_1\}$

- a. $h_0(F1) = \llbracket \text{every} \rrbracket = \lambda g_a . \lambda w_s . \lambda P_{\langle e,t \rangle} . \lambda Q_{\langle e,t \rangle} . \forall x (P(x) \rightarrow Q(x))$
- b. $h_1(F1) = \llbracket \text{some} \rrbracket = \lambda g_a . \lambda w_s . \lambda P_{\langle e,t \rangle} . \lambda Q_{\langle e,t \rangle} . \exists x (P(x) \wedge Q(x))$

¹¹²Here I illustrate the movement of “every protester” from the edge of *vP* to the Spec,TP position, without the initial passivization step from the complement position of the verb. The position of the lower copy is not crucial for the illustration here.

Given an ordinary assignment function g , we can compute $\llbracket vP \rrbracket (h \parallel g)$ for each $h \in H$. Because the value of $\llbracket vP \rrbracket$ is not dependent on the focus assignment h , however, $\llbracket vP \rrbracket (h_0 \parallel g) = \llbracket vP \rrbracket (h_1 \parallel g)$. Therefore it can never be the case that $\llbracket vP \rrbracket (h_0 \parallel g)$ and $\llbracket vP \rrbracket (h_1 \parallel g)$ will differ in terms of relative likelihood, explaining why *even* cannot be used felicitously in this structure.

The derivation in (262) therefore predicts the inability of *even* to associate with the determiner heading a DP which is interpreted outside of the scope of *even*. The corresponding lower copy of movement will not contain the F-marked node necessary for the intended focus association. With no F-marking, the complement of *even* will not be focus-assignment-dependent, violating the Focus Assignment Dependency Requirement (43).

A possible alternative explanation of the badness of (259b) might be that *even* is unable to associate with a determiner which is part of a DP that has moved out of *even*'s scope. However, this is not the case. In example (264) below, *even* is able to associate backwards with the quantifier "every" inside the restriction of the moved DP. This is precisely what is predicted by the explanation presented here. Determiner Replacement replaces the head of the moved DP with a definite determiner, in order to correctly interpret the movement chain. Additional determiners inside the restrictor of the moved DP are left untouched. Therefore, the lower copy of the moved DP in (264) will include the F-marked "every," making it possible for *even* to associate with this determiner.

(264) *Even can associate backwards with a determiner in the restrictor:* (David Pesetsky, p.c.)

Context: For Sue, the problem isn't just a worry that some students won't respond.

✓[Hearing from [every]_F student] wouldn't *even* solve the problem.

The inability to associate backwards with an F-marked determiner is predicted for all focus-sensitive operators, not just *even*. Because the F-marking on the determiner will be lost in the process of Determiner Replacement, backwards association should be impossible with all focus-sensitive operators. We see in (265) below that *only* also cannot associate backwards with the F-marked determiner of a DP. However, this is not terribly surprising given that *only* is not able to associate backwards with F-marking in the restrictor either, as we have seen before.

(265) *Only cannot associate backwards with a determiner:*

- a. There are rumors that we arrested every protester. ✓Actually, we *only* arrested [some]_F protesters.
- b. There are rumors that every protester was arrested. *Actually, [some]_F protesters were *only* arrested.

The interesting contrast is instead in the behavior of *even*. Recall that it is possible for *even* to associate with the moved DP's restrictor (or a part thereof), as we saw in previous sections, but

not with the determiner heading the DP. We can make this contrast even sharper, for example through the minimal pair in (266) below. Backwards association with the restrictor of a moved DP is possible, but backwards association with the determiner heading the DP—in an identical structural configuration—is not.

(266) **Backwards association possible with a restrictor but impossible with the determiner:**

- a. ✓ Every [protester]_F was *even* arrested.
 b. * [Every]_F protestor was *even* arrested. (=259b)

Note that the contrast in (266) cannot be explained by a view where *even* takes scope in a position higher than indicated by its surface position. Consider a view where *even* is able to move covertly and extend its scope, and this makes apparent backwards association possible. The availability of backwards association in (266a) would then tell us that *even* covertly moves over the subject “every protestor” (267a). However, given that association with determiners is generally possible, as shown in (264), this LF configuration would then predict backwards association with “every” to be possible in (267b). Because “every protestor” is a constituent, it is not possible for *even* to move to take “protestor” in its scope, but not “every.”

(267) **Hypothetical scope theory LFs for (266):**

- a. *even* [DP every [protester]_F] was arrested.
 ⇒ predicts backwards association to be possible
- b. *even* [DP [every]_F protestor] was arrested.
 ⇒ predicts backwards association to be possible, contrary to fact ☹

This movement of *even* is in fact what is proposed by one version of the so-called scope theory of *even*, which will be discussed in detail in chapter 6. The contrast here will act as an important argument against using the scope theory of *even* as an explanation for backwards association.

In contrast to this scope theory approach, my proposal predicts the inability of *even* to associate with a determiner of a DP that has moved out of *even*'s scope. Apparent backwards association happens when *even* associates with F-marking in the lower copy of a movement chain, in its scope. If the DP's quantificational material is interpreted outside of *even*, the process of Trace Conversion will destroy the determiner in the lower copy of the DP, inside of *even*, making it impossible to associate with the intended F-marked constituent.

5.2 Association with the moved DP

In this section I will show that *even* and *only* are unable to associate backwards with an entire DP which is F-marked. This too is predicted by the semantics of *even* and *only*, although the explanation will be different from what I proposed above for the inability of *even* to associate backwards with determiners.

Example (268) below is a case of attempted backwards association with a DP. Here the context makes it clear that the preajcent DP “every student” contrasts with “two-thirds of the professors.”

(268) **New degree requirements:** (based on 200)

Context: David wants to change the requirements for the Linguistics degree program. The proposal will pass if either (a) at least two-thirds of the professors vote for the change or (b) every student votes for the change. *David knows that it's very unlikely that all the students will be happy with the changes; therefore he is hoping that he can pass the change by getting two-thirds of the faculty on board.*

The votes have been counted and the change was successful. Over two-thirds of the professors voted for the change. But it turns out that the proposal would have passed anyway, even if more of the faculty had voted against it. That's because, surprisingly,

* [Every student]_F (had) *even* voted for the change.¹¹³

One confound that makes constructing and testing such examples difficult is that, due to the process of focus projection—the PF mapping between F-marking and prosodic prominence placed on a subpart of the F-marked constituent—a variant with F-marking on the restrictor “student” is predicted to be pronounced in a similar—if not identical—manner. Such a sentence is indeed grammatical, albeit in a different context which highlights that it is the restrictor, not the entire DP, that is contrastive:

(269) Context: Every professor voted for the change.

✓ Every [student]_F (had) *even* voted for the change.

This confound can be avoided by testing simplex quantificational DPs, such as “everyone.” Example (261) from the previous section shows that backwards association of *even* with “everyone” is also not possible:

(270) ***Even* can't associate backwards with an entire DP:**

Of course some protesters were arrested. *[Everyone]_F was *even* arrested. (=261)

¹¹³The perfect auxiliary “had” is used here to make sure that we are testing the adverb *even*, rather than some kind of post-nominal *even*. This sentence seems to be ungrammatical with the intended association both with and without “had.”


In the remainder of this section I will show how my proposal for the semantics of *even* predicts that this pattern of association is not possible. Before doing so, though, it's worth noting that there are a number of apparent counterexamples to this claim that backwards association with an entire F-marked DP is not possible. Two such examples are given in (271). I will return to such cases in section 5.3, after I present the prediction of my proposal.

(271) **Some apparent counterexamples:**

- a. ✓ [Mary]_F, John *even* saw ___ at the party. (=4b)
 b. I loved that rice. ✓ [This corn]_F, I *even* loved ___ . (Chris Tancredi, p.c.)

Here for concreteness I will first present the derivation for (268) and demonstrate how the use of *even* in this configuration is never licensed, and then discuss how this general problem will always occur in cases of backwards association with an entire F-marked DP.

(272) **Expected derivation for (268):**

- a. Narrow syntax: [_{DP,F1} Every student] PAST *even* [_{vP} [_{DP,F1} every student] vote for the change]

 b. LF: [_{DP,F1} Every student] λ1 PAST *even* [_{vP} [_{DP,F1} the student ①] vote for the change]

Let's look at the semantics for the DP in the scope of *even*, represented by "the student ①" above. The preajacent focus assignment h_0 is defined so that $h_0(F1)$ will return the semantic denotation of the node with index F1, as computed from its constituent parts. Here there are two nodes in the workspace with index F1—one outside of the scope of *even* and one inside. For the purposes of computing the scalar inference of *even*, I will assume that only the nodes in its complement are visible for the determination of H , and therefore $h_0(F1) = \llbracket \text{the student } \textcircled{1} \rrbracket$. $h_0(F1)$ would therefore be of extensional type e (superintensional type $\langle a, \langle s, e \rangle \rangle$).

(273) $h_0(F1) = \llbracket \text{the student } \textcircled{1} \rrbracket = \lambda g_a . \lambda w_s : g(1) \text{ is a student} . g(1)$

Note that the denotation of the higher copy of DP cannot generally be used to define h_0 in this situation, as it will immediately lead to a problem in interpreting the movement chain in the assertion. If $h_0(F1) = \llbracket \text{every student} \rrbracket$, the truth condition for $\llbracket \text{even } vP \rrbracket (h_0 \parallel g)$ will be $\llbracket vP \rrbracket (h_0 \parallel g) = (\lambda w_s . \text{every student votes for the change in } w)$. This has effectively undone the Trace Conversion operation applied to the lower copy. For this example, where what is moved is a quantificational subject, this predicts that the subject will obligatorily have narrow scope. For examples where the DP originates in a non-subject position, undoing Trace Conversion in this way will force the lower copy of the DP to have the higher, quantificational type (here, extensional type $\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$),

which will be unable to compose with its sister for simple type reasons. See section 6.2 on QR as type-driven movement and the necessity of Trace Conversion on the lower copy.

The question now is what the other elements in the set of focus assignment functions will be. Recall that, for each F-index, all the focus assignments must return an object with the same semantic type. This is a definitional property of the set H in my system, repeated below in (274).

(274) **Focus Assignment Type Uniformity:** (=34)

For any F-index $Fi \in \mathbb{N}_F$ and any two $h', h'' \in H$, the set of relevant focus assignment functions, either $\text{typeof}(h'(Fi)) = \text{typeof}(h''(Fi))$ or both $h'(Fi)$ and $h''(Fi)$ are undefined.

In this case, because the prejacent $h_0(F1)$ has extensional type e , the alternatives for F1 must be other individuals of extensional type e . The denotations of the contextually relevant DPs “every student” and “two-thirds of the professors,” which both have extensional type $\langle\langle e, t \rangle, t\rangle$, therefore cannot be in H . If H is the singleton set $\{h_0\}$, we are done, as the complement of *even* will not vary with different choices of focus assignments in H , violating the Focus Assignment Dependency Requirement (43).

Suppose the relevant domain of individuals include John, Mary, and Bill, who are students, and Irene, Kai, and Martin, who are professors. I will assume for demonstration here that the alternative focus assignments in H map F1 to John, Mary, Bill, Irene, Kai, and Martin. The inclusion of both students and professors—and more generally, the makeup of these individuals who are contextually relevant—is not crucial for the argument here. I will return to this assumption later.

(275) $H = \{h_0, h_1, h_2, h_3, h_4, h_5, h_6\}$

- a. $h_0(F1) = \llbracket \text{the student } \textcircled{1} \rrbracket = \lambda g_a . \lambda w_s . g(1) \text{ is a student} . g(1)$ (=273)
- b. $h_1(F1) = \llbracket \text{John} \rrbracket = \lambda g_a . \lambda w_s . \text{John}$
- c. $h_2(F1) = \llbracket \text{Mary} \rrbracket = \lambda g_a . \lambda w_s . \text{Mary}$
- d. $h_3(F1) = \llbracket \text{Bill} \rrbracket = \lambda g_a . \lambda w_s . \text{Bill}$
- e. $h_4(F1) = \llbracket \text{Irene} \rrbracket = \lambda g_a . \lambda w_s . \text{Irene}$
- f. $h_5(F1) = \llbracket \text{Kai} \rrbracket = \lambda g_a . \lambda w_s . \text{Kai}$
- g. $h_6(F1) = \llbracket \text{Martin} \rrbracket = \lambda g_a . \lambda w_s . \text{Martin}$

At this point we can start to see what the problem will be. The prejacent value for the node with index F1 is the ordinary-assignment-dependent individual $\textcircled{1}$. However, recall that the index 1 is not bound in the scope of *even*, and therefore the semantics of *even* will quantify generically over the domain of this quantifier (chapter 3). Because *even* does not allow a “bound” reading when projecting through quantification, we have no information about what individual the original

prejacent corresponding to the DP is, and therefore we will not be able to construct a meaningful scalar inference.

Let's compute the other ingredients necessary for computing the inference of *even*. This includes the denotation of the complement of *even*, vP , and the set G of relevant ordinary assignment values. G is used to quantify over the variable with index 1, which is unbound in the scope of *even*. I assume here that G includes the assignment functions which map the index 1 to the same set of relevant individuals from the discussion above: the three students, John, Mary, and Bill, and the three professors, Irene, Kai, and Martin.

$$(276) \quad \llbracket vP \rrbracket (h \parallel g) = \lambda w . h(F1)(g)(w) \text{ voted for the change in } w$$

$$(277) \quad G = \{ \{1 \mapsto x\} \mid x \in D_e \}$$

$$a. \quad g_1(1) = \text{John}$$

$$b. \quad g_2(1) = \text{Mary}$$

$$c. \quad g_3(1) = \text{Bill}$$

$$d. \quad g_4(1) = \text{Irene}$$

$$e. \quad g_5(1) = \text{Kai}$$

$$f. \quad g_6(1) = \text{Martin}$$

(Note there is no g_0 .)

Now consider the expected scalar inference of *even*:

(278) **Predicted scalar inference of *even*:**

$$\begin{aligned} & \sim \text{GEN}(g \in G, h \in H) (h \neq h_0 \rightarrow \llbracket vP \rrbracket (h_0 \parallel g) <_{\text{likely}} \llbracket vP \rrbracket (h \parallel g)) && \text{(by definition)} \\ & \iff \text{GEN} \left(\begin{array}{l} g \in \{g_1, g_2, g_3, g_4, g_5, g_6\}, \\ h \in \{h_1, h_2, h_3, h_4, h_5, h_6\} \end{array} \right) \left(\begin{array}{l} (\lambda w . h_0(F1)(g)(w) \text{ voted for the change in } w) <_{\text{likely}} \\ (\lambda w . h(F1)(g)(w) \text{ voted for the change in } w) \end{array} \right) \\ & \iff \text{GEN} \left(\begin{array}{l} g \in \{g_1, g_2, g_3, g_4, g_5, g_6\}, \\ h \in \{h_1, h_2, h_3, h_4, h_5, h_6\} \end{array} \right) \left(\begin{array}{l} \overbrace{(\lambda w . g(1) \text{ voted for the change in } w)}^{\text{prejacent proposition}} <_{\text{likely}} \\ \underbrace{(\lambda w . h(F1)(g)(w) \text{ voted for the change in } w)}_{\text{alternative proposition}} \end{array} \right) \end{aligned}$$

Now let's consider what the prejacent and alternative propositions will actually range over:

(279) **Prejacent propositions in (278) for different ordinary assignments g :**

$$a. \quad g_1 \Rightarrow (\lambda w . \text{John voted for the change in } w)$$

$$b. \quad g_2 \Rightarrow (\lambda w . \text{Mary voted for the change in } w)$$

$$c. \quad g_3 \Rightarrow (\lambda w . \text{Bill voted for the change in } w)$$

- d. $g_1 \Rightarrow (\lambda w . \text{Irene voted for the change in } w)$
- e. $g_2 \Rightarrow (\lambda w . \text{Kai voted for the change in } w)$
- f. $g_3 \Rightarrow (\lambda w . \text{Martin voted for the change in } w)$

(280) **Alternative propositions in (278) for different focus assignments h :**

- a. $h_1 \Rightarrow (\lambda w . \text{John voted for the change in } w)$
- b. $h_2 \Rightarrow (\lambda w . \text{Mary voted for the change in } w)$
- c. $h_3 \Rightarrow (\lambda w . \text{Bill voted for the change in } w)$
- d. $h_4 \Rightarrow (\lambda w . \text{Irene voted for the change in } w)$
- e. $h_5 \Rightarrow (\lambda w . \text{Kai voted for the change in } w)$
- f. $h_6 \Rightarrow (\lambda w . \text{Martin voted for the change in } w)$

The set of prejacent propositions computed by quantification over the ordinary assignment g (279) is predicted to be identical to the set of alternative propositions computed by quantification over the focus assignment h (280). The scalar inference of *even* requires that, in general, a proposition in the former set is less likely than a proposition in the latter set. Clearly, if we cross the propositions pairwise, the resulting inequalities in the domain of generic quantification will be symmetric. Assuming that generic quantification requires that its nuclear scope hold of at least half of its domain of quantification, we predict that this scalar inference of *even* can never be true.

The problem observed here will occur generally in this configuration of attempted backwards association with an F-marked DP that is moved out of the scope of *even*. This problem occurs in precisely this configuration because the DP node whose denotation is *ordinary* assignment dependent, due to it being a Trace Converted lower copy of movement, is also a node whose denotation is *focus* assignment dependent, due to it being what is F-marked. *Even* will quantify over different ordinary assignments for this node for the prejacent, ranging over the relevant set of individuals, and also quantify over possible alternatives to an *e*-type individual in the context, which will also range over this set of individuals. The membership of this set of relevant individuals does not actually matter—for example, G and H above could have been limited to functions returning only different students for $F1$, or only different professors, etc. The scalar inference computed for *even* in this configuration will be a logical contradiction, making such use of *even* not felicitous with the intended association.

Backwards association with an F-marked DP is also impossible with *only*. Consider the following example, where we contrast the DPs “every student” and “at least two-thirds of the professors.”¹¹⁴

¹¹⁴With *even* above, I tested backwards association with a DP using the simplex quantifier “everyone” (261) in addition to a context contrasting “every student” with “at least two-thirds of the professors” (268). A parallel test for backwards

(281) **New degree requirements, with *only*:** (based on 268)

Context: David wants to change the requirements for the Linguistics degree program. The proposal will pass if *both* (a) at least two-thirds of the professors vote for the change *and* (b) every student votes for the change.

The votes have been counted and the proposal did not pass, because less than two-thirds of the faculty voted for the change.

* [Every student]_F (had) *only* voted for the change.¹¹⁵

The inability of *only* to associate backwards with a whole DP is less surprising than with *even*, as *only* is not able to associate backwards with F-marking in the restrictor, either. However, it's worth showing how this is predicted under my account. We begin with the LF for (281) below:

(282) **Expected derivation for (281):**

a. **Narrow syntax:** [_{DP,F1} Every student]_{PAST} *only* [_{DP,F1} every student]_{vP} vote for the change]

b. **LF:** [_{DP,F1} Every student] λ1 PAST *only* [_{DP,F1} the student ①]_{vP} vote for the change]

I will assume again that the preadjacent focus assignment h_0 will be determined based on the denotation of the lower DP with focus index F1, "the student ①," which has extensional type e . Other relevant focus assignments will be constant functions returning individuals, as in (275) above. The association with *only* would be as in (i) below, which is indeed ungrammatical. However, given the independent unacceptability of the baselines in (ii), this evidence is not conclusive.

(i) There are rumors that everyone was arrested at the rally yesterday. *In reality, [someone]_F was *only* arrested.

(ii) **Baseline:** There are rumors that everyone was arrested at the rally yesterday.

a. * In reality, *only* [someone]_F was arrested.

b. * In reality, it was *only* the case that [someone]_F was arrested.

The same goes for "noone":

(iii) There are rumors that someone was arrested at the rally yesterday. *In reality, [noone]_F was *only* arrested.

(iv) **Baseline:** There are rumors that someone was arrested at the rally yesterday.

a. * In reality, *only* [noone]_F was arrested.

b. * In reality, it was *only* the case that [noone]_F was arrested.

Such examples instead suggests that "someone" and "noone" generally resist being the associate of *only*. I do not know why *only* cannot associate with these simplex quantifiers, even within its scope.

¹¹⁵The perfect auxiliary "had" is important to make sure we are testing the adverb *only* rather than a post-nominal *only*. See also footnote 113.

denotation of the complement vP will also be the same as above, in (276). Our semantics for *only* (42) gives us the following denotation for “*only vP*”:¹¹⁶

$$\begin{aligned}
 (283) \quad \llbracket \textit{only vP} \rrbracket (g) & \\
 &= \lambda g_a . \lambda w . \forall h \in H (h \neq h_0 \rightarrow \llbracket vP \rrbracket (h \parallel g)(w) \text{ is false}) \\
 &= \lambda g_a . \lambda w . \forall h \in \{h_1, h_2, h_3, h_4, h_5, h_6\} (h(F1)(g)(w) \text{ did not vote for the change in } w) \\
 &= \lambda g_a . \lambda w . \text{John, Mary, Bill, Irene, Kai, and Martin all didn't vote for the change in } w
 \end{aligned}$$

There are two problems with this result. First, note that the resulting denotation for “*only vP*” is not dependent on the ordinary assignment function g that it takes. In particular, because $\llbracket \textit{only vP} \rrbracket$ is not dependent on the ordinary assignment’s value for index 1, the quantificational higher copy and its λ -binder with index 1 will not contribute to the interpretation of the sentence. The semantic interpretation of (281) will therefore include vacuous binding.¹¹⁷

Second, consider the prejacent presupposition introduced by *only* in (281). The prejacent presupposition of $\llbracket \textit{only vP} \rrbracket$ is that, for an ordinary assignment g and world w , “ $g(1)$ *did* vote for the change in w .” This presupposition will project universally through the quantificational subject, leading to the presupposition that every student voted for the change. This presupposition is incompatible with the assertion derived from (283). I believe that backwards association configurations as in (282) generally run into this type of problems when considering the prejacent presupposition and its projection behavior, making the assertion either a contradiction or a tautology and therefore not a useable meaning, but it is difficult to prove that this issue occurs in general terms. I will leave a more general explanation for why this configuration is not possible for future research.

In this section I showed that both *even* and *only* are unable to associate backwards with an F-marked DP which has moved out of its scope. F-marked DPs, then, pattern together with F-marked determiners from the last section in not allowing backwards focus association. It’s important to

¹¹⁶For simplicity, I will not illustrate the prejacent presupposition of *only* here. The contribution of the prejacent will be discussed below.

¹¹⁷The vacuous binding issue could be avoided by adding a bound variable, as in (i). Note that backwards association is not improved in (i), making the second reason why this configuration is out—based on the interaction of *only*’s assertion and prejacent presupposition—a more general reason for the ill-formedness of this configuration.

- (i) Context: The all-female Linguistics department is voting on this year’s MVP. Due to previous incidents where the vast majority of students voted for themselves, but none of the faculty did, the department has adopted the following rule: no MVP will be named if *both* (a) every student votes for herself *and* (b) no professor votes for herself. It looks like there will be an MVP this year because only one of these conditions for nullification was satisfied:

* [Every student]_{F,i} (had) *only* voted for herself_i.

note that the reasons why backwards association is disallowed in these cases are not the same. Backwards association with the determiner is predicted to be impossible for *all* focus-sensitive operators, as the determiner is replaced in the process of Trace Conversion. On the other hand, backwards association with a DP is not possible with *even* and *only* due to the operators' semantics. An example of a focus-sensitive operator which *does* allow backwards association with an entire F-marked DP (*also*) will be discussed in the appendix to this chapter.

5.3 Backwards association and DP structure

In the previous sections I've shown that when a DP moves out of the scope of *even* and *only*, the operator is unable to associate "backwards" with the entire DP or with the DP's determiner. This behavior contrasts with cases of F-marking in the *restrictor* of a moved DP, where we know that *even*, but not *only*, is able to associate backwards. These patterns are schematized below:

- (284) *Even* can associate backwards with the restrictor, but not with the head D or entire DP:
- a. $\underbrace{[\text{DP D } \dots \alpha_F \dots]} \dots \checkmark \textit{even} / * \textit{only} [\dots _ \dots]$ chapter 4
 - b. $* \underbrace{[\text{DP D}_F \text{ NP}]} \dots \textit{even} / \textit{only} [\dots _ \dots]$ section 5.1
 - c. $* \text{DP}_F \dots \textit{even} / \textit{only} [\dots _ \dots]$ section 5.2

In this section I will extend the discussion to additional varieties of DPs and use backwards association with *even* as a diagnostic for the DP's structure. Specifically, because backwards association with *even* is only possible with F-marking in the restrictor of a moved DP, successful backwards association will show that the F-marked constituent is properly contained in the complement of the DP's head.

I begin with proper names. As we have seen previously, *even* is able to associate backwards with a proper name which has moved out of *even*'s scope:

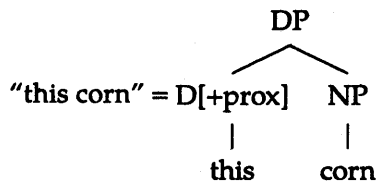
- (285) *Even* can associate with a proper name moved out of its scope, but not *only*: (=166)
- a. \checkmark [Mary]_F, John *even* saw $_$ at the party.
 - b. $*$ [Mary]_F, John *only* saw $_$ at the party. (based on Tancredi, 1990, ex. 57b)

If proper names are themselves maximal DP categories, the behavior of proper names in (285) would be at odds with the results of section 5.2. Note in particular that the problem that occurs in such configurations (in section 5.2) will occur regardless of the content of the DP in question.

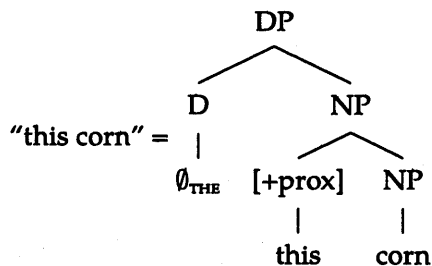
(288) **Backwards association with demonstratives and demonstrative DPs:**

- a. [These]_F protesters were *even* arrested. (David Pesetsky, p.c.)
 b. I loved that rice. [This corn]_F, I *even* loved. (Chris Tancredi, p.c.)

One simple hypothesis would be to analyze demonstratives as D heads, as schematized in (289) below. Under this hypothesis, the results in (288) above would be surprising. Example (288a) would be a case of backwards association with a D head, which is predicted to be ungrammatical under my analysis (section 5.1); and similarly example (288b) would be a case of backwards association with an entire DP, also predicted to be ungrammatical (section 5.2).

(289) **Demonstratives as D heads: makes false predictions**

The evidence from (288) above thus supports a view where demonstratives are actually a part of the restrictor of the DP, rather than the D head itself, as argued for example by Dryer (1992) and in subsequent research. The DP would then be headed by a null variant of the definite determiner, as with proper names above, which is not included in the F-marking in the examples in (288). An example of such a structure is in (290) below.

(290) **Demonstratives as modifiers:**

Again, cross-linguistic evidence suggests that demonstratives are independent of the definite determiner. For example, in Hebrew demonstratives are post-nominal, patterning with adjectives rather than quantifiers. Definite DPs in Hebrew show the definite marker *ha-* on the head noun and all adjectives, which we observe on the demonstrative as well in example (291) below. See Dryer (1992) for additional cross-linguistic evidence of this form.

(291) **Definite marking independent of demonstratives:**

Ha-teza ha-zot lo ra'a.
 DEF-thesis DEF-this_f not bad_f

'This thesis is not bad.'

Hebrew (Hadas Kotek, p.c.)

A similar argument can be made with numerals. Example (292) below shows that *even* can associate backwards with a simplex numeral in a DP without an overt determiner. This shows that the numeral must be a modifier in the restrictor of the DP, rather than a D head heading the DP itself.

(292) **Backwards association with a numeral:**

Of course ten protesters were arrested. ✓[Fifty]_F protesters were *even* arrested.

In this chapter I discussed the structure of proper names and DPs with demonstratives and numerals in English. The fact that *even* can only associate backwards with the restrictor of a moved DP, but not with the D head or the entire DP, can be used to clarify the structure of these DPs. In all of these cases, the availability of backwards association with *even* shows that these elements are in the restrictor of a DP headed by a null definite determiner. In the next section I will continue this investigation by looking at a particular family of DPs, *wh*-phrases.

5.4 Association with *wh*-phrases

In this section I will investigate the possibility of backwards association with *wh*-phrases. I begin with overtly moved *wh*-phrases in simplex *wh*-questions. As noted by Tancredi (1990), the *wh*-questions in (293) are ungrammatical with the intended interpretations where *only* associates with the *wh*-word.¹¹⁸

(293) ***Only* cannot associate with a *wh*-word moved above it:**

- a. * [Who]_F did you *only* meet ___? (based on Tancredi, 1990, ex. 57ci)
 Intended: Who *x* is such that you met only *x*?
- b. * [Which boy]_F did you *only* meet ___?
 Intended: Which boy *x* is such that you met only *x*?
- c. * [Which]_F boy did you *only* meet ___?
 Intended: Which boy *x* is such that you met only that boy *x*?

The corresponding echo questions in (294) are grammatical with the intended interpretation, suggesting that the differing positions of the *wh*-words—in the c-command domain of *only* in (294) but not in (293)—are indeed the source of this contrast.¹¹⁹ The echo question intonation itself involves pitch accent on the echo-*wh*-replaced phrase,¹²⁰ see Bartels (1997, chapter 6) for more information on echo *wh*-question intonation.

- (294) A: I went to a party with a bunch of senators, but I *only* met [Jim Montague]_F.
 B1: ✓ You *only* met [who]_F? (based on Tancredi, 1990, ex. 57cii)
 B2: ✓ You *only* met [which senator]_F?

One factor potentially contributing to the ungrammaticality of the examples in (293) is the fact that constituent questions are interpreted as requests for exhaustive answers (Groenendijk and Stokhof, 1984, a.o.). That is, the question “Who did you meet?” expects as an answer the list of all the individuals satisfying the predicate ($\lambda x . \text{you met } x$), i.e. the only individuals that satisfy ($\lambda x . \text{you met } x$). We might wonder whether the addition of *only* in (293) is then redundant. If so, the degraded status of (293) could be due to blocking by the simpler *only*-less variants of the questions.

This hypothesis can be tested by modifying the question so that the addition of *only* is not redundant, for example by the addition of a bound variable.¹²¹ Consider example (295), which is also judged as ungrammatical with the intended association of *only* with the object of “meet.” We would expect this question, if it were grammatical, to have a meaning quite different from its variant with-

¹¹⁸The intended meaning here is a reasonable one. A linguist who prefers to remain anonymous reports a conversation with his child where this type of structure was constructed by the child (approximately age 3).

- (i) Context: Linguist pulls child’s hat low over the child’s eyes.
 Child: What can I *only* see?
 Linguist: I don’t know, what?
 Child: Your beard.

We can understand what was intended by the child’s question, but the question is nonetheless ungrammatical for adult speakers of English. I have nothing to say here regarding the acquisition path of focus which made this utterance possible. I thank the anonymous bearded linguist for sharing this conversation.

A common strategy for asking the intended question in English is the use of *the only*, as in “Who is the only one that you met?” See Herdan (2005); McNally (2008) for discussion of this DP-internal *only*.

¹¹⁹There is of course a question as to where the *wh*-words are interpreted at LF; the availability of association with *only* here suggests that they are not covertly moving to C at LF. This accords with, for example, Sobin’s (2010) approach to these types of echo questions, which he terms “syntactic echo questions.” See Aoun and Li (1993) for this line of argumentation applied to *wh*-in-situ in Mandarin Chinese.

¹²⁰In these examples the echo-*wh*-phrase is also the F-marked constituent that *only* associates with.

¹²¹I thank Danny Fox for bringing this issue to my attention and for suggesting these types of examples as a test.

out *only*, in (296). In particular, if you met both John and Mary on John's birthday, "John" would be a possible weak answer for the question in (296) but not for the question in (295). Example (295) has an expected semantics clearly distinct from the meaning of (296), but it is still ungrammatical with the intended association.

- (295) * [Which boy]_{F,i} did you *only* meet ____i on his_i birthday?
 Intended: which boy x is such that you met only x on x 's birthday?
- (296) ✓ [Which boy]_i did you meet ____i on his_i birthday?

Therefore we can safely say that the ungrammaticality of questions such as (293) is not simply due to pragmatic blocking of these forms because their use of *only* is somehow redundant. *Only* is unable to associate with a fronted *wh*-phrase, nor with a part of it, whether that fronting is overt or covert. This is predicted by the generalization that *only* cannot ever associate backwards, regardless of the type of material F-marked, and my proposal for these facts.

Above I compared fronted *wh*-phrases with in-situ *wh*-phrases in echo questions, but there are other types of in-situ *wh*-phrases as well. In English multiple *wh*-questions, only one *wh*-phrase is fronted, while other *wh*-phrases stay in-situ. There are two schools of thought regarding the question of where these in-situ *wh*-phrases are interpreted at LF: either all *wh*-phrases must be near the complementizer for interpretation, and therefore surface-in-situ *wh*-phrases must covertly move to C (Karttunen, 1977, a.o.); or *wh*-phrases may be interpreted in-situ at LF, and therefore surface-in-situ *wh*-phrases need not move covertly (Hamblin, 1973, a.o.). See Kotek (2014) for a recent review of these two views.

One possibility, proposed by Dayal (1996, 2002), is that pair-list readings of multiple *wh*-questions require covert movement of the in-situ *wh*-word to C, but single-pair readings do not. Example (298) below, we see that when *only* associates with the in-situ *wh*-phrase, the pair-list reading is no longer available. This can be explained under Dayal's proposal: association with *only* blocks covert movement of the in-situ *wh*-phrase to C, making the pair-list reading unavailable. The single-pair reading, however, does not require covert movement of the in-situ *wh*-phrase, and is therefore compatible with the association with *only* in (298).¹²²

- (297) *Baseline*: I can tell you [which student read which book].
 ✓ single pair, ✓ pair list
- (298) I can tell you [which student *only* read [which book]_F].
 ✓ single pair, *pair list

¹²²Note that this is a superiority-obeying multiple *wh*-question, so it is not an intervention effect of the sort in Pesetsky (2000); Beck (2006). See Kotek (2014, §4.2) for a different view, that this configuration does instantiate an intervention effect, and for further discussion of this example.

Let's now turn to cases of overt *wh*-movement with *even*. Consider first a case with a complex *wh*-phrase "which president," with pitch accent on "president" (299). This question is grammatical with *even* associating backwards. As this *wh*-question is embedded, its grammaticality cannot be due to an echo question reading.

(299) *Even* can associate with a *which*-phrase moved above it:

✓ He told me [which PRESIDENT he *even* met ____].

Note that the pitch accent on "president" in (299) could be compatible with two different F-markings: F-marking on the DP "which president" or F-marking on the restrictor "president." We can distinguish between these two F-markings by considering the scalar inference introduced by *even*. The utterance is compatible with a reading where *even* introduces a scalar inference expressing the unlikelihood of the person in question meeting presidents (300a), corresponding to F-marking on the predicate "president." The utterance does not have a reading where *even* introduces a scalar inference expressing the unlikelihood of the person in question meeting the person that they met as opposed to other people in general (300b), corresponding to F-marking on the entire DP "which president."

(300) *Even* can associate with *the restrictor* in a *which*-phrase moved above it:

He told me [which PRESIDENT he *even* met ____]. (299)

- a. ✓ He told me [which [president]_F he *even* met ____].
 ~ it is unlikely for him to meet presidents, as opposed to other sorts of people.
- b. * He told me [[which president]_F he *even* met ____].
 ~ it is unlikely for him to meet the person that he met, as opposed to other people.

The fact that only the interpretation with F-marking on the restrictor (300a) is possible is even clearer when we use a restrictor that doesn't have a reasonable alternative in the context. Consider example (301) below. The two possible parses, differing in the position of F-marking, are given below with predicted paraphrases for the inference of *even*.

(301) Mary has made a series of surprising marriages in her life. #You'll never guess which PERSON she *even* married last year. (David Pesetsky, p.c.)

- a. # You'll never guess [which [person]_F she *even* married last year].
 ~ it is unlikely for Mary to marry people, as opposed to other entities (?).
- b. * You'll never guess [[which person]_F she *even* married last year].
 ~ it is unlikely for Mary to have married the person that she married, as opposed to other people.

The reading in (301a) is available but strange because—under normal circumstances—we expect Mary to marry a person, and therefore it is not clear what the other, more likely alternatives are. If we knew ahead of time that Mary was, for example, an alien, the use of *even* in this sentence becomes felicitous. The reading predicted by (301b), in contrast, is a reasonable inference that is supported by the context. The fact that (301) is judged as semantically anomalous tells us that the only reading that the grammar allows for this utterance is the reading in (301a), where F-marking is on the restrictor.

The contrasts above accord with our generalizations regarding backwards association with *even*, and are explained by my proposal. *Even* is able to associate backwards with an F-marked constituent in the *restrictor* of a DP which has moved out of its scope, but not with the entire DP (section 5.2).

This view is further supported by the behavior of simplex *wh*-words: example (302) below with *even* associating backwards with *who* is ungrammatical, even though the parallel echo question in (303) is grammatical. In the case of a simplex *wh*-phrase, there is no option to F-mark only its restrictor independent of the entire DP, as is possible with a complex *wh*-phrase (300a).

(302) *Even* also cannot associate with *who* moved above it:

* You'll never guess [[*who*]_F I *even* met ____].

Intended: You'll never guess [who *x* is such that I met *x*]

↪ it is unlikely for the speaker to meet *x*, as opposed to other people.

(303) ✓ You *even* met [*who*]_F?

The availability of the interpretation in (300a) shows us that focus association of an operator with material which has *wh*-moved outside of the operator's scope is not impossible, contrary to what has been claimed in Tancredi (1990) and subsequent work. However, we see that association in this configuration is indeed very restricted, possible with *even* but not *only*, and with association with a predicate in the restrictor of a *wh*-phrase, but not with the entire *wh*-phrase.

The contrasts presented above are explained by my proposal and my discussion in section 5.2, based on the behavior of non-*wh* DPs. *Even* is able to associate backwards with F-marking in the *restrictor* of a DP which has moved out of *even*'s scope, but not with the entire moved DP. The proposal presented here therefore predicts the patterns of possible backwards association in *wh*-questions, presented in this section.

5.5 Summary

In chapter 4 I presented my proposal for the patterns of backwards association of *only* and *even* with F-marking in the restrictor of a DP which has moved out of the scope of the operator. In this chapter I extended this approach to patterns of backwards association with F-marked determiners and with F-marking on the entire moved DP. I showed that, in contrast to backwards association with material in the restrictor, backwards association in these configurations is impossible with both *only* and *even*. These patterns of association are summarized in the following table:

(304) Summary of backwards association patterns:

LF configuration	<i>only</i>	<i>even</i>
$\underbrace{[\text{DP D } \dots \alpha_F \dots]} \dots \text{Op} [\dots \text{ } \dots]$	×	○
$\text{DP}_F \dots \text{Op} [\dots \text{ } \dots]$	×	×
$\underbrace{[\text{DP DP}_F \dots]} \dots \text{Op} [\dots \text{ } \dots]$	×	×

In section 5.1 I proposed that F-marked determiners in the lower copy of movement systematically lose their F-marking as a consequence of Trace Conversion. In section 5.2 I showed that the semantics of both *only* and *even* predict backwards association with the entire DP to be impossible. The source of the explanation in these two cases is different: I predict backwards association with F-marked determiners to be impossible regardless of the semantics of the focus-sensitive operator, whereas backwards association with the entire F-marked DP is impossible because of the semantics of *only* and *even*. In the appendix to this chapter I discuss an extension of my proposal to *also* and show that backwards association with the entire DP is in fact possible, as predicted by the semantics of *also*. Finally, in this section I used these findings to investigate the structure of certain DP constructions and patterns of backwards association with *wh*-phrases.

Appendix: Backwards association with *also*

In this chapter I showed that both *even* and *only* are unable to associate backwards with F-marked determiners and F-marked DPs. In this appendix I will extend the discussion to the additive operator *also*, based in part on the work of Rullmann (2003).

Let's first look back at the shape of the proposals I presented in this chapter. In the case of F-marked determiners, backwards association is impossible because the F-marked determiner in the lower copy is replaced in the process of Trace Conversion, predicting that backwards association

will be impossible for any focus-sensitive operator. In contrast, in the case of F-marking on the entire moved DP, my explanation for the lack of backwards association with *even* and *only* was based on these operators' semantic contributions. An important prediction is that there could exist a different focus-sensitive operator with a semantics distinct from *even* or *only* which allows backwards association in one case but not the other.

Here I will argue that *also* substantiates this prediction. The semantics of the additive inference predicts backwards association with an entire DP to be possible. This explains various differences in the backwards association patterns of *even* and *also*. This approach to backwards association with *also* has been proposed by Rullmann (2003), without the Copy Theory, as association with an F-marked trace. While the semantics of additives allows such a non-Copy-Theoretic approach to backwards association to work, the contrasts between *even* and *also* further motivate the need for a Copy-Theoretic approach to patterns of backwards association.

I will begin by presenting some basic data on backwards association with additives. Krifka (1998) first noted the ability of additives, cross-linguistically, to associate with preceding material. In some cases, such as English *too*, this corresponds to the fact that the focus-sensitive operator is adjoined to the right. But it can also be observed with the adverb *also*, which allows for canonical backwards association configurations of the form I have studied here.

- (305) a. Peter ate [pasta]_F. He *also* ate [polenta]_F.
 b. [Peter]_F ate pasta. [Pia]_F *ALSO* ate pasta.
- (306) Every [student]_F will come to the party. Every [professor]_F will *ALSO* come to the party.
- (307) [John]_F, I *ALSO* consider ___ intelligent.

Interestingly, as Krifka (1998) discusses, additives which associate backwards receive pitch accent themselves, as indicated by capital letters, and their preceding focus associates are realized with prosody characteristic of contrastive topics. Due to this, Krifka (1998) argues that when additives associate backwards, they are always associating with a contrastive topic. Rullmann (2003) challenges this view, however, and argues that backwards association with additives (in particular, *either*) is dependent on the associate having moved out of the additive's scope. Example (308) is a contrast Rullmann (2003) presents to support this argument. Both "John" and "Mary" could be contrastive topics, given an appropriate discourse context, but there is no way that "John" originated within the scope of the additive *either*. As Rullmann (2003) discusses, the additive *either* is an NPI and therefore must be adjoined in the embedded clause.

- (308) Rullmann (2003, ex. 86):
 a. # [John]_F said [Mary didn't come *EITHER*].

- b. ✓ John said [[Mary]_F didn't come *EITHER*].

Arguments parallel to those I presented for *even* in section 4.1 can be constructed to show that the backwards associate of *also* must have originated from within the scope of *also*:

(309) *Also* associating with leftward topic requires base position in *also*'s scope: (cf 171)

- a. ✓ The [Canadians]_F, the report *ALSO* convinced the judges that we spied on ____.
 b. ✓ The [Canadians]_F, the report convinced the judges that we *ALSO* spied on ____.
 c. ✓ The [judges]_F, the report *ALSO* convinced ____ that we spied on the Canadians.
 d. * The [judges]_F, the report convinced ____ that we *ALSO* spied on the Canadians.

(Ungrammatical with the intended association of *judges* with *also*.)

Now I will turn to backwards association with determiners and DPs, which has been the topic of this chapter. *also* is unable to backwards associate with the F-marked determiner of a moved DP, as illustrated in (310) below. This is predicted by my analysis from section 5.1.

(310) *Also* cannot associate backwards with a determiner.¹²³

- a. I knew they arrested [some]_F protesters on some nights. ✓ But it turns out they *also* arrested [every]_F protester on some nights.
 b. I knew [some]_F protesters were arrested on some nights. *But it turns out [every]_F protester was *also* arrested on some nights.

However, in contrast to *even* and *only*, *also* is able to associate backwards with an entire moved DP:

(311) *New degree requirements, with also*: (based on 268/281)

Context: David wants to change the requirements for the Linguistics degree program. The proposal will pass if *both* (a) at least two-thirds of the professors vote for the change *and* (b) every student votes for the change.

The votes have been counted and the proposal passed. Almost all of the faculty voted for the change and...

- ✓ [Every student]_F (had) *also* voted for the change.

That *also* can backwards associate with a DP is also very clear in the following *wh*-phrase example in (312), where the use of *also* is felicitous. The felicitous reading is the one in (312b) which involves

¹²³The scope-bearing "on some nights" is necessary in order to make the alternative with "every" not strictly stronger than the alternative with "some" that it is congruent to. This is related to a requirement of additives that, as Beaver and Clark (2008, p. 73) state it, "the pre-existing partial answer is not entailed by the prejacent."

associating with the individual who Mary married, i.e. the entire individual DP, rather than the restrictor “person.” This contrasts from the *even* variant in (301) which was judged as semantically odd, due to the lack of a DP focus parse. This also explains the fact that *also* can backwards associate with an entire simplex *wh*-phrase, in (313), in contrast to *only* (293a) and *even* (302).

- (312) Mary has made a series of surprising marriages in her life. We knew she married Bill last year, and that was strange enough. But it turns out she had also married someone else last year. ✓You’ll never guess which PERSON she ALSO married last year.¹²⁴
- a. # You’ll never guess [which [person]_F she ALSO married last year].
 ~ Mary married someone who is not a person last year.
- b. ✓ You’ll never guess [[which person]_F she ALSO married last year].
 ~ Mary married someone else (besides Bill) last year.
- (313) ✓ You’ll never guess [who]_F I ALSO met.
 ‘You’ll never guess [who *x* is such that I met *x*].’
 ~ the speaker met someone besides *x*.

We see that *also* is able to associate backwards with F-marking in the restrictor of a moved DP, and with the entire moved DP which is F-marked, but not with an F-marked determiner. This can be explained by the semantics of *also*. I assume here that *also* introduces an additive presupposition that at least one non-prejacent alternative is true.¹²⁵ For illustration purposes, consider the following derivation for (312b):

(314) Computing (312b):

- a. LF: ...[_{CP} [_{DP,F1} which person] λ1 Mary also [_{vP} married [_{DP,F1} the person ①] last year]]
- b. $H = \{h_0, h_1, h_2, h_3\}$
- i. $h_0(F1) = \llbracket \text{the person } \textcircled{1} \rrbracket = \lambda g_a . \lambda w_s : g(1) \text{ is a person} . g(1)$
- ii. $h_1(F1) = \llbracket \text{John} \rrbracket = \lambda g_a . \lambda w_s \text{John}$
- iii. $h_2(F1) = \llbracket \text{Sue} \rrbracket = \lambda g_a . \lambda w_s \text{Sue}$

¹²⁴In contrast to the *even* variant in (301) above, it seems that no pitch accent or contrastive topic prosody is required on “person” here. Here I use SMALL CAPS to indicate the expected locus of this prosody, if present.

¹²⁵It has been argued by Kripke (1990), based on *too*, that the purely existential characterization of the additive inference is not correct. See also Heim (1992). The argument here will also work for their anaphoric semantics for additives, in which the additive requires a particular, contextually-determined alternative to be true. Just one complication arises when combining this view with my approach to backwards association of *also* with a DP: I predict the anaphoric individual which satisfies the additive must be of type *e*, the type of the Trace Converted lower copy of movement, rather than higher type nominals. I will leave this question open for future research.

- iv. $h_3(F1) = \llbracket \text{Bill} \rrbracket = \lambda g_a . \lambda w_s . \text{Bill}$
- c. $\llbracket vP \rrbracket (h \parallel g) = \lambda w_s . \text{Mary married } h(F1)(g)(w) \text{ last year}$
- d. $\underline{\text{also}} \rightsquigarrow \exists h \in H . h_0(F1)(g)(w) \neq h(F1)(g)(w) \wedge \text{Mary married } h(F1)(g)(w) \text{ last year}$

I assume that the non-prejacent alternatives in H return constant functions for F1, ranging over relevant individuals—here, John, Sue, and Bill. The additive inference of *even* in (314d) is satisfied by there being one person, Bill—which one focus assignment, h_3 returns for focus index F1—that Mary had married last year. Unlike in the case of *even*, where F-marking on the lower copy of a DP meant that we result in a scalar inference that is not satisfiable, here we yield an existential inference that is satisfied in the context. Intuitively, the difference is that *even* as well as *only* need to know which alternative is the prejacent, whereas *also* simply needs to ensure that there are true alternatives, and does not actually need to know which alternative (value for F1 by $h \in H$) is the prejacent value.

One additional example of note is the following contrast between *even* and *also* in backwards association with subject control constructions:

(315) *Even* can't backwards associate across control:

- * JOHN wants to *even* come. (based on 174b)
- a. * [John]_F wants [PRO to *even* come].
- b. * John wants [PRO_F to *even* come].

(316) Additives can backwards associate across control:

- a. ✓ John wants [PRO_F to *also* come].
- b. ✓ John wants [PRO_F to come *too*]. (Heim, 1992, fn. 13)

The ability of additives, but not *even* or *only*, to associate with the PRO in this configuration is noted as a puzzle by Heim (1992, fn. 13). Furthermore, as Heim (1992) argues, discussing (316b), the additive in such examples takes scope over the nonfinite embedding, rather than the matrix clause, because the inference of the additive is that someone else will come, not that someone else wants to come. Under my analysis here, this contrast is explained through the more general difference in backwards association with entire DPs between *even* and *also*.

In conclusion, *also* is able to associate backwards with the entire moved DP in a way that neither *even* nor *only* are able to do, and it is also able to associate with F-marking in the restrictor of a moved DP. These possible patterns of association are summarized in the following table, updated from (304) above:

(317) Summary of backwards association patterns:

LF configuration	<i>only</i>	<i>even</i>	<i>also</i>
$[\text{DP } D \dots \alpha_F \dots] \dots \text{Op} [\dots _ \dots]$ 	x	○	○
$\text{DP}_F \dots \text{Op} [\dots _ \dots]$ 	x	x	○
$[\text{DP } \text{DP}_F \dots] \dots \text{Op} [\dots _ \dots]$ 	x	x	x

In this appendix I have discussed the focus-sensitive operator *also* and its backwards association patterns. As proposed by Rullmann (2003) before, *also*'s semantics allows it to associate backwards with an "F-marked trace," which corresponds under the Copy Theory to F-marking on the entire lower copy of movement. However, the range of structural configurations under which *also* is able to associate backwards differs from that of *even*: in particular, as we saw in section 5.2 in this chapter, *even*'s semantics does not allow it to associate backwards with an entire F-marked DP, although it can associate with F-marking in the DP's restrictor. These differences between the backwards association configurations of *even* and *also* thus support the Copy Theory-based proposal presented in this dissertation, which is able to distinguish between F-marking within the DP's restrictor and F-marking on the entire DP, unlike in Rullmann's (2003) proposal which does not adopt the Copy Theory.

Chapter 6

Even in downward-entailing contexts

In this chapter I will extend the discussion to examples with *even* in the scope of downward-entailing operators. It has been known since Karttunen and Peters (1979) that the scalar inference introduced by *even* is reversed in such contexts. For example, consider the pair of examples in (318), each of which is felicitous in the context given. (In this section, downward-entailing operators will be in bold.)

(318) **The scalar inference of *even* is reversed in a downward-entailing environment:**

a. **Context:** Hadas normally doesn't like seafood. But today she was feeling adventurous and ate many things.

✓ Hadas *even* ate the [shrimp]_F.

↪ $(\lambda w . \text{Hadas ate the shrimp in } w) <_{\text{likely}} (\lambda w . \text{Hadas ate the } <\dots\text{alternative}\dots > \text{ in } w) \dots$

b. **Context:** Hadas normally loves eating shrimp. But today she was feeling sick and didn't eat anything.

✓ Hadas **didn't *even*** eat the [shrimp]_F.

↪ $(\lambda w . \text{Hadas ate the shrimp in } w) >_{\text{likely}} (\lambda w . \text{Hadas ate the } <\dots\text{alternative}\dots > \text{ in } w) \dots$

The addition of sentential negation (in bold) in (318b) has the effect of flipping the scale in the inference expressed. The use of *even* in (318a) expresses that Hadas is relatively *less* likely to eat shrimp as opposed to other foods, whereas its use in (318b) expresses that it is *more* likely for Hadas to eat shrimp.

Different proposals have been made for explaining this *scale reversal* behavior of *even*. One family of proposals, called the "scope theory" of *even*, involves the adverb *even* taking wider scope than its surface position indicates (Karttunen and Peters, 1979; Wilkinson, 1996; Guerzoni, 2004; Nakanishi, 2012, a.o.). Another approach to the scale reversal of *even*, called the *lexical ambiguity theory*, assumes two different but related, polarity-sensitive lexical entries for *even*, using opposite scales

for the computation of *even*'s scalar inference (Rooth, 1985; von Stechow, 1991a; Rullmann, 1997; Schwarz, 2005; Giannakidou, 2007, a.o.).

In my discussion thus far, I have assumed that adverbial *even* is interpreted in its surface position. If instead it is possible for *even*'s scope to expand, this may lead to an alternative explanation for *even*'s ability to associate backwards, as I will show below. In this chapter I therefore investigate the behavior of *even* in such downward-entailing environments, and in particular what different theories for *even*'s scope reversal behavior predict for backwards association. I will show that the scope theory of *even* cannot explain the ability of *even* to associate with material which has moved out of its scope. Instead, the behavior of *even* in downward-entailing contexts, with and without backwards association, is best explained by the lexical ambiguity theory of *even*.

6.1 The lexical ambiguity theory

I will begin by presenting the lexical ambiguity theory of *even*, and show how this approach can accurately interpret the inference of *even* both with and without backwards association.

The lexical ambiguity theory of *even*, first laid out in Rooth (1985), proposes that there are two variants of English *even*, whose distribution depends on their environment. The standard *even* which introduces an inference of the relative *unlikeliness* of the prejacent is a positive polarity item (here *even*_{PPI}). In addition, Rooth (1985) proposes that there is a reverse scale *even* which introduces an inference of the relative *likeliness* of the prejacent and is a negative polarity item (here *even*_{NPI}). English VP-*even* is then interpreted in its pronounced position. The existence of these two types of *evens* is supported by the fact that some languages lexicalize these two items differently. See König (1991); von Stechow (1991a); Rullmann (1997) and others for German, Giannakidou (2007) for Greek, Lahiri (2008) for Spanish.

Below are the definitions for *even*_{PPI} and *even*_{NPI} based on the definition for *even* that I developed in chapter 3. The *even* that we have discussed previously corresponds to *even*_{PPI} here.

(319) *Even*_{PPI} and *even*_{NPI} under the lexical ambiguity theory:

- a. $[[\textit{even}_{\text{PPI}}]]_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle} \equiv \lambda g_a \cdot \lambda w_s \cdot \lambda P_{\langle a, \langle s, t \rangle \rangle} : \quad (=151)$
 $\underbrace{\text{GEN } (g' \in G, h \in H) (h \neq h_0 \rightarrow P(g' \parallel h_0) <_{\text{likely}} P(g' \parallel h))}_{\text{scalar inference}} \cdot \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\text{truth condition}}$
- b. $[[\textit{even}_{\text{NPI}}]]_{\langle a, \langle s, \langle \langle a, \langle s, t \rangle \rangle, t \rangle \rangle} \equiv \lambda g_a \cdot \lambda w_s \cdot \lambda P_{\langle a, \langle s, t \rangle \rangle} : \quad (=152)$
 $\underbrace{\text{GEN } (g' \in G, h \in H) (h \neq h_0 \rightarrow P(g' \parallel h_0) >_{\text{likely}} P(g' \parallel h))}_{\text{scalar inference}} \cdot \underbrace{P(h_0 \parallel g)(w) \text{ is true}}_{\text{truth condition}}$

Let's see how these two lexical entries for *even* can account for the scale reversal behavior of *even* in example (318) above. I begin with example (318a). Here, because *even* is not in a downward-

entailing context, the positive polarity item $even_{PPI}$ must be used. $Even_{PPI}$ has the same semantics as the regular $even$ that we have looked at previously, leading to an inference that it is relatively less likely that Hadas eat the shrimp than that she eat other foods. The truth conditions of (318a) are unaffected by $even$, and will be true if Hadas ate the shrimp.

(320) Interpreting (318a) with PPI $even$:

“Hadas $even_{PPI}$ ate the [shrimp]_F.”

a. Narrow syntax: $[\emptyset_{\text{the}} \text{Hadas}] \underbrace{even_{PPI} [_{vP} [\emptyset_{\text{the}} \text{Hadas}] \text{ate the [shrimp]}_{F1}]}_{\text{scope}}$

b. LF: $[\emptyset_{\text{the}} \text{Hadas}] \lambda 1 \underbrace{even_{PPI} [_{vP} [\text{the Hadas } \textcircled{1}] \text{eat the [shrimp]}_{F1}]}_{\text{scope}}$

c. $[[vP]] (h \parallel g) = \lambda w_s : g(1) \text{ is Hadas} . g(1) \text{ eats the } h(F1)(g)(w) \text{ in } w$

d. $H = \{h_0, h_1, h_2\}$

i. $h_0(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is shrimp in } w$

ii. $h_1(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is fish in } w$

iii. $h_2(F1) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is meat in } w$

e. \sim inference of $even_{PPI}$:

true in context (318a)

$$\text{GEN } (g' \in G, h \in H)(h \neq h_0 \rightarrow (\lambda w . \text{LA}([vP](g' \parallel h_0)(w))) <_{\text{likely}} (\lambda w . \text{LA}([vP](g' \parallel h)(w))))$$

$$\iff \text{GEN } (g' \in G, h \in \{h_1, h_2\})((\lambda w . g'(1) \text{ is Hadas and eats the } h_0(F1)(g')(w) \text{ in } w) <_{\text{likely}} (\lambda w . g'(1) \text{ is Hadas and eats the } h(F1)(g')(w) \text{ in } w))$$

$$\approx ((\lambda w . \text{Hadas eats the shrimp in } w) <_{\text{likely}} (\lambda w . \text{Hadas eats the fish in } w) \wedge$$

$$(\lambda w . \text{Hadas eats the shrimp in } w) <_{\text{likely}} (\lambda w . \text{Hadas eats the meat in } w))$$

Next I turn to example (318b). Here $even$ is interpreted in its surface scope, and this means that it is in the scope of a downward-entailing operator. The NPI $even$, $even_{NPI}$ will be used. This leads to a reversed scalar inference—that it is more likely for Hadas to eat the shrimp than to eat other foods—as desired. The truth conditions will again be unaffected by $even$, and here (318b) will be true if Hadas did not eat the shrimp.

(321) Interpreting (318b) with NPI $even$:

“Hadas didn’t $even_{NPI}$ eat the [shrimp]_F.”

a. Narrow syntax: $[\emptyset_{\text{the}} \text{Hadas}] \underbrace{\text{didn't } even_{NPI} [_{vP} [\emptyset_{\text{the}} \text{Hadas}] \text{eat the [shrimp]}_{F1}]}_{\text{scope}}$

b. LF: $[\emptyset_{\text{the}} \text{Hadas}] \lambda 1 \underbrace{\text{Neg } even_{NPI} [_{vP} [\text{the Hadas } \textcircled{1}] \text{eat the [shrimp]}_{F1}]}_{\text{scope}}$

c. \sim inference of $even_{NPI}$:

true in context (318b)

$$\text{GEN } (g' \in G, h \in H)(h \neq h_0 \rightarrow (\lambda w . \text{LA}([vP](g' \parallel h_0)(w))) >_{\text{likely}} (\lambda w . \text{LA}([vP](g' \parallel h)(w))))$$

$$\begin{aligned}
&\iff \text{GEN } (g' \in G, h \in \{h_1, h_2\})((\lambda w . g'(1) \text{ is Hadas and eats the } h_0(\text{F1})(g')(w) \text{ in } w) >_{\text{likely}} \\
&\quad (\lambda w . g'(1) \text{ is Hadas and eats the } h(\text{F1})(g')(w) \text{ in } w)) \\
&\approx ((\lambda w . \text{Hadas eats the shrimp in } w) >_{\text{likely}} (\lambda w . \text{Hadas eats the fish in } w) \wedge \\
&\quad (\lambda w . \text{Hadas eats the shrimp in } w) >_{\text{likely}} (\lambda w . \text{Hadas eats the meat in } w))
\end{aligned}$$

Let's now look at a case where *even* associates backwards, in a downward-entailing context. I will show that the lexical ambiguity theory can account for such examples as well. Consider example (322) in a context where "student" contrasts with "professor." The sentence is grammatical with the intended association of *even* with the leftward "student," with a scalar inference that it is considered more likely for a student to be to the party than for a professor to be there.

(322) ✓ No [student]_F *even* came to the party.

The derivation of (322) is given below in (323). Under my proposal, *even* associates with F-marking in the subject's lower copy of movement. As the lower copy of movement undergoes Trace Conversion—illustrated in (323b)—its quantificational part is overwritten. *Even* will then associate with the F-marked "student" in "the student *x*" in the complement of *even*. Because *even* is within the scope of a downward-entailing operator "no," it will be interpreted as *even*_{NPI}.

(323) Interpreting (322) with NPI *even*:

"No [student]_F *even*_{NPI} came to the party."

a. Narrow syntax: [No [student]_{F1}] *even*_{NPI} [_{vP} [no [student]_{F1}] came to the party]

b. LF: [No [student]_{F1}] λ1 *even*_{NPI} [_{vP} [the [student]_{F1} ①] came to the party]

c. [[vP]] (*h* || *g*) = λ*w*_s : *g*(1) is a *h*(F1)(*g*)(*w*) in *w* . *g*(1) came to the party in *w*

d. *H* = {*h*₀, *h*₁}

i. *h*₀(F1) = λ*g*_a . λ*w*_s . λ*x*_e . *x* is a student in *w*

ii. *h*₁(F1) = λ*g*_a . λ*w*_s . λ*x*_e . *x* is a professor in *w*

e. ~ inference of *even*_{NPI}:

GEN (*g*' ∈ *G*, *h* ∈ *H*)(*h* ≠ *h*₀ → (λ*w* . LA([[vP]] (*g*' || *h*₀)(*w*))) >_{likely} (λ*w* . LA([[vP]] (*g*' || *h*)(*w*))))

⇔ GEN (*g*' ∈ *G*)((λ*w* . *g*'(1) is a *h*₀(F1)(*g*'(1))(*w*) in *w* ∧ *g*'(1) came to the party in *w*) >_{likely} (λ*w* . *g*'(1) is a *h*₁(F1)(*g*'(1))(*w*) in *w* ∧ *g*'(1) came to the party in *w*))

⇔ GEN (*x* ∈ *D*_e)((λ*w* . *x* is a student and came to the party in *w*) >_{likely}

(λ*w* . *x* is a professor and came to the party in *w*))

As we can see, the derivation in (323) yields the correct inference that we intuitively assign to this sentence. If *even*_{PPI} were used here in place of *even*_{NPI}, we would yield the wrong inference for (322), that it is *less* likely for a student to come to the party than for a professor to be there.

Consequently, we see that the lexical ambiguity theory of *even* is able to correctly interpret cases of backwards association of *even* in downward entailing environments, while preserving our working assumption that *even* is interpreted in its pronounced position.

6.2 The scope theory of *even*

In this section I consider the so-called *scope theory* of *even* in downward-entailing contexts (Karttunen and Peters, 1979; Wilkinson, 1996; Guerzoni, 2004; Nakanishi, 2012, a.o.). Under this view, *even* covertly moves to a higher position to take scope over the downward-entailing operator.¹²⁶ Including the downward-entailing quantifier in propositions used to construct the scalar inference results in the apparent scale reversal, without the need for multiple homophonous *evens* as in the lexical ambiguity theory.

The scope theory is important to consider in relation to the puzzle of backwards association. If the scope theory of *even* is correct, this opens up the possibility that the backwards association phenomenon I have studied here is in fact illusory. *Even* is able to scope at LF in a position higher than indicated by its surface position. If this LF scope of *even* now includes the F-marked material, as in example (324) below, the apparent backwards association is no longer a mystery.

(324) The scope theory as potential solution to backwards association:

- a. PF: A [professor]_F will [*even* [come to the party]]
- b. LF: *even* [A [professor]_F will [___ [come to the party]]]

Furthermore, note that the behavior of *only*, in contrast, does not show any sensitivity to downward-entailing contexts. Suppose this difference corresponds to a difference in the ability of these two focus-sensitive operators to covertly move: *even* is able to move covertly and expand its scope, but *only* cannot. This hypothesis may be able to explain why *even* is able to associate apparently backwards, but *only* cannot.

Below I will consider two variants of the scope theory for VP-*even*, differing in exactly what material is moved. I will show that both variants are unable to derive the correct pattern of backwards

¹²⁶Karttunen and Peters (1979), who first proposed the scope theory, did not describe this scope taking as a form of covert movement, instead proposing an “Even Rule” which governs the mapping between the pronounced position and its interpreted scope. The effect of their Even Rule can be thought of as the effect of covert movement, though, and this is the way that the scope theory has been described in more recent literature.

association. My proposal for backwards association with *even* but not *only*, laid out in chapter 4, is therefore independently necessary. Furthermore, I will show that one common conception of the scope theory of *even* makes clearly incorrect predictions for the backwards association behavior of *even*, and cannot be maintained.

6.2.1 Covert movement of *even* alone

I first consider the variant of the scope theory where the adverb *even* can covertly move alone to a higher projection with propositional type. This variant of the scope theory is assumed in Wilkinson (1996); Guerzoni (2004); Nakanishi (2012).¹²⁷

I will first illustrate this scope theory with our example of *even* in a downward-entailing context, example (318b). I will illustrate this computation here with covert movement of *even* adjoining to TP, above the subject. For illustration purposes, the movement of the non-quantificational subject from Spec,vP to Spec,TP is not illustrated.

(325) Interpreting (318b) with covert movement of *even*:

“Hadas didn’t *even* eat the [shrimp]_{F1}.”

a. Narrow syntax: [TP Hadas didn’t *even* eat the [shrimp]_{F1}]

b. Even moves at LF: *even* [TP Hadas didn’t eat the [shrimp]_{F1}]

c. $\llbracket \text{TP} \rrbracket (h \parallel g) = \lambda w_s . \text{Hadas didn't the } h(F1)(g)(w) \text{ in } w$

d. \sim inference of *even*:

true in context (318b)

$\text{GEN } (g' \in G, h \in H)(h \neq h_0 \rightarrow \llbracket \text{TP} \rrbracket (g' \parallel h_0) <_{\text{likely}} \llbracket \text{TP} \rrbracket (g' \parallel h))$

¹²⁷Much of both Wilkinson (1996) and Nakanishi (2012) are based on Karttunen and Peters’s (1979) scope theory for constituent-marking *even*, where *even* and the constituent it adjoins covertly moves together, as a covert movement which we might now call QR. For example, all of the derivations in Wilkinson (1996) explicitly illustrated with trees are movements of this form.

However, Wilkinson (1996) also discusses a scope theory for adverbial *even*, where “*even* has to move to the higher Infl” (p. 199). Similarly, Nakanishi (2012) investigates cases where the focus associate of *even* moves out of the surface scope of *even*, stating that “the embedded VP-*even* must also be adjoined to the matrix IP”—in other words, interpreting the adverbial *even* in an LF position much higher than its surface position. It is this movement of adverbial *even* that I discuss in this section.

Note that this movement does not leave a semantically contentful trace. In the LF representations given by Wilkinson (1996), Guerzoni (2004), and Nakanishi (2012) which are meant to reflect this movement of *even*, there is no trace position or associated λ -binder.

$$\begin{aligned} &\iff \text{GEN } (g' \in G, h \in \{h_1, h_2\}) \left(\begin{array}{l} (\lambda w . \text{Hadas didn't eat the } h_0(\text{F1})(g')(w) \text{ in } w) <_{\text{unlikely}} \\ (\lambda w . \text{Hadas didn't eat the } h(\text{F1})(g')(w) \text{ in } w) \end{array} \right) \\ &\approx \left(\begin{array}{l} (\lambda w . \text{Hadas didn't eat the shrimp in } w) <_{\text{unlikely}} (\lambda w . \text{Hadas didn't eat the fish in } w) \wedge \\ (\lambda w . \text{Hadas didn't eat the shrimp in } w) <_{\text{unlikely}} (\lambda w . \text{Hadas didn't eat the meat in } w) \end{array} \right) \end{aligned}$$

The scalar inference computed in (325) accords with our intuitions. *Even* here expresses that it is relatively *unlikely* for Hadas to *not* eat the shrimp as compared to other foods. This is identical to the inference computed for this example using a stationary NPI *even* in (321) above, which expressed that it is relative *likely* for Hadas to eat the shrimp as compared to other foods.

Let's now look at how the scope theory might help with cases of apparent backwards association. We begin again by considering example (322) from above, repeated below as (326). At LF, *even* covertly moves to a position above the downward-entailing operator (326a). The subject will then be in the scope of *even*. The movement of the subject from Spec,vP to Spec,TP is again not illustrated here for simplicity.

(326) Interpreting (322) with covert movement of *even*:

- a. Narrow syntax: No [student]_{F1} *even* came to the party
- b. Even moves at LF: *even* [_{TP} no [student]_{F1} came to the party]
- c. $\llbracket \text{TP} \rrbracket (h \parallel g) = \lambda w_s . \nexists x (x \text{ is a } h(\text{F1})(g)(w) \text{ and came to the party in } w)$
- d. $H = \{h_0, h_1\}$
 - i. $h_0(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a student in } w$
 - ii. $h_1(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a professor in } w$
- e. \rightsquigarrow inference of *even*:

$$\begin{aligned} &\text{GEN } (g' \in G, h \in H) (h \neq h_0 \rightarrow \llbracket \text{TP} \rrbracket (g' \parallel h_0) >_{\text{likely}} \llbracket \text{vP} \rrbracket (g' \parallel h)) \\ &\iff \text{GEN } (g' \in G) \left(\begin{array}{l} ((\lambda w . \nexists x (x \text{ is a } h_0(\text{F1})(g')(w) \text{ and came to the party in } w)) <_{\text{unlikely}} \\ (\lambda w . \nexists x (x \text{ is a } h_1(\text{F1})(g')(w) \text{ and came to the party in } w)) \end{array} \right) \\ &\iff (\lambda w . \nexists x (x \text{ is a student and came to the party in } w)) <_{\text{unlikely}} \\ &\quad (\lambda w . \nexists x (x \text{ is a professor and came to the party in } w)) \end{aligned}$$
- f. $\iff (\lambda w . \exists x (x \text{ is a student and came to the party in } w)) >_{\text{likely}}$
 $\quad (\lambda w . \exists x (x \text{ is a professor and came to the party in } w))$

The scalar inference introduced by *even* is then as in (326e), expressing the relative *unlikelyness* that no student came the party. This can be restated, by factoring out the negation, as expressing the relative *likeliness* that some student came to the party (326f). This reflects the scale-reversing behavior of *even* due to the presence of the downward-entailing operator. The result in (326) accords with our intuitions about the felicity of this expression.

However, this scope theory makes incorrect predictions regarding the distribution of backwards association with *even*. Consider, for example, the pair of examples in (327) below, modified from (174) above. We observe that the raising raising example in (327a) is judged as grammatical, while the control counterpart in (327b) is ungrammatical with the intended association.

(327) **Scope theory incorrectly predicts no contrast between raising and control:**

a. ✓ **No** [student]_F seems to *even* be at the party. (cf 174a)

b. ***No** [student]_F wants to *even* be at the party. (cf 174b)

Expected scope theory LF: *even* [no [student]_F wants to ___ be at the party]

The scope theory does not predict any difference in the ability of *even* to move covertly out of certain kinds of nonfinite embeddings but not others. Therefore we predict the LF in (327b) above. In this LF, the overt F-marked “student” is within the scope of *even*, and we therefore predict *even* to be able to associate with the predicate “student” in the subject, contrary to fact. The scope theory therefore predicts that backwards association should be possible in the control embedding, just like it is in raising contexts.

Note further that we cannot rescue this scope theory by stipulating that the covert movement step of *even* is not possible across a control clause boundary. Under the scope theory of *even*, *even* would have to move in the exact same configuration as in the expected LF for (327b) to explain the scale reversal of *even* in other, grammatical examples with control embeddings:

(328) ✓ **No one** wants to *even* read the [abstract]_F of this terrible paper.

Scope theory LF: *even* [no one wants to ___ read the [abstract]_F of this terrible paper]

The contrast in grammaticality between the basic example (322) and the control example (327b) therefore acts as an argument against this scope theory of *even*, as the scope theory is unable to predict this contrast.

More generally, the scope theory of *even* predicts that backwards association does *not* require the F-marked constituent to have originated in the surface scope of *even*. The scope theory is therefore unable to explain the facts presented in section 4.1 to show that the movement chain is crucial for backwards association.

Furthermore, simply expanding the scope of *even* then predicts that any constituent in the LF scope of *even* is a potential target for apparent backwards association. Assuming that backwards association with the restrictor of a particular moved DP is possible (329a), this scope theory approach then predicts that *even* could just as well associate with the quantifier of that moved DP (329b), or with the entire DP (329c). However, as I showed in chapter 5, *even* is unable to associate backwards with the quantificational determiner of a moved DP, nor with an entire F-marked DP. The

scope theory approach to backwards association does not predict this sensitivity. In contrast, the proposal I offer here is able to explain these contrasts.

(329) *Even* cannot associate backwards with determiners and entire DPs:

- a. ✓ $\left[\text{DP Quantifier [NP]}_F \dots \text{even} [\dots] \dots \right]$
 Scope theory LF: *even* $\left[\text{DP Quantifier [NP]}_F \dots [\dots] \dots \right]$
- b. * $\left[\text{DP [Quantifier]}_F \text{NP} \dots \text{even} [\dots] \dots \right]$ see section 5.1
 Predicted scope theory LF: *even* $\left[\text{DP [Quantifier]}_F \text{NP} \dots [\dots] \dots \right]$
- c. * $\text{DP}_F \dots \text{even} [\dots] \dots$ see section 5.2
 Predicted scope theory LF: *even* $\text{DP}_F \dots [\dots] \dots$

The challenge to the scope theory presented here is partially recognized by Wilkinson (1996), who notes that *even* cannot associate with material in higher clauses, although her theory predicts that it could be within of the scope of *even* at LF. Wilkinson (1996, p. 199) therefore states that adverbial *even* must identify its associate at S-structure. If we add this surface configurational requirement to *even*, we are unable to use the scope theory to account for backwards association.

To conclude, the scope theory with Wilkinson's surface configurational requirement explicitly disallows backwards association, whereas the scope theory without this requirement would over-generate backwards association. Therefore, the scope theory with covert movement of *even* is unable to explain the patterns of backwards association with *even* presented here.

6.2.2 Covert movement of *even* and its associate

We can imagine another variant of the scope theory—the hypothesis described as the “scope theory” in Rooth (1985), as his hypothetical formalization of the theory sketched in Anderson (1972). In this variant, the adverb *even* and its focus associate form a constituent at LF—call it an “*even*-phrase”—akin to what we observe when *even* adjoins directly to an argument.¹²⁸ If the *even*-phrase is in the scope of a downward-entailing operator, the *even*-phrase will move covertly to outscope the downward-entailing operator.

I will first demonstrate this variant using example (318b) from above.

¹²⁸See also Wagner (2006), which proposes this same syntactic movement for the adverb *only*, in particular for discussion of the movement step in (330)'s LF step 1.

(330) Construction and covert movement of an *even*-phrase in (318b):

“Hadas didn’t *even* eat the [shrimp]_F.”

- a. Narrow syntax: [TP Hadas didn’t *even* [_{vP} eat [DP the [shrimp]_{F1}]]]
- b. LF step 1: [TP Hadas didn’t [_{even-phrase} *even* [DP the [shrimp]_{F1}]]] λx [_{vP} eat x]
- c. LF step 2: [_{even-phrase} *even* [DP the [shrimp]_{F1}]] λy [TP Hadas didn’t y λx [_{vP} eat x]]

Note that under this theory, the adverb *even* is not interpreted adjoined to a constituent of propositional type. We would instead have to define a semantics to allow *even* to compose with two arguments: the first a constituent (here a DP) of extensional type σ which contains focus alternatives and second a constituent of extensional type $\langle \sigma, t \rangle$. Here is such a variant of *even*, based on the proposition-taking *even* in (151):

(331) Semantics for two-place *even*:

$$\begin{aligned} \llbracket \textit{even} \rrbracket_{\langle a, \langle s, \langle \langle a, \langle s, \sigma \rangle \rangle, \langle a, \langle s, \langle \sigma, t \rangle \rangle \rangle \rangle \rangle} &\equiv \lambda g_a . \lambda w_s . \lambda P_{\langle a, \langle s, \sigma \rangle \rangle} . \lambda Q_{\langle a, \langle s, \langle \sigma, t \rangle \rangle \rangle} : \\ \text{GEN } (g' \in G, h \in H) &\left(h \neq h_0 \rightarrow \frac{(\lambda w' . Q(g' \parallel h_0)(w')(P(g' \parallel h_0)(w')))) <_{\textit{likely}}}{(\lambda w' . Q(g' \parallel h)(w')(P(g' \parallel h)(w')))} \right) \\ &\underbrace{\hspace{10em}}_{\text{scalar inference}} \\ &\underbrace{Q(h_0 \parallel g)(w)(P(h_0 \parallel g)(w)) \text{ is true}}_{\text{truth condition}} \end{aligned}$$

This *even* composes with its first argument—here, the DP “the [shrimp]_{F1}”—using Superintensional Functional Application (44) and then combines with its second argument—here, the derived property (λx . Hadas didn’t eat x)—again using Superintensional Functional Application. These modes of application are enforced by the type signature of *even*. We might independently want a lexical entry for *even* of this form, which takes two arguments, because English has DP- and PP- adjoining focus particles, in addition to the adverb studied here.

Below I compute the inference of *even* in (330):

(332) Interpreting the *even*-phrase in (318b/330):

- a. $\llbracket \text{TP} \rrbracket (h \parallel g) = \lambda w_s . \lambda x_e . \text{Hadas didn't eat } x \text{ in } w$
- b. $\llbracket \text{DP} \rrbracket (h \parallel g) = \lambda w_s . \text{the } h(\text{F1})(g)(w)$
- c. $H = \{h_0, h_1, h_2\}$
- i. $h_0(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is shrimp in } w$
 - ii. $h_1(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is fish in } w$
 - iii. $h_2(\text{F1}) = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is meat in } w$

- d. i. $\llbracket \text{DP} \rrbracket (h_0 \parallel g) = \lambda w_s . \text{the shrimp in } w$
 ii. $\llbracket \text{DP} \rrbracket (h_1 \parallel g) = \lambda w_s . \text{the fish in } w$
 iii. $\llbracket \text{DP} \rrbracket (h_2 \parallel g) = \lambda w_s . \text{the meat in } w$
- e. $\llbracket \text{even DP} \rrbracket_{\langle a, \langle s, \langle a, \langle s, \langle \sigma, t \rangle \rangle \rangle \rangle} = \lambda g_a . \lambda w_s . \lambda Q_{\langle a, \langle s, \langle e, t \rangle \rangle} :$
 $\text{GEN } (g' \in G, h \in H) \left(h \neq h_0 \rightarrow \begin{array}{l} (\lambda w' . Q(g' \parallel h_0)(w')(\llbracket \text{DP} \rrbracket (g' \parallel h_0)(w'))) <_{\text{likely}} \\ (\lambda w' . Q(g' \parallel h)(w')(\llbracket \text{DP} \rrbracket (g' \parallel h)(w'))) \end{array} \right)$
 $Q(h_0 \parallel g)(w)(\llbracket \text{DP} \rrbracket (h_0 \parallel g)(w)) \text{ is true}$
 $= \lambda g_a . \lambda w_s . \lambda Q_{\langle a, \langle s, \langle e, t \rangle \rangle} :$
 $\text{GEN } (g' \in G, h \in \{h_1, h_2\}) \left(\begin{array}{l} (\lambda w' . Q(g' \parallel h_0)(w')(\text{the shrimp in } w')) <_{\text{likely}} \\ (\lambda w' . Q(g' \parallel h)(w')(\text{the } h(\text{F1})(g')(w') \text{ in } w')) \end{array} \right)$
 $Q(h_0 \parallel g)(w)(\text{the shrimp in } w) \text{ is true}$
- f. \sim inference from *even* in (331):
 $\approx \text{GEN } (g' \in G) \left(\begin{array}{l} (\lambda w' . \llbracket \text{TP} \rrbracket (g' \parallel h_0)(w')(\text{the shrimp in } w')) <_{\text{likely}} \\ (\lambda w' . \llbracket \text{TP} \rrbracket (g' \parallel h)(w')(\text{the fish in } w')) \wedge \\ (\lambda w' . \llbracket \text{TP} \rrbracket (g' \parallel h_0)(w')(\text{the shrimp in } w')) <_{\text{likely}} \\ (\lambda w' . \llbracket \text{TP} \rrbracket (g' \parallel h)(w')(\text{the meat in } w')) \end{array} \right)$
 $\iff \text{GEN } (g' \in G) \left(\begin{array}{l} (\lambda w' . \text{Hadas didn't eat the shrimp in } w') <_{\text{likely}} \\ (\lambda w' . \text{Hadas didn't eat the fish in } w') \wedge \\ (\lambda w' . \text{Hadas didn't eat the shrimp in } w') <_{\text{likely}} \\ (\lambda w' . \text{Hadas didn't eat the meat in } w') \end{array} \right)$
 $\iff (\lambda w' . \text{Hadas didn't eat the shrimp in } w') <_{\text{likely}} (\lambda w' . \text{Hadas didn't eat the fish in } w') \wedge$
 $(\lambda w' . \text{Hadas didn't eat the shrimp in } w') <_{\text{likely}} (\lambda w' . \text{Hadas didn't eat the meat in } w')$

This yields the correct scalar inference. The scale reversal behavior of *even* is reflected in this inference by the inclusion of the negation within the scope of *even*'s second argument.

Unfortunately this *even*-phrase based scope theory approach is not helpful for deriving the backwards association of *even*. For example, consider our basic example of backwards association with a subject (215), repeated here:

(333) A [professor]_F *even* came to the party. (=215)

At LF, the DP containing the F-marked material must move to become the first argument of *even*. However, in this and other cases of apparent backwards association, this movement would be *downward* and therefore banned under common assumptions regarding possible movement configurations. This hypothetical derivation is given below:


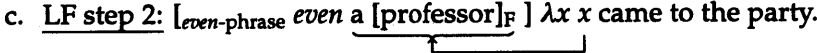
(334) Backwards association under the *even*-phrase approach requires downward movement:

LF: A [professor]_F [*even*] came to the party.

Assuming that movement at LF can only be upward, under this view the adverb *even* will only be able to associate with material which is in its surface scope.

We could also consider a variant where *even* first covertly moves to a higher position, and then moves its associate to be its first argument. This variant has the potential to explain the backwards association in (215), through the following derivation at LF:

(335) **Backwards association by moving *even* high, and then forming an *even*-phrase:**

- a. Narrow syntax: A [professor]_F *even* came to the party.
- b. LF step 1: *even* a [professor]_F _____ came to the party.

- c. LF step 2: [*even*-phrase *even* a [professor]_F] λx x came to the party.


Unfortunately, once *even* moves alone to a higher position (LF step 1), we expect *even* to be able to attract any constituent in its new scope to be its first argument. This option therefore suffers from the problems which were presented in section 6.2.1 for the version of the scope theory which moves just the adverb *even*.

The ability for *even* to associate with material which has moved out of its scope remains unexplained by this variant of the scope theory which builds *even*-phrases.

6.3 Summary

It is well known that *even* interacts with downward-entailing operators in a way that other focus-sensitive operators do not. Namely, when *even* is the surface scope of a downward-entailing operator, its scalar inference is reversed. In this chapter I presented the two common approaches to this scale reversal behavior of *even* and its relation to the problem of backwards association. Under the lexical ambiguity view, the semantics of *even* is polarity-sensitive, but *even* is always interpreted in its surface position. Under the scope theory, *even* is proposed to covertly move out of the scope of the downward entailing operator. This covert movement, proposed to be unique to *even*, is able to explain the greater ability of *even* to associate backwards.

I showed that the scope theory of *even* is not able to explain the distribution of backwards association with *even*. I showed that one variety of the scope theory overgenerates backwards association, while another undergenerates. In contrast, I showed that my proposal, where *even* is interpreted in its pronounced position, but can associate with the lower copy of F-marking in its scope, together with the lexical ambiguity theory, predicts the correct distribution of backwards association in downward-entailing environments.

Appendix: Nakanishi's (2012) argument for the scope theory

In this chapter I've argued for the adoption of the lexical ambiguity theory of *even*'s scale reversal behavior, and presented arguments against the scope theory of *even*. In this appendix I will discuss a set of complex data presented in Nakanishi (2012) as an argument against the lexical ambiguity view and in support of the scope theory. I will show that my proposal for backwards association is able to defuse a subset of the data making up this argument, but not all of it.

Nakanishi's argument comes from examples involving antecedent-contained deletion (ACD), a variety of VP ellipsis where the ellipsis gap is, on the surface, contained within the antecedent VP. Let's first discuss the interpretations and derivations of the ACD example in (336) below. The VP ellipsis gap indicated by a Δ in (336) can be interpreted as either "lift" or "fail to lift."

(336) Antecedent-contained deletion with two possible ellipsis interpretations:


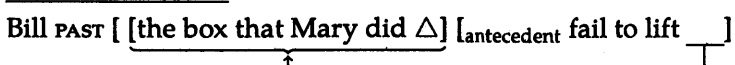
Bill [_{VP1} failed to [_{VP2} lift [the box that Mary did Δ]]].

- a. $\checkmark \Delta = \text{"lift"}$
- b. $\checkmark \Delta = \text{"fail to lift"}$

The question is how these interpretations for (336) are possible. VP ellipsis requires a semantically identical antecedent VP. The readings in (336) seem to correspond to the two VPs labeled VP1 and VP2, but these VPs themselves include the ellipsis gap, and therefore cannot be the antecedent for ellipsis.

As discussed in section 4.6, I along with Nakanishi (2012) follow the common view, first introduced by Sag (1976) and May (1985) and more recently argued for by Kennedy (1997) and Fox (2002), that ACD resolution uses QR to construct an appropriate antecedent VP which does not contain the ellipsis site. (337) below gives a possible LF for each of the readings in (233). The object DP "the box that..." which contains the ellipsis site, can be QRed to different heights. In order to create an antecedent for VP ellipsis which does not itself contain the ellipsis site, the DP must be QRed to a position at least as high as the edge of the intended antecedent VP. Therefore the QR is required to be higher in (337b), to create the antecedent VP "fail to lift," than in (337a), to create the antecedent VP "lift."

(337) LFs for each reading of (233):

- a. $\checkmark \Delta = \text{"lift"}$:
 Bill failed to [[the box that Mary did Δ] [_{antecedent lift}]]

- b. $\checkmark \Delta = \text{"fail to lift"}$:
 Bill PAST [[the box that Mary did Δ] [_{antecedent fail to lift}]]


The larger and smaller antecedent VPs are possible in (336) in part because both are possible antecedents for a VP ellipsis site in the complement of “did.” We can modify (336) so that the two antecedent VPs involve different auxiliaries, so that only one interpretation is possible at a time. This is illustrated through the following examples:

(338) ACD with one possible interpretation at a time, depending on the auxiliary:

- a. Bill has [_{VP1} failed to [_{VP2} lift [the box that Mary *did* Δ]]].
 - i. $\checkmark \Delta$ = “lift”
 - ii. $* \Delta$ = “fail to lift”
- b. Bill has [_{VP1} failed to [_{VP2} lift [the box that Mary *has* Δ]]].
 - i. $* \Delta$ = “lift”
 - ii. $\checkmark \Delta$ = “fail to lift”

Because VP1 is in the perfect, it must be the complement of the perfect auxiliary “have.” When the default auxiliary “did” is used adjacent to the ellipsis gap, the elided VP is unambiguously interpreted as “lift,” corresponding to VP2. When the auxiliary “have” is used, the elided VP is unambiguously interpreted as “fail to lift,” as in VP1. Because of the antecedent containment problem, the ellipsis resolution in (338b) must involve a higher QR of the object, as in (337b), than for (338a), as in (337a).

We can now turn to the data presented in Nakanishi (2012). I will concentrate on one example, reproduced here as (339) below. (“Failed” is bolded here as a downward-entailing operator, which will become important below.)

(339) Nakanishi’s (2012) example (37) of scale-reversed *even* with ACD:¹²⁹

Mary tried to lift the piano, the desk, and the box, but couldn’t lift any of them. Bill said that he can lift all of them. However, he has failed to lift the piano that Mary has failed to lift, and has also failed to lift the desk that she has failed to lift. Moreover, he has failed to *even* lift the [_{box}]_F that she has Δ .

Let’s first consider the interpretation of *even* in the last sentence in (339). The inference of *even* here is that boxes are more likely or easier to lift than other items under discussion, such as the piano and desk. This makes the use of *even* in the last sentence felicitous in the context. This is

¹²⁹Nakanishi (2012, ex 37) represents the F-marking in this example as on “the box,” to the exclusion of the relative clause, which is not a constituent. I believe the correct analysis of the intended example is with F-marking on the head noun “box” alone.

the scale-reversed interpretation of *even*, which is possible here because the sentence has an NPI-licensing main verb, “fail.” Under the lexical ambiguity view, the correct interpretation of *even* comes from the use of *even*_{NPI}, which must be in the scope of the NPI-licensing main verb, “fail.”

The problem comes from the QR step involved to construct the correct antecedent VP. The choice of auxiliary adjacent to the ellipsis gap in (339) forces the larger ACD resolution, “fail to (even) lift.”¹³⁰ In order to undo the antecedent containment of the ellipsis gap, the object “the box...,” including the focused constituent, must move out of the antecedent VP. This is schematized in (340) below.

- (340) LF for (339) necessary for ellipsis resolution: (based on Nakanishi, 2012, ex. 38b)
 He has_[PERF] [[the _{[box]_F] that she has_[PERF] Δ] _[antecedent fail to *even*_{NPI} lift]]}
-

Given the requirement that focus-sensitive operators take their focus associate in their scope (the “c-command requirement” below), Nakanishi (2012) argues that this and similar data pose a problem for the lexical ambiguity theory and its assumption that *even* is interpreted in its surface position:

“The lexical theory would say that *even* in the last sentence in (339) is an NPI, in which case *even* has to stay in the scope of *fail* at LF, as in (340). However, in this LF *even* cannot c-command the focus, and thus we would have to abandon the c-command requirement. Alternatively, it might be possible to assume that *even* in (340) is associated with the trace that is focused (cf. Rullmann, 2003). However, note that it is not the entire QR-ed NP that is focused, but only part of it. It is not clear how we can obtain the intended focus association from the assumption that the trace of the entire NP is focused.”

— Nakanishi (2012) (example numbers modified)

My proposal offers a way to do what Nakanishi dismisses—to “associate with the trace”—different from Rullmann’s (2003) proposal that she alludes to.¹³¹ Under my proposal, the base position of movement includes a copy of the F-marked constituent “box.” *Even* is able to associate with this F-marked constituent in its scope, leading to a correct inference. Here I will follow Fox (2002) in late-merging the gap-containing relative clause after the object “the box” has been QR-ed.

¹³⁰Whether *even* is part of the interpreted ellipsis site is a separate issue here, which is empirically difficult to detect. I will assume here for discussion that the ellipsis resolution is “fail to *even* lift.”

¹³¹Nakanishi’s criticism of the Rullmann (2003) approach of associating with an F-marked trace, though quick, is ultimately correct for *even*. See section 5.2 and the appendix to chapter 5 for discussion.

(341) Derivation and interpretation of (339) under my proposal:

a. QR of the object "the box":

Bill has[**PERF**] [[the [box]_F] [antecedent failed to *even*_{NPI} [_{vP} PRO₁ lift [the [box]_{F1}]]]]

b. Late merger of the relative clause (adjunct):

Bill has[**PERF**] [[the [box]_F [that Mary has[**PERF**] Δ]]₂
[antecedent failed to *even*_{NPI} [_{vP} PRO₁ lift [the [box]_F]₂]]]

c. LF after Trace Conversion:

Bill has[**PERF**] [[the [box]_F [that Mary has[**PERF**] Δ]]
λ₂ [antecedent failed to *even*_{NPI} [_{vP} PRO₁ lift [the [box]_F ①]]]]

d. $H = \{h_0, h_1, h_2\}$

i. $h_0(F1) = \llbracket \text{box} \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a box in } w$

ii. $h_1(F1) = \llbracket \text{desk} \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a desk in } w$

iii. $h_2(F1) = \llbracket \text{piano} \rrbracket = \lambda g_a . \lambda w_s . \lambda x_e . x \text{ is a piano in } w$

e. $\llbracket vP \rrbracket (h \parallel g) = \lambda w_s : g(2) \text{ is a } h(F1) . g(1) \text{ lifts } g(2) \text{ in } w$

f. ~ inference of *even*_{NPI}:

$\text{GEN} (g' \in G, h \in H) (h \neq h_0 \rightarrow (\lambda w . \text{LA}(\llbracket vP \rrbracket (g' \parallel h_0)(w))) >_{\text{likely}} (\lambda w . \text{LA}(\llbracket vP \rrbracket (g' \parallel h)(w))))$

$\Leftrightarrow \text{GEN} (g' \in G) \left(\begin{array}{l} (\lambda w . g'(2) \text{ is a } h_0(F1)(g')(w) \text{ in } w \wedge g'(1) \text{ lifts } g'(2) \text{ in } w) >_{\text{likely}} \\ (\lambda w . g'(2) \text{ is a } h_1(F1)(g')(w) \text{ in } w \wedge g'(1) \text{ lifts } g'(2) \text{ in } w) \wedge \\ (\lambda w . g'(2) \text{ is a } h_0(F1)(g')(w) \text{ in } w \wedge g'(1) \text{ lifts } g'(2) \text{ in } w) >_{\text{likely}} \\ (\lambda w . g'(2) \text{ is a } h_2(F1)(g')(w) \text{ in } w \wedge g'(1) \text{ lifts } g'(2) \text{ in } w) \end{array} \right)$

$\Leftrightarrow \text{GEN} (x, y \in D_e) \left(\begin{array}{l} (\lambda w . x \text{ is a box in } w \wedge y \text{ lifts } x \text{ in } w) >_{\text{likely}} \\ (\lambda w . x \text{ is a desk in } w \wedge y \text{ lifts } x \text{ in } w) \wedge \\ (\lambda w . x \text{ is a box in } w \wedge y \text{ lifts } x \text{ in } w) >_{\text{likely}} \\ (\lambda w . x \text{ is a piano in } w \wedge y \text{ lifts } x \text{ in } w) \end{array} \right)$

The scalar inference of *even* computed in (341) requires that, for a generic person y , it is *more likely* that y lift an object which is a box than to lift an object which is a desk or a piano. This inference is satisfied in the given context. This lets us interpret (339) with *even* interpreted its surface position using the lexical ambiguity theory, associating successfully with the F-marked "box" in the lower copy of the moved object. This is made possible by my analysis which is able to capture backwards association with *even*. As discussed in section 4.6 above, association with *even* does not restrict the height of QR and, therefore, size of ACD resolution. The derivation above acts as an existence

proof that a derivation for (339) is possible with the lexical ambiguity view, and does not force us to adopt the scope theory of *even*, contrary to Nakanishi's (2012) claims.

Unfortunately this counterargument to Nakanishi (2012) does not extend to the full set of data presented by her. That is, the approach I demonstrated above in (341) for example (339) does not work for a class of examples that she presents. The relevant difference has to do with the position of F-marking. In example (339) above, which I am able to capture using my proposal and the lexical ambiguity theory of *even*, the F-marking is on the *head noun* of the object. Following Fox's (2002) analysis of ACD, this F-marking will be present in the base position of the DP, within the scope of *even*. However, a number of the examples Nakanishi (2012) uses in her argument have F-marking in the *relative clause* adjoined to the object. Under Fox's (2002) analysis of ACD, this relative clause is late-merged to the QRed object, outside of antecedent VP, and therefore the intended F-marking will not be included in the scope of *even*. An example of this form from Nakanishi (2012) is reproduced here as (342):

- (342) **Example (36) from Nakanishi's (2012) argument, with F-marking in the relative clause:**¹³²
 Joe always tries to solve every problem that other people try to solve. He is trying to solve every problem that his classmate is trying to solve, and he is also trying to solve every problem that his tutor is trying to solve. Moreover, he_i is trying to *even* solve every problem that his_i [supervisor]_F is Δ .

Nakanishi (2012) argues that the final sentence in (342) is grammatical. The ellipsis gap must be interpreted as "trying to (*even*) solve," requiring QR of the object to a position outside of the antecedent VP. Assuming Fox's (2002) approach to ACD, the grammaticality of this derivation is unexpected by my approach, where *even* is interpreted in its surface position and does not move, and therefore cannot contain the F-marked constituent in its scope. In this way, my counterargument to Nakanishi (2012) only applies for a subset of the data. I will leave the final word on these ACD constructions with *even* for future research.

¹³²As noted for example (339) above, I again disagree with Nakanishi (2012) on the extent of the F-marking in the test sentence. Nakanishi gives this example with "his supervisor" F-marked, but given this context I believe the F-marked constituent is "supervisor."

Chapter 7

Reconstruction and association

In this chapter, I return to the relationship between reconstruction and focus association. This issue was considered briefly in section 4.2. Moved DPs are able to reconstruct in English—getting interpreted as if they are in their base positions—but this reconstruction is unable to feed focus association. I argued there that a common model for scope reconstruction—syntactic reconstruction—therefore cannot be correct for English as it incorrectly predicts configurations of backwards association with *only*. (This argument will be reviewed and expanded on below.)

In this chapter I revisit this issue and show that a more nuanced view is necessary. I begin in section 7.1 by taking a closer look at the role of reconstruction in cases of association with F-marking in the restrictor of a moved DP. This will confirm what I have argued before, that backwards association with *even* does not force scope reconstruction and that scope reconstruction does not feed focus association with *only*. Then in section 7.2 I show that there *are* instances where focus association is fed by reconstruction. These are cases of exceptional backwards association with quantificational determiners, which I showed in section 5.1 to generally not be possible: I will show that in contexts where reconstruction of the DP is independently licensed, this can feed backwards association with the determiner.

This puzzling state of affairs will motivate a new mechanism for syntactic reconstruction—Inverse Trace Conversion—which I present in section 7.3. I will claim that reconstruction in English must use Inverse Trace Conversion, and that total syntactic reconstruction by deletion of the higher copy at LF is not an option in English. Finally, in section 7.4, I will discuss German and Dutch which differ from English in allowing reconstruction to feed focus association of F-marking in the restrictor with *only*. I will argue that in such cases, German and Dutch utilize total syntactic reconstruction, which is unavailable in English.

7.1 Reconstruction and association with F-marking in the restrictor

I begin in this section by considering cases with F-marking in the restrictor of a DP moved out of the scope of the focus-sensitive operator. This was the configuration considered in chapter 4. I showed there that *even* but not *only* is able to associate “backwards” with the F-marked constituent in such configurations. I proposed that this is due to *even* associating with the F-marked constituent in the lower copy of the DP, using the Copy Theory of movement. The same configuration with *only* was shown to yield an unusable meaning. This LF is schematized in (343).

(343) **Associating with F-marking in the lower copy:** (=168)

- a. Narrow syntax: [DP Quantifier ... α_F ...] ... [Op ... [DP Quantifier ... α_F ...] ...]
- b. LF: [DP Quantifier ... α_F ...] λx ... [Op ... [DP the ... α_F ... x] ...]

In section 4.2, I briefly considered the possibility of syntactic reconstruction being the mechanism by which *even* is able to associate backwards in this configuration. The idea was that, when *even* is apparently associating “backwards” with material which has moved out of its scope, the moved DP is interpreted in a lower position, within the scope of *even*. This possibility is schematized in (186), repeated below as (344).

(344) **Backwards association by syntactic reconstruction:** (=186)

- a. NS/PF: [DP Quantifier ... α_F ...] ... [Op [...] ...]
- b. LF: ... [Op [... [DP Quantifier ... α_F ...] ...]] allowing *Op* to associate with α

In this section I revisit and expand upon this discussion. What we learn here will be the same as the lessons from section 4.2: that *even*'s ability to associate backwards is independent of the LF scope of the DP, and that scope reconstruction does not feed focus association with *only*. In addition to syntactic reconstruction, I will also consider two other proposals for scope reconstruction: PF movement and semantic reconstruction. The conclusion will be that none of these three common approaches to reconstruction is able to derive the full set of backwards association patterns.

Let's begin with cases of backwards association with *even*, where I will present an argument for my proposal over the forced reconstruction alternative. Under my proposal, backwards association places no restriction on the interpreted position of the quantificational DP. In contrast, reconstruction as schematized in (344) predicts that in cases of backwards association with *even*, the DP containing F-marking will take scope in its lower position within the scope of *even*.

Here I show that backwards association with *even* is independent of the DP's interpreted LF position. Consider the sentence in (190). Here we are able to interpret this sentence with *even*

associating with the predicate “student” in the leftward subject. Crucially, (190) is compatible both with surface scope and inverse scope between the universal subject and negation. I assume that inverse scope of universal quantifiers in subject position under negation is due to reconstruction of the subject into its *vP*-internal position (McCloskey, 1997; Sauerland, 2003).

(345) **Subject association with *even* is compatible with different scopes for the subject: (=190)**

Every [student]_F didn't *even* come to the party. $\checkmark \forall > \neg, \checkmark \neg > \forall$

a. $\checkmark \forall > \neg$: \Rightarrow No student came.

We planned a party with free food. We expected at least some students to come, because they love free food, but we didn't particularly expect professors to come. No one attended the party. We weren't surprised that every professor didn't come ($\forall > \neg$). What surprised us, though, is that every [student]_F didn't *even* come to the party.

b. $\checkmark \neg > \forall$: \Rightarrow Not every student came, but some may have.

We planned a party with free food. We expected all the students to come, because they love free food, but we didn't expect all the professors to come. Not all of the students came and not all of the professors came. We weren't surprised that every professor didn't come ($\neg > \forall$). What surprised us, though, is that every [student]_F didn't *even* come to the party.

The two scopes in (190) show that the possibility of *even* associating with “student” in the subject is independent of the scope of the quantificational subject, and therefore that the association of *even* with material in the leftward subject does not force scope reconstruction of the subject into its base position. But the availability of (190b) also teaches us that backwards association configurations are not incompatible with scope reconstruction.

I will now verify that both of these readings in (190) can be derived through my proposal. Note that *even* is interpreted in a downward-entailing context, and therefore its scalar inference will be reversed, as discussed in section 6.1. I will now show that the lexical ambiguity theory of *even*, defended in section 6.1, correctly derives the scalar inferences for *even* in both readings in (190).

I begin with the surface scope reading, (190a). The narrow syntax derivation and LF representation for this reading are given below.

(346) **Surface scope interpretation for (190):**

a. Narrow syntax: [Every [student]_F didn't *even* [[every [student]_F] come to the party]

b. LF after TC: [Every [student]_F $\lambda x . \neg$ *even*_{NPI} [[the [student]_F x] come to the party]

$\Rightarrow \forall > \neg$ assertion

c. *even*_{NPI}: \sim GEN x ((x is a student and came to the party) $>_{likely}$ (x is a professor and came to the party))

Because *even* computes its scalar inference based on the F-marked material in its scope (section 4.4), *even* can associate with the F-marked predicate “student” in (346). As *even* is interpreted in its surface position, within the scope of the NPI-licensing negation, the scale-reversed $even_{NPI}$ is used. This yields the correct scalar inference for *even*, as well as the correct truth conditions with the universal quantifier taking scope over negation.

Next I derive the inverse scope reading, (190b). Here I will first consider syntactic reconstruction, where the higher copy of the chain is deleted or ignored at LF, and the lower copy does not undergo Trace Conversion. Note that in section 4.2, when first discussing the issue of reconstruction and association, I dismissed the possibility of syntactic reconstruction. As I am reconsidering that conclusion in this chapter, I will (re)consider the syntactic reconstruction option here (as well as the similar PF movement option, below).

The derivation for the inverse scope reading based on syntactic reconstruction is shown below:

(347) **Inverse scope interpretation for (190) through syntactic reconstruction:**

- a. Narrow syntax: $\underbrace{[\text{Every } [student]_F \text{ didn't } even \text{ } [every \text{ } [student]_F \text{ come to the party}]]}_{\text{scope}}$
- b. Ignore higher copy at LF:
 $\{[Every \text{ } [student]_F\} \text{ NEG } even_{NPI} \text{ } [every \text{ } [student]_F \text{ come to the party}]$
 $\iff \text{ NEG } even_{NPI} \text{ } [every \text{ } [student]_F \text{ come to the party}] \quad \Rightarrow \neg > \forall \text{ assertion}$
- c. $even_{NPI}$: $\sim \rightarrow$ (every student came to the party) $>_{likely}$ (every professor came to the party)

In this derivation, there is only one instance of F-marking at LF, which is in the scope of *even*, and therefore syntactic reconstruction has turned this case of apparent backwards association into a standard case of *even* associating with F-marking in its complement. Here too, “even” is within the scope of negation and is therefore interpreted as $even_{NPI}$, deriving the correct scalar inference.

A similar option is a derivation where the subject moves at PF, instead of in the narrow syntax (Sauerland, 1998; Sauerland and Elbourne, 2002). Such movement will feed linearization, but not interpretation, resulting in the illusion of scope reconstruction. For our purposes, PF movement will be identical to what is derived by syntactic reconstruction, and therefore is a possible derivation for this reading:

(348) **Inverse scope interpretation for (190) through PF movement:**

- a. Narrow syntax: $\text{NEG } even \text{ } [every \text{ } [student]_F \text{ come to the party}]$
- b. PF: $\underbrace{[Every \text{ } [student]_F \text{ didn't } even \text{ } [every \text{ } [student]_F \text{ come to the party}]]}_{\text{scope}}$
- c. LF = NS: $\text{NEG } even_{NPI} \text{ } [every \text{ } [student]_F \text{ come to the party}] \quad \Rightarrow \neg \forall \text{ assertion}$
- d. $even_{NPI}$: $\sim \rightarrow$ (every student came to the party) $>_{likely}$ (every professor came to the party)

Unfortunately, these mechanisms of syntactic reconstruction and PF movement both incorrectly predict that backwards association with *only* should be possible when the DP is reconstructed. Consider example (191), repeated here as (349).

(349) **Scope reconstruction does not feed focus association with *only*: (=191)**

* Every [professor]_F will *only* come to the party.

Intended: \approx *Only* every [professor]_F will come to the party.

I will first show how syntactic reconstruction predicts backwards association to be possible in (349). The higher copy of the subject will be ignored at LF, resulting in an LF representation with only one instance of F-marking, which is in the scope of *only*. This predicts *only* to be able to associate with the F-marked constituent.

(350) **Syntactic reconstruction predicts backwards association with *only* to be possible:**

a. Narrow syntax: [Every [professor]_F] will *only* [[every [professor]_F] come to the party]

b. Ignore higher copy at LF:

{every {professor}_F} will *only* [[every [professor]_F] came to the party]

\Leftrightarrow *only* [every [professor]_F came to the party]

\Rightarrow predicts association to be possible, contrary to fact (349) 🧑

Now consider the PF movement approach. In this case, the F-marked predicate “professor” in the subject never leaves the scope of *only* in narrow syntax or LF, predicting backwards association with *only* to be possible.

(351) **PF movement predicts backwards association with *only* to be possible:**

a. Narrow syntax: *only* [every [professor]_F came to the party]

b. Move subject at PF: [Every [professor]_F] *only* [___ came to the party]

\Rightarrow “Every [professor]_F *only* came to the party.”

c. LF = NS: *only* [every [professor]_F came to the party]

\Rightarrow predicts association to be possible, contrary to fact (349) 🧑

One concern here may be that the example in (349) is not exactly parallel to our example with *even* in (190), as (190) additionally includes sentential negation. The examples below show that


the addition of sentential negation to (349), either above or below *only*, does not change the lack of backwards association with *only*.¹³³

- (352) a. * Every [professor]_F will not *only* come to the party.
 b. * Every [professor]_F will *only* not come to the party.

If syntactic reconstruction or PF movement were generally possible, we predict scope reconstruction to be able to feed focus association with *only*, contrary to fact. Furthermore, the readings of example (190) with *even* show that it is not the case that DPs containing F-marking cannot reconstruct into the scope of a focus-sensitive operator. I therefore concluded in section 4.2 that syntactic reconstruction cannot generally be available in English. This same conclusion holds for PF movement based on the discussion here.

Let's now consider a different approach to scope reconstruction: semantic reconstruction. Under this approach, the base position of the movement chain is interpreted using a high-type trace, which is then abstracted over by a correspondingly high-type λ -binder (Cresti, 1995; Rullmann, 1995, a.o.). In this case the base position of the movement chain will be completely replaced by the variable T_i —a variable T_i of the type of the moved quantifier, $\langle\langle e, t \rangle, t\rangle$.

(353) Interpreting the base position using a high-type trace leads to focus-insensitivity:

- a. LF: [Every [professor]_F]_i λi NEG *even* [_{vP} T_i came to the party]
 b. $\llbracket T_i \rrbracket = \lambda g_a . \lambda w_s . g(i)$
 c. $\llbracket vP \rrbracket$ is not focus-assignment-dependent \Rightarrow cannot derive backwards association 

This trace node will not include any information on the F-marking within the restrictor of the moved DP.¹³⁴ Therefore there will be no F-marked constituent in the complement of *even* at LF, and therefore *even* will not have any alternatives to use for its scalar inference. Semantic reconstruction can derive the correct truth conditions for the inverse scope reading of (190), but is unable to compute the inference of *even* at the same time.

To summarize, the evidence from quantifier scope presented here shows that backwards association of *even* with F-marking in the restrictor of a DP does not force reconstruction of the DP.

¹³³In particular, one might wonder whether the presence of sentential negation was essential for allowing the reconstruction in (190), as reconstruction of the quantifier under negation leads to a difference in relative scope interpretation. See discussion of Scope Economy (Fox, 1995) in section 7.2 below.

¹³⁴Formally, this comes down to the question of what the focus semantic value of the base position of the movement chain is. Because the high-type trace node T_i is not itself F-marked, its denotation will not be focus-assignment-dependent. To my knowledge, no attempt has ever been made to preserve syntactic material from the lower copy of movement, as in the Copy Theory with Trace Conversion (Fox, 2002), when semantic reconstruction is invoked.

This can be explained by my proposal, where the existence of the lower copy of F-marking in the scope of *even* is enough to allow for backwards association, if syntactic reconstruction or PF movement is used for scope reconstruction. However, I also showed that syntactic reconstruction and PF movement incorrectly predict that scope reconstruction will feed focus association with *only*.

The following table summarizes the observed patterns of association when the DP carrying F-marking scope reconstructs, and what is predicted under the theories of syntactic reconstruction/PF movement and semantic reconstruction.

(354) Observed and predicted patterns of association with reconstruction (first version):

Configuration at PF	Observed	Predicted	
		SynR/PF	SemR
[DP D ... α_F ...] ... <i>even</i> [...]	○	○	×
[DP D ... α_F ...] ... <i>only</i> [...]	×	○	×

Before presenting my solution to this puzzle, in the next section I will further complicate the empirical picture by showing that in certain cases, reconstruction *does* feed focus association.

7.2 Reconstruction feeds backwards association with determiners

The examples we discussed in section 7.1 above were cases of association with material in the restrictor of a moved DP. However, my proposal predicts a different interaction with reconstruction in cases of association with a quantificational determiner.

Consider example (259) from section 5.1, repeated here as (355):

(355) *Even* cannot associate backwards with a determiner: (=259)

- a. Of course we arrested some protesters. ✓We *even* arrested [every]_F protester.
- b. Of course some protesters were arrested. ??[Every]_F protester was *even* arrested.

In section 5.1, I explained the difference between *even*'s ability to backwards associate with F-marking in the restrictor of a DP, and its inability to backwards associate with F-marking on the determiner. The lower copy in a copy theoretic movement chain must be modified for interpretation using Trace Conversion. The Determiner Replacement step of Trace Conversion will replace the lower copy of the determiner "every" with a definite article, and, as a consequence, eliminate the F-marking in the lower copy. Since there is no F-marking in the lower copy, the Focus Assignment Dependency Requirement (43) is violated and *even* is unable to find an associate.

(356) **Hypothetical derivation of (259b): (=262)**

- a. Narrow syntax: $[[\text{Every}]_F \text{protester}]_{\text{F}} \text{PAST } \textit{even} \underbrace{[[\text{every}]_F \text{protester}]_{\text{F}}}_{\uparrow} \text{BE arrested}$
- b. LF after TC: $[[\text{Every}]_F \text{protester}] \lambda x \text{PAST } \textit{even} [{}_{vP} \text{the protester } x] \text{BE arrested}$
- c. $[{}_{vP}]$ is not focus-assignment-dependent \Rightarrow predicts no backwards association \odot

In our previous discussion of this example, the possibility of syntactic reconstruction was not discussed. However, if the subject in the derived Spec,TP position is able to undergo syntactic reconstruction into the lower Spec,*v*P position at LF, we might expect a resulting structure where the F-marked node “every” is in the scope of *even*, and is not destroyed through Trace Conversion.

At first glance, the marked status of (355b) seems to teach us that this reconstruction of the subject is not possible. However note that the inability to scope reconstruct—and thereby feed focus association—in (355b) is independently explained by Fox’s (1995) Scope Economy condition. Fox (1995) proposes that covert scope-shifting operations such as Quantifier Raising and Quantifier Lowering (here described uniformly as scope reconstruction) are subject to an *Economy* condition: these syntactic operations are not possible if they do not introduce new scope possibilities.

Some complex examples are necessary in order to see the effects of Scope Economy on reconstruction. Consider example (357) below, which has an inverse scope reading. One possible source of this inverse scope reading is through a combination of scope reconstruction of the subject to its Spec,*v*P base position and QR of the object to the edge of *v*P above the subject’s interpreted position (see Johnson and Tomioka, 1998).

(357) **One path to inverse scope: QR of the object and reconstruction of the subject**

A guard is $[{}_{vP}$ standing in front of every church]

Now consider the examples in (358) from Fox (1995). We assume that in these examples, the subject has moved across-the-board from the Spec,*v*P position in each of the conjuncts. Example (358a) has an inverse scope reading, where every church has a guard in front of it and every mosque has a guard in front of it. This reading cannot be generated by QRing the objects “every church” and “every mosque” above the surface position of the subject, as that would trigger a violation of the Coordinate Structure Constraint, among other problems. This inverse scope must have instead been derived by two, parallel applications of scope reconstruction of the subject with QR of the object above it, schematized in (357). Fox (1995) notes that this inverse scope reading is made unavailable if *just one* of the conjuncts is modified so that scope reconstruction into it does not change scope relations. This is done in example (358b) by changing the second conjunct’s object from the scope-bearing “every mosque” to the non-scope-bearing “this mosque.” Assuming that

scope reconstruction into a coordinate structure, like the original overt movement itself, must be across-the-board, Scope Economy explains the lack of an inverse scope reading between “a guard” and “every church” in (358).

(358) **A contrast in across-the-board scope reconstruction (Fox, 1995, p. 320):**

- a. A guard is [standing in front of every church] and [sitting at the side of every mosque].
 $\checkmark \forall > \exists$
- b. # A guard is [standing in front of every church] and [sitting at the side of this mosque].
 $*\forall > \exists$

Returning now to example (355b), reconstruction of the subject into its lower Spec, *v*P position does not affect the truth-conditional interpretation of the sentence. Therefore, Scope Economy predicts that the subject is interpreted in its surface position at LF, outside of the scope of *even*.

A prediction of this Scope Economy view is that backwards association of *even* with a determiner (as in 355b) will become possible if scope reconstruction of the subject is independently licensed by Scope Economy. This is supported by the contrast in (359) below. Example (359a) is a baseline similar to (355b), but embedded in a higher clause possibility modal. Here backwards association with the determiner “every” is judged as very difficult. Note that reconstruction of the subject “every student” into a position within the scope of *even* does not yield a new scope relation. In example (359b), the existential modal “could” is placed in the same clause as the subject and within the subject’s surface scope. The addition of this modal allows reconstruction of the subject to make a truth-conditional difference, leading to an inverse scope reading “could > \forall .” Speakers judge the backwards association of “every” with *even* to be grammatical in this example, with the assertion interpreted with the scope relation “could > \forall .”

(359) **Backwards association with the determiner improves with a modal:**

Context: We’re expecting a large turnout at this party.

- a. * It’s possible that [every]_F student will *even* come.
- b. \checkmark [Every]_F student could *even* come. could > \forall

Consider the derivation for (359b) using syntactic reconstruction, in (360) below. Here for illustration purposes I assume that the contextually-determined alternative to the determiner “every” is “some.”

(360) Syntactic reconstruction derivation for (359b):

- a. Narrow syntax: $[[\text{Every}]_F \text{ student}] \text{ could } \textit{even} \underbrace{[[[\text{every}]_F \text{ student}]]}_{\text{LF}}$ come to the party]
- b. Ignore higher copy at LF:
 $\{\{\text{every}\}_F \text{ student}\} \text{ could } \textit{even} [[[\text{every}]_F \text{ student}] \text{ came to the party}]$
 $\iff \text{ could } \textit{even}_{\text{PPI}} [[[\text{every}]_F \text{ student}] \text{ came to the party}]$
 \Rightarrow leads to “could > \forall ” scope, satisfying Scope Economy
- c. $\textit{even}_{\text{PPI}}$: $\sim ((\text{every student came to the party}) <_{\textit{likely}} (\text{some students came to the party}))$
 \Rightarrow predicts association to be possible \odot

Under syntactic reconstruction, the higher copy of the DP is ignored at LF and Trace Conversion is not used. Therefore the lower copy of the DP—the only copy at LF—is unmodified and thus retains the F-marking on its determiner. This allows for backwards association of *even* with the determiner. If the DP were not reconstructed, however, the determiner in the lower copy would be replaced in the process of Trace Conversion, and focus association would become impossible (section 5.1).

Reconstruction feeding focus association with the determiner is also possible using the PF movement approach to scope reconstruction (Sauerland, 1998; Sauerland and Elbourne, 2002). This variant is illustrated below:


(361) PF movement derivation for (359b):

- a. LF = NS: could $\textit{even}_{\text{PPI}} [[[\text{every}]_F \text{ student}] \text{ came to the party}]$
 \Rightarrow leads to “could > \forall ” scope, satisfying Scope Economy
- b. PF: $[[[\text{Every}]_F \text{ student}]] \text{ could } \textit{even} \underbrace{[[[\text{every}]_F \text{ student}]]}_{\text{LF}}$ come to the party]
 \Rightarrow “EVERY student could *even* come to the party.”
- c. $\textit{even}_{\text{PPI}}$: $\sim ((\text{every student came to the party}) <_{\textit{likely}} (\text{some students came to the party}))$
 \Rightarrow predicts association to be possible \odot

Under the PF movement approach, the DP with F-marked determiner never moved out of the scope of *even*, and therefore we predict that *even* will be able to associate with the determiner.

The ability to associate with the determiner in (359b) is not expected, however, using semantic reconstruction. In semantic reconstruction, the base position of movement is replaced with a high-type trace. The complement of *even* at LF will not include any F-marked material, predicting *even* to violate the Focus Assignment Dependency Requirement (43).

(362) **Interpreting the base position using a high-type trace leads to focus-insensitivity:**

- a. LF: $[[\text{Every}]_F \text{ student}]_i \lambda i \text{ could even } [{}_{vP} T_i \text{ come to the party}]$
- b. $[[T_i]] = \lambda g_a . \lambda w_s . g(i)$
- c. $[[vP]]$ is not focus-assignment-dependent \Rightarrow cannot derive backwards association 

The interaction of scope reconstruction (when licensed by Scope Economy) and backwards association with the determiner is explained by a syntactic reconstruction or PF movement approach to reconstruction, but is unexpected under the semantic reconstruction approach.

So far in this section I have discussed *even*, with scope reconstruction feeding association with the determiner. However, the explanation here in terms of syntactic reconstruction/PF movement feeding focus association with the determiner should apply regardless of the particular focus-sensitive operator involved. We therefore predict that scope reconstruction of the DP, for example under a modal as in (359b), could feed focus association with *only*.

I have found that for some speakers, there is a contrast which points in this direction, although there is substantial inter-speaker variation. For many but not all speakers consulted, the example in (363b) with backwards association across a modal is markedly better than (363a) without the modal.¹³⁵ The context in (363b) is designed to support the inverse scope reading with “can > only > \exists .” For these speakers, just as we saw with *even*, scope reconstruction licensed by Scope Economy is able to feed backwards association with the determiner.

(363) **Scope reconstruction feeding focus association of *only* with a determiner:**

- a. Context: The department traditionally funds all requests from students for research funds. But this year they received many more applications than in past years, and the department was unable to fund every proposal. When the news was announced, the students were disappointed that...
* $[\text{Some}]_F$ proposals had *only* been accepted.
- b. Context: The department traditionally funds all requests from students for research funds. But this year they received many more applications than in past years. At the faculty meeting, the administrator reported that not every proposal can be accepted this round...
% $[\text{Some}]_F$ proposals $\left\{ \begin{array}{l} \text{can} \\ \text{should} \end{array} \right\}$ *only* be accepted.

The finding here is surprising given that, in all of the examples considered thus far in this dissertation, *only* has *never* been able to associate backwards. In particular, in section 7.1 we explicitly

¹³⁵For all other speakers, both examples were judged as equally ungrammatical. No speakers judged both examples in (363) as grammatical.

considered the effect of scope reconstruction of DPs containing F-marking in their restrictors, and found that reconstruction does *not* feed focus association in such cases. This is the puzzle I intend to resolve in the following section.

Being parasitic on scope reconstruction means that this pattern of association is limited by independent constraints on scope reconstruction of the DP. Scope Economy (Fox, 1995), which states that covert scope shifting operations such as scope reconstruction cannot occur if they are semantically vacuous, can explain the contexts in which backwards association with a determiner is judged as grammatical. Furthermore, the pattern of association presented in this section teaches us an important characteristic of the notion of Scope Economy itself: covert scope-shifting operations can only be used when they make a difference in the truth conditions of the utterance, not if they simply make *any* interpretational difference. In particular, *feeding focus association is not sufficient for licensing a covert scope-shifting operation*.

The patterns of association possible with the F-marked *determiner* of a DP, discussed here, is importantly different from the patterns observed with F-marking in the *restrictor* of a DP, as in the previous section. If the DP has moved out of the scope of the operator, backwards association with F-marking in the restriction is always possible with *even* but always impossible with *only*, independent of whether the DP itself reconstructs. In contrast, if F-marking is on the DP's determiner, backwards association is possible if and only if the DP scope reconstructs.

These facts are summarized in the table below, together with predictions of different mechanisms for reconstruction, building on the table in (354) above. As I have discussed above, association with a reconstructed determiner is expected under syntactic reconstruction/PF movement, but is not predicted under semantic reconstruction.

(364) Observed and predicted patterns of association with reconstruction (expanded):

Configuration at PF	Observed	Predicted	
		SynR/PF	SemR
$[\underline{\text{DP D ...}\alpha_{\text{F}}\dots}] \dots \text{even} [\dots \text{---} \dots]$	○	○	×
$[\underline{\text{DP D ...}\alpha_{\text{F}}\dots}] \dots \text{only} [\dots \text{---} \dots]$	×	○	×
$[\underline{\text{DP D}_{\text{F}} \dots}] \dots \text{only/even} [\dots \text{---} \dots]$	○	○	×

To summarize, we observe that both syntactic reconstruction/PF movement and semantic reconstruction predict uniformity with respect to *even* and *only*'s ability to backwards associate with a DP that has moved out of their scope: either backwards association is always possible, regardless of whether association is with the determiner heading the moved DP or with part of its restrictor, or

backwards association is always impossible. Therefore, none of these existing approaches to scope reconstruction can explain the observed patterns of backwards association discussed here.

7.3 Reconstruction via Inverse Trace Conversion

In this section I propose a new mechanism for scope reconstruction at LF using the Copy Theory of movement, which I call *Inverse Trace Conversion*. This view of reconstruction is able to explain the data we have seen in this chapter, which we have seen cannot be explained by off-the-shelf approaches to scope reconstruction. I will also show that Inverse Trace Conversion has some other useful applications, but its use must also be constrained in order to derive the correlations between scope and binding reconstruction in Fox (1999).

In the discussion in the previous sections, we have seen that, in general, syntactic reconstruction/PF movement fares better than semantic reconstruction in terms of the predictions it makes with regard to patterns of focus association. Semantic reconstruction is unable to derive backwards association with F-marking in the restrictor of a reconstructed DP (section 7.1). We also saw that scope reconstruction is able to feed new patterns of focus association with determiners (section 7.2), which is unexpected by semantic reconstruction. However, as discussed first in section 4.2 above, syntactic reconstruction also predicts that scope reconstruction of a DP with F-marking in its restrictor will feed backwards association with *only*, contrary to fact.

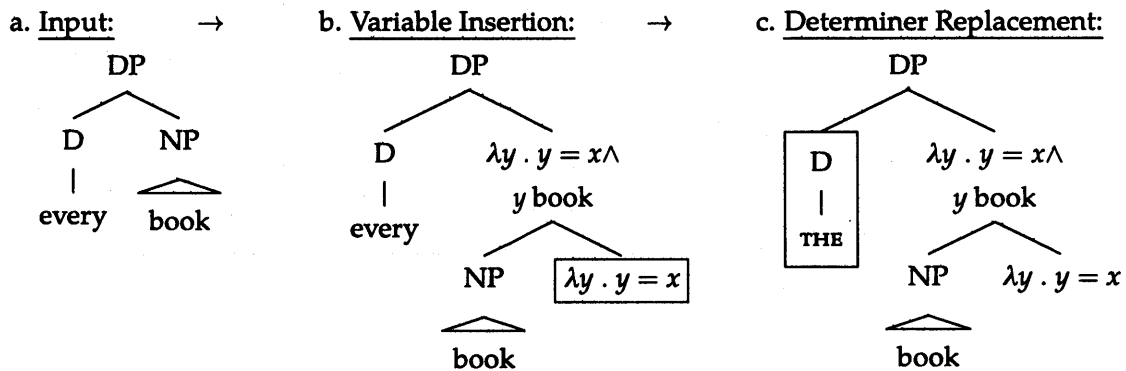
What we are looking for is a way to tweak syntactic reconstruction so that it feeds backwards association with the determiner (section 7.2), but preserves the regular pattern of focus association in cases of F-marking in the restrictor—backwards association possible with *even* but not *only*. It will be instructive, then, to briefly review my approach to backwards focus association with *even* but not *only* in chapter 4. I derive this contrast using the Copy Theory of movement, by having both copies in a movement chain active in the interpretation of the chain. When a lower copy is in the scope of *only* but a higher copy is not, the focus alternative substitution which happens as part of the semantics of *only* will result in conflicting requirements on the variable being quantified over, and hence this configuration across *only* will be ungrammatical. When a lower copy is in the scope of *even*, the F-marking in just the lower copy can be used for association by *even*. *Even* does not modify the assertion, so the assertive content will be able to compose above *even* without problems.

Inverse Trace Conversion (ITC) will be able to derive the patterns of focus association observed by using *both* copies of the restriction in the interpretation of the movement chain at LF. Importantly, none of the traditional approaches to reconstruction have this property: under PF movement, there is no copy at LF; under syntactic reconstruction, the higher copy is completely deleted or ignored

at LF; and under semantic reconstruction, the lower copy is completely replaced by a high-type variable. What I propose here, instead, is an approach to scope reconstruction which preserves this characteristic of both copies of the restrictor being interpreted.

My proposal will be very similar to Trace Conversion, the process by which lower copies of DP movement are converted into a type e bound variable which also includes an active restrictor (Rullmann and Beck, 1998; Fox, 2002). The individual steps involved in modifying the lower copy using Trace Conversion are repeated here from (53) above:

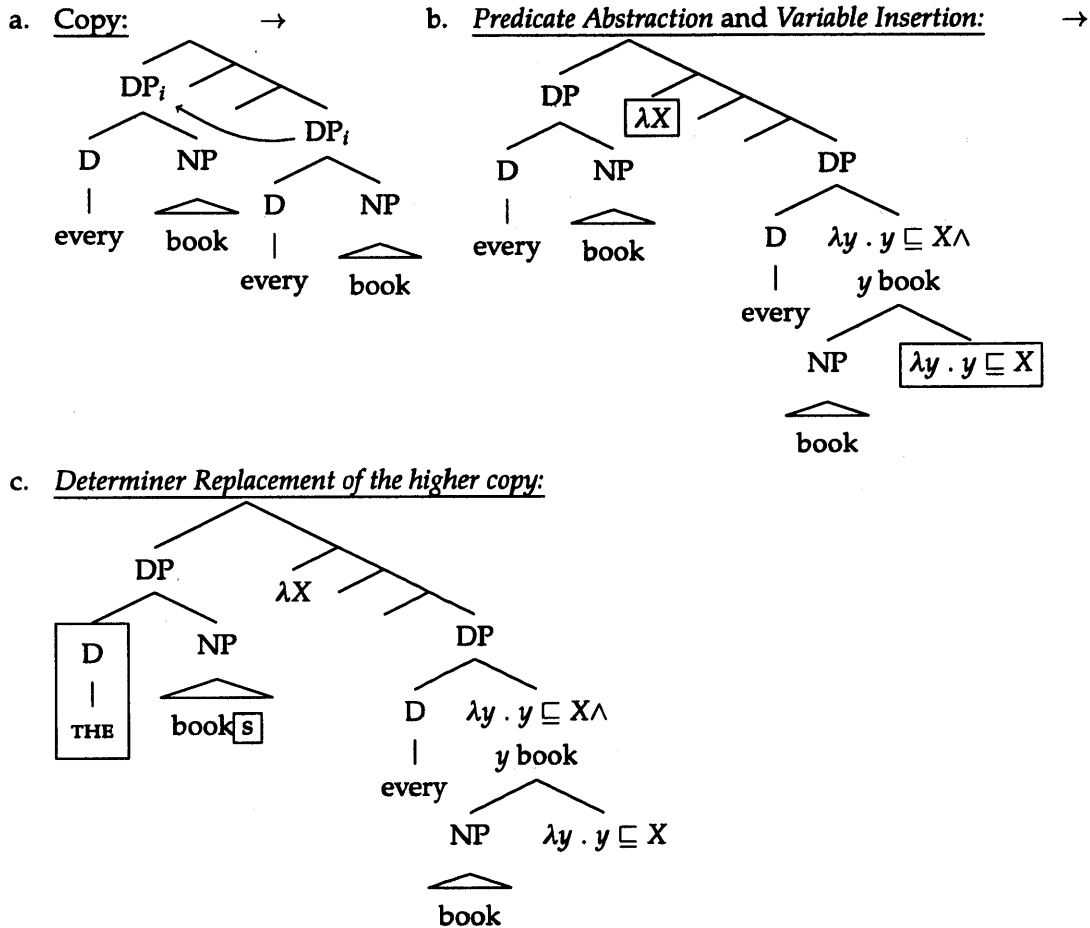
(365) Trace Conversion of the lower copy: (=53/258)



ITC can be thought of as a different application of the individual steps of Trace Conversion (365). The result will be that both copies of the restrictor are interpreted, but the quantificational material in the *lower* copy of movement is ultimately interpreted, unlike in Trace Conversion, where the *higher* copy of movement is interpreted at LF. Specifically, the higher copy will be converted into a plural definite description, corresponding to the *restrictor* of the lower quantifier.

This process is illustrated in (366) below. The starting point is a copy-based movement chain, here illustrated by movement of a QP “every book.” Predicate Abstraction inserts a λ -binder right below the higher copy of movement, just as in the normal interpretation of movement chains. Variable Insertion targets the lower copy, as in Trace Conversion, introducing the variable bound by the λ -binder. Here, in contrast to standard Trace Conversion, a part-of relation (\sqsubseteq) is used instead of equality to the variable. Finally, Determiner Replacement turns the *higher* copy into a definite description, and if it were originally singular, it is modified so that it will also range over plural individuals.

(366) Inverse Trace Conversion (basic, plural individual version):



I will first demonstrate how this proposal can derive a simple inverse scope reading. Consider the example in (367), where the subject universal is interpreted under negation via reconstruction. In the narrow syntax, the subject is copied from its *v*P-internal base position to Spec,TP. After ITC at LF, we yield an assertion which computes the set of “the professors” *X* and then claims that it is not the case that every professor in *X* came to the party. In this way, both the higher and lower copies of the restriction (here, “professor”) contribute to the interpretation, and the quantifier takes scope in the lower, base position of movement.

(367) **An Inverse Trace Conversion derivation for inverse scope:**

Every professor didn't come to the party.

 $\checkmark \forall > \neg, \checkmark \neg > \forall$ a. Narrow syntax: [every professor] _{NEG} [[every professor] come...]b. Pronounce higher copy at PF: "Every professor didn't come to the party."c. LF after Inverse Trace Conversion:[the professors] $\lambda X . \neg$ [every professor in X came to the party] $\neg > \forall$

One immediate problem with this formulation of ITC comes from the existence presupposition of the higher copy which is interpreted as a definite description. For example, in the derivation in (367) above, the definite description "the professors" will introduce the presupposition that there exist (relevant) professors. In the case of (367), this presupposition is unproblematic. But consider (368) below. The correct interpretation of example (368) involves scope reconstruction of the subject within the scope of negation and "seem." If this reconstruction uses ITC, the higher copy will introduce a presupposition that there exists at least one even prime number greater than two, contrary to fact.

(368) **The higher copy of ITC can introduce a problematic presupposition:**

✓ An even prime number greater than two does not seem to exist.

 $\neg > \text{seem} > \exists$

(Martin Hackl and Irene Heim, p.c.)

a. Narrow syntax: [An even prime number greater than two] _{NEG} seem

[[an even prime number greater than two] exist]

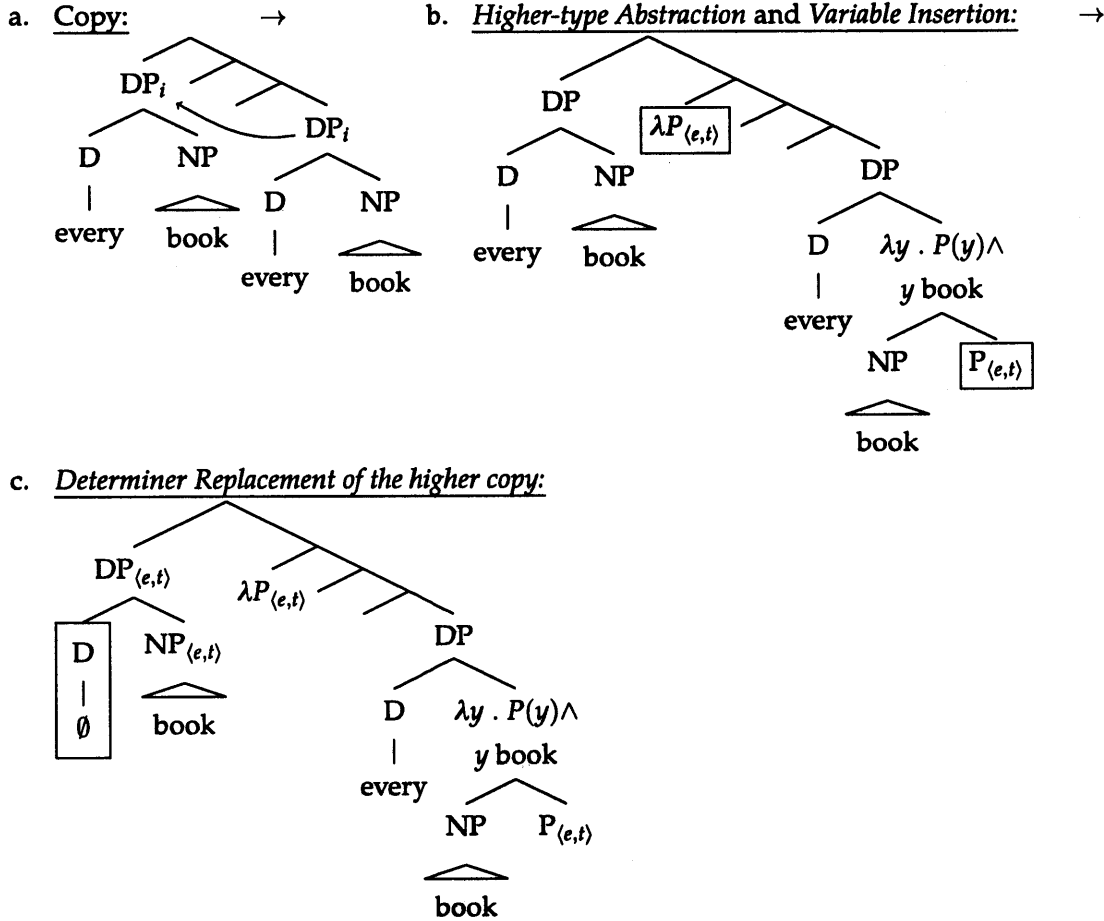
b. Pronounce higher copy at PF:

"An even prime number greater than two does not seem to exist."

c. LF after Inverse Trace Conversion:[The even prime numbers greater than two] $\lambda X . \neg$ seem [an even prime number greater than two in X exists] $\neg > \text{seem} > \exists$ d. Presupposition of the higher copy: $\exists x . x$ is an even prime number greater than two

One possible fix for this issue is to make the higher copy not a definite description returning a plural individual, but instead return the restrictor of the higher copy, as a *property* of superintensional type $\langle a, \langle s, \langle e, t \rangle \rangle \rangle$. The λ -abstractor will correspondingly abstract over a variable of extensional type $\langle e, t \rangle$ adjoined to the lower copy's restrictor. In the Determiner Replacement step, the higher copy's determiner will now be replaced by a semantically inert dummy (\emptyset below). This revised process of ITC is schematized below using extensional types, again using QR of "every book" for illustration.

(369) Inverse Trace Conversion (property version):



Both variants of ITC presented here have the characteristics that allow it to explain the distribution of reconstruction and focus association in English: reconstruction occurs, feeding backwards association with the determiner, but it does not feed association of *only* with F-marking in the restrictor. The important property of these ITC proposals is that the restrictor will be interpreted both inside and outside of the scope of *only*. *Only* attempting to associate with F-marking in the restrictor will lead to an uninterpretable structure. This follows the same logic as for cases where the higher copy of the quantifier is interpreted, using Trace Conversion on the lower copy, in section 4.3. This is demonstrated below with our example (349), using the basic ITC (366):

(370) ITC blocks *only* associating with F-marking in the restrictor (349):

- a. Narrow syntax: $\underbrace{[\text{every } [\text{professor}]_F] \text{ only } [[\text{every } [\text{professor}]_F] \text{ come...}]}$
- b. LF after Inverse Trace Conversion:
 $[\text{the } [\text{professors}]_F] \lambda X \text{ only } [[\text{every } [[\text{professor}]_F \text{ in } X]] \text{ came to the party}]$
 $\iff [\text{the professors}] \lambda X . \neg [\text{every } [\text{student in } X] \text{ came to the party}]$
 \Rightarrow predicted to be uninterpretable (see section 4.3)

The structure in (370) becomes uninterpretable because one copy of the restrictor is inside the scope of *only*, where its F-marked part will get replaced by alternatives, while the other copy is interpreted outside of the scope of *only*. In this example, the variable *X* being quantified over will be a set of professors, according to the higher copy, but in the lower quantifier, we are restricted to the students in *X*. See section 4.3 for further discussion based on the non-scope-reconstructed case.

In contrast, the same structure with *even* (371 below) will be interpretable since *even* computes its scalar inference on the basis of the content in the complement of *even*, and the computed inference does not compose with material above (see section 4.4 above).

(371) ITC allows *even* to associate with F-marking in the restrictor (190b):

- a. Narrow syntax: $\underbrace{[\text{every } [\text{student}]_F] \text{ NEG } \text{even} [[\text{every } [\text{student}]_F] \text{ come...}]}$
- b. LF after Inverse Trace Conversion:
 $[\text{the } [\text{students}]_F] \lambda X . \neg \text{even}_{\text{NPI}} [[\text{every } [[\text{student}]_F \text{ in } X]] \text{ come to the party}] \Rightarrow \neg > \forall$
- c. even_{NPI} : $\sim ((\text{every student comes to the party}) >_{\text{likely}} (\text{every professor comes to the party}))$
 \Rightarrow predicts correct scalar inference

Now consider the case of association with an F-marked determiner. We saw in section 7.2 that scope reconstruction of the DP, independently licensed by Scope Economy, is able to feed backwards association with an F-marked determiner. Because the interpreted quantifier is the quantifier in the lower copy of the DP, within the scope of the focus-sensitive operator, ITC predicts backwards association to be possible in this configuration with both *even* and *only*, as schematized here:

(372) ITC allows backwards association with the determiner:

a. Narrow syntax: $[[\text{Quantifier}]_F \text{ student}] \text{ MODAL } \textit{only/even} [[[\text{Quantifier}]_F \text{ student}] \dots]$

b. LF after Inverse Trace Conversion:

$[\text{the students}] \lambda X \text{ MODAL } \textit{only/even} [[\text{Quantifier}]_F [\text{student in } X] \dots]$

⇒ predicts association to be possible

The following table presents the various patterns of focus association that we have observed here, together with the predictions made by a grammar utilizing ITC. If scope reconstruction in English always involves a variety of ITC, then, we correctly predict the distribution of focus association in all cases of scope reconstruction.

(373) Observed and predicted patterns of association and reconstruction, using ITC:

Construction at PF	Observed	Predicted by ITC
$[\text{DP } D \dots \alpha_F \dots] \dots \textit{even} [\dots \text{ } \dots]$	○	○ (371)
$[\text{DP } D \dots \alpha_F \dots] \dots \textit{only} [\dots \text{ } \dots]$	×	×
$[\text{DP } D_F \dots] \dots \textit{only/even} [\dots \text{ } \dots]$	○	○ (372)

ITC allows for a dissociation between the position(s) where the restriction is interpreted and where the quantificational force is interpreted. This dissociation could be thought of as a positive or negative result, depending on the particular phenomenon in question. In the remainder of this section, I will present one candidate for an independent application of ITC, and then discuss a problematic prediction of ITC.

The ITC process presented here can be applied to explain the the scopal behavior of floated quantifiers. Consider the paradigm in (374). When the quantifier *all* is adjoined to the subject in its surface position (374a), we are able to interpret it with surface scope or inverse scope with respect to sentential negation. When the quantifier is floated as in (374b–c), the interpreted scope of the quantifier is restricted in such a way that it must take scope in the floated position (Williams, 1982; Dowty and Brodie, 1984; Déprez, 1994, a.o.).

(374) Quantifier scope is fixed when floated:

- a. All the students will not graduate on time. √∀ > ¬, √¬ > ∀
- b. The students will *all* not graduate on time. √∀ > ¬, *¬ > ∀
- c. The students will not *all* graduate on time. *∀ > ¬, √¬ > ∀

Following Sportiche (1988), I will assume here a *stranding* approach to quantifier float. The idea is that the quantifier (here, *all*) was initially adjoined to the associated DP and can be stranded in intermediate positions that the subject moves through, between its theta position and its final Spec,TP position.¹³⁶ Under our Copy Theory view of movement, we can think of the quantifier and its restrictor DP as being pronounced in different positions of the shared DP movement chain.

Now consider how the fixed scope of a floated quantifier is interpreted. One straightforward approach might be to require that a DP completely reconstruct into the position where its quantifier was stranded, for example through standard syntactic reconstruction, whereby the higher copy (pronounced at PF) is completely deleted or ignored at LF. This possibility is schematized in (375):

(375) **Interpreting quantifier float in (374c) via standard syntactic reconstruction:**

a. Narrow syntax: [the students]_i will not [all [the students]_i] graduate on time.

b. LF: ~~[the students]_i~~ will not [all [the students]_i] graduate on time. $\neg > \forall$

However, under this syntactic reconstruction view, there is no DP “the students” in the derived Spec,TP subject position at all at LF. Assuming that Binding Condition A is evaluated at LF (Reinhart and Reuland, 1993; Fox, 2000; Lebeaux, 2009, a.o.), this approach would predict the example in (376) below to be ungrammatical, contrary to fact. The reflexive binding in example (376), then, acts as an argument that “the students” is interpreted in its higher, derived position at LF.

(376) ✓ The students_i seem to themselves_i to all be geniuses.

A partial application of the first version of ITC in (366) could be the solution to such cases.¹³⁷ Assuming that the quantifier is stranded in narrow syntax, we can use the Predicate Abstraction and Variable Insertion steps from (366) above. ITC gives us the behavior observed in quantifier float: the definite description is interpreted high at LF, while a quantifier which ranges over the individuals in the definite description takes scope low (377). A derivation for the example with reflexive binding, (376), is also presented below in (378).

¹³⁶Quantifiers cannot be stranded in all positions which we expect the subject to have moved through—for example, quantifiers cannot be stranded in the postverbal object position of a passive clause. Various modifications to Sportiche’s stranding account have therefore been proposed. For example, Bošković (2004) proposes that floatable quantifiers cannot be adjoined to a DP in its theta-position, as doing so would interrupt the theta-assignment process. Such details of the analysis of quantifier float are incidental to the discussion of quantifier scope here.

¹³⁷The second variety of ITC in (369), where the higher copy is interpreted as the property of its domain, does not apply to the derivation of quantifier float. In quantifier float we overtly see the higher copy as a plural definite description, which is what is expected in the first ITC. Furthermore we do not expect a property of extensional type $\langle e, t \rangle$ to be an antecedent for the reflexive in (376).

(377) ITC derives the low scope of the floated quantifier in (374c):

a. Narrow syntax: [The students] will not [all [the students]] graduate on time.

b. LF: [the students] λX . \neg [all [the students in X]] graduate on time. $\neg > \forall$

(378) ITC derives the high antecedent for binding in (376):

a. Narrow syntax: [The students]_i seem to themselves_i to [all [the students]] be geniuses.

b. LF: [the students] λX . it seem to themselves_x
that [all [the students in X]] are (individually) geniuses

The behavior of quantifier float—on the assumption that they are stranded quantifiers rather than adverbs—thus acts as support for a mechanism such as ITC to exist in the grammar. ITC allows the grammar to interpret (a copy of) the restrictor of a DP in a higher position than where the DP's quantificational part actually takes place, just as we saw overtly with floated quantifiers.

There is, however, one remaining problem with allowing ITC in the grammar. In general, it has been argued that the position where a DP is interpreted at LF for scope is also the position relevant for binding. Here I will focus on English A-movements, as has been our focus in this discussion of reconstruction. Fox (1999) makes this argument using the examples in (379), below, which involve the A-movement of an existential subject across the raising predicate “seems.” Fox argues that in example (379a), if the subject “a student of David’s” were interpreted in its base position, it would yield a Condition C violation, and this explains the fact that the sentence cannot be interpreted with inverse scope, with “seem” taking scope over the existential quantifier. Example (379b), where the reconstruction of the subject would not produce a Condition C violating structure, is not limited in scope possibilities. Fox (1999) argues that, therefore, the scope reconstruction of the subject into its base position must be one to one with its reconstruction for binding purposes.¹³⁸

(379) Scope reconstruction feeds Condition C (Fox, 1999, p. 179):

a. [A student of David_i's] seems to him_i ___ to be at the party. $\checkmark \exists > \text{seem}; * \text{seem} > \exists$

b. [A student of his_i] seems to David_i ___ to be at the party. $\checkmark \exists > \text{seem}; \checkmark \text{seem} > \exists$

For Fox (1999), the “seem $> \exists$ ” readings in (379) should be analyzed using syntactic reconstruction via deletion of higher copies. The scope-trapping effect in (379a) follows because the restrictor “student of David’s” must have been merged in the lower position, as the higher copy will be deleted by LF.

¹³⁸As to why the lower copy of the restriction “student of David’s” in (379a) is not enough to trigger a Condition C violation, Takahashi and Hulsey (2009) proposes that the entire restrictor “student of David’s” is wholesale late-merged in the higher, Case-receiving position.

In contrast, the technology of ITC, taken together with at least partial late merger of the restrictions of A-moved DPs (Takahashi and Hulsey, 2009), predicts that the restriction of a DP can be interpreted high while the DP takes scope low. This conflicts with the data in (379). I demonstrate here how the “seem > \exists ” reading of (379a) is predicted to be possible using ITC:

(380) **Inverse Trace Conversion may incorrectly predict inverse scope for (379a):**

a. Narrow syntax:

[A student late-merged of David_i's]_i seems to him_i [a student]_i to be at the party.

b. LF after Inverse Trace Conversion:

[The students of David_i's] λX seems to him_i [[a [student in X]] to be at the party]

⇒ predicts “seem > \exists ” to be possible in (379a) ☹

In order to rule out this configuration, I propose a ban on modifications of quantificational domains above the point of quantification:

(381) **The Ban on Domain Modification Above Quantification:**

In a chain of copied DPs, one copy of the DP will be interpreted at LF with the original quantificational determiner: call this DP “DP_Q.” All copies of this DP which c-command DP_Q must have the same restrictor as in DP_Q.

Note that, without ITC, the interpreted quantifier of a moved DP at LF was necessarily in its highest LF copy, and therefore configurations violating The Ban on Domain Modification Above Quantification would never be generated. I have shown that these previous approaches to reconstruction are unable to derive the correct pattern of focus association with reconstruction, but ITC can. Together with The Ban on Domain Modification Above Quantification, ITC can also derive the correlation of interpreted scope position and the interpreted position for Condition C purposes (Fox, 1999).

We began this chapter with a puzzle: the behavior of backwards association with F-marked determiners shows that syntactic reconstruction exists in English and can feed focus association, but reconstruction nonetheless is unable to feed association of *only* with F-marking in the restrictor. In this section I introduced Inverse Trace Conversion, a new approach to syntactic reconstruction where the higher copy's restrictor also contributes to interpretation. I showed how ITC is able to derive the full pattern of scope reconstruction and backwards association observed in English. However, I also presented a number of complications that come about when using ITC.

It's worth highlighting the strangeness of the English behavior that we have observed. Work such as Fox (1999) convincingly shows that the correct notion of reconstruction in English is *syntactic* reconstruction. Aside from this inability to feed backwards association, to my knowledge, there is

no other way in which a reconstructed constituent is not simply interpreted in its lower position. A natural question, then, is whether this behavior is a universal fact about reconstruction cross-linguistically or is somehow specific to English. I will answer this question in the next section by presenting the behavior of German and Dutch, which in fact exhibits the behavior that one would straightforwardly expect from syntactic reconstruction, further underscoring the surprising behavior of English.

7.4 Reconstruction feeds association with *only* in German and Dutch

The focus in this dissertation has been on association with the English focus-sensitive adverbs *even* and *only*. I have shown that in cases of F-marking in the restrictor of moved DPs, *even* is able to associate backwards with the F-marking but *only* is unable to, even if the DP can be reconstructed (section 7.1). The English data follows the schema below, repeated from (167).

(382) In English, *even* but not *only* can associate with material which has moved out: (=167)

$$\underbrace{[\text{DP } \dots \alpha_F \dots]} \dots [\checkmark \text{even} / * \text{only} [\dots \text{---} \dots]]$$

In this section, I note that movement bleeding focus association with *only* is not universal. In particular, certain movements in German and Dutch are documented to allow backwards association with *only*, in contrast to the behavior of similar movements in English.¹³⁹ A close look at these movements will show that backwards association with *only* forces reconstruction. This is explained by these movements utilizing a traditional syntactic reconstruction option rather than the Inverse Trace Conversion option used by English. The primary evidence here will come from the obligatory movement to initial position in V2 clauses, a type of movement that does not exist in English.

Let's begin by looking at movement to the *prefield*—the first position of V2 clauses—in German and Dutch. Barbiers (1995) reports that both *only*'s and *even*'s can associate with constituents moved to clause-initial position, as exemplified below.¹⁴⁰

(383) \checkmark [TWEE boeken]_i denk ik dat Jan *pas/maar* ____i heeft gekocht.
 [[two]_F books] think I that John just/only has bought
 'TWO books, I think that John has bought just/only ___.' (based on Barbiers, 1995)

¹³⁹I thank Patrick Grosz and Coppe van Urk for help in constructing much of the data tested here, for German and Dutch respectively.

¹⁴⁰Beaver and Clark (2008) reports that (383) with another variant of 'only,' *alleen maar*, is ungrammatical. Coppe van Urk (p.c.) reports that (383) with *alleen maar* is indeed slightly degraded, but is still grammatical, and judges (383) with *maar* as completely grammatical.

- (384) ✓ [TWINTIG boeken]_F; denk ik dat Jan *zelfs* __, heeft gekocht.
 [[twenty]_F books] think I that Jan even has bought
 'TWENTY books, I think that John has bought *even* __.'

Similar facts hold for German, as seen in example (385).

- (385) [Jedes BUCH] hat der Hans *nur* gelesen... (ZEITSCHRIFTEN hat er keine gelesen.)
 every [book]_F has the Hans only read [magazines]_F has he none read
 'It was only every BOOK that Hans read. He didn't read any MAGAZINES.'

I propose that in such cases of backwards association with *only*, the fronted DP is reconstructed into its base position, within the scope of *only*, through either standard syntactic reconstruction or PF movement.¹⁴¹ The following examples provide evidence for this claim.

First consider a basic case with movement of a quantificational DP to first position, in (386) below. The most straightforward reading for this utterance is with surface scope, with the quantifier in first position taking scope over the subject.

- (386) **Movement to first position has scope consequences:**

[Jeden Soldaten] hat eine Kugel getötet.
 every soldier has a bullet killed
 'Every soldier was killed by a bullet.'

= 'For every soldier x , there was some bullet or other y , such that y killed x ' $\forall > \exists$

Now consider a variant of (386) where *nur* 'only' associates backwards with part of the DP in first position. In this example, the more natural interpretation would be one with surface scope, where the object "every infantry person" takes scope over the subject "a bullet." However, such a reading is not available (387a). Instead, the sentence is interpreted as a claim about a specific bullet, taking scope over every infantry person (387b).

- (387) **Backwards association with *only* forces scope reconstruction:**

[Jeden InfanterISTEN] hat eine Kugel *nur* getötet... (Die ArtillerISTEN waren
 every [infantry person]_F has a bullet only killed the [artillery people]_F were
 unverletzt.)
 unharmed

- a. * *only* > $\forall > \exists$: 'Only every [infantry person]_F x was such that a bullet y killed x .'
 b. ✓ $\exists >$ *only* > \forall : 'For some (specific) bullet, it *only* killed every [infantry person]_F.'

'...The artillery people were unharmed.'

¹⁴¹ As noted above, these two options will yield the same predictions for our LF purposes.

Backwards association with *nur* 'only' has forced the DP containing the F-marking to reconstruct into *nur*'s scope, and therefore within the scope of the subject "a bullet." As the F-marking is on the restrictor of the moved DP, this behavior is predicted by a standard syntactic reconstruction/PF movement approach to scope reconstruction, but not by the Inverse Trace Conversion approach proposed for English in the previous section. At LF, this material is only in the base position, within the scope of *only*.

The results suggest that English can only use Inverse Trace Conversion for reconstruction, while German and Dutch have the standard syntactic reconstruction option. This raises a number of questions regarding the deeper difference between these two types of languages. However, the comparison is not completely parallel. The form of movement I have studied here—movement to clause-initial position in V2 clauses—is a form of movement which does not exist in English.¹⁴² I will therefore tentatively suggest that the availability of different reconstruction strategies correlates with the nature of the original movement chain, and movement to first position in V2 clauses is a type of movement that allows syntactic reconstruction.

A related movement which exists in German and Dutch but not in English is the head-movement of a verb to verb-second position. As the following German examples show, both *nur* 'only' and *sogar* 'even' are able to associate backwards with a main verb moved to V2 position:

- (388) Martina TANZT vermutlich *nur* mit Hans-Dietrich.
 Martina [DANCES]_i presumably only with Hans-Dietrich ____i
 'Presumably Martina only DANCES with Hans-Dietrich
 (...but does not go to bed with him).' (Bayer, 1996, chapter 1, fn. 13)
- (389) Martina SCHLÄFT vermutlich *sogar* mit Hans-Dietrich.
 Martina [SLEEPS]_i presumably even with Hans-Dietrich ____i
 'Presumably Martina even SLEEPS with Hans-Dietrich.'

Patrick Grosz (p.c.) also observes the following naturally-occurring example of *nur* associating backwards with a finite verb in second position:

- (390) Example from *Mannheimer Morgen* newspaper, February 22, 2013 (Patrick Grosz, p.c.):
 Context: I don't deal with this kind of stuff...
 das [verwirrt]_i mich *nur* ____i.
 that [confuses]_F me only
 'That only [confuses]_F me.'

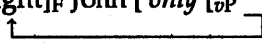
¹⁴²Taglicht (1984, ch. 4, fn. 35) notes that this inability to backwards associate with *only* is apparently a newer innovation of English: "It is interesting to note that this constraint did not exist in earlier English; see the quotations in OED, s.v. *only* adv., i.c, for example: 1721 *The eldest son shall only inherit his father.*"

At first glance it seems that English again differs from German in this regard. English also has verbal head movement, for example in question formation. Example (391) is a baseline showing that *only* is able to associate with the modal auxiliary “might.” Example (392) below shows that backwards association of *only* with the modal is not possible.

(391) Baseline: John *only* [might]_F read the book.

(392) * What [might]_F John *only* read?

Assumed derivation: What [might]_F John [*only* [_{vP} ___ [read ___]]]?



However, there is reason to believe that (392) is not a fair test of backwards association with *only*—that is, that the assumed narrow syntax derivation in (392) is not correct. First, note that this position for *only* before the modal auxiliary is in general a marked position for *only* to be if it associates with material past the auxiliary, as we would expect it to be able to do if it is indeed an adverb (393).

(393) John (?**only*) might (✓*only*) read the [book]_F... (not the [magazine]_F.)

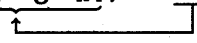
Instead, the pre-modal *only* in (391) may be best analyzed as a constituent *only* modifying the auxiliary alone, as schematized in (394) below. This view is further supported by the fact that *only* can front together with the modal to yield the intended reading of (392), as in (395).¹⁴³

(394) *Only* adjoined directly to the modal in (391):

John [[*only* [might]_F] [read the book]].

(395) ✓ What *only* [might]_F John read? (cf 392)

Possible derivation: What [*only* [might]_F] John ___ [read ___]?



The ungrammaticality of (392) may then reflect an inability to subextract from an *only*-phrase, rather than any problem with the adverb *only* associating backwards with a head-moved verb. Verbs which clearly originate within the scope of the adverb *only* in English are main verbs, which do not head-move in question formation, unlike in German. We are therefore unable to directly compare the German behavior with parallel data in English.

¹⁴³In turn, however, example (395) does raise questions regarding the nature of head movement itself, on the assumption that “*only* might” here is not fused into a single morphophonological word or syntactic terminal.

7.5 Summary

In chapter 4, I briefly considered the use of syntactic reconstruction as a general solution to the problem of backwards association. There, I observed that if syntactic reconstruction can generally feed focus association, we would overgenerate backwards association with *only*. I therefore set aside the possibility of syntactic reconstruction in much of the dissertation.

In this final chapter, I returned to the issue of syntactic reconstruction. I presented certain cases of backwards association with an F-marked determiner which are only possible when reconstruction is independently licensed by Scope Economy. These examples showed that covert scope shifting operations cannot occur solely for the purpose of feeding focus association. These examples also show that syntactic reconstruction *does* exist in English and furthermore that—when it is licensed—it can feed focus association.

This leaves us with a puzzle: syntactic reconstruction exists and feeds focus association with the determiner, but it does not feed focus association with material in the restrictor. As a solution to this problem, I proposed a new view of syntactic reconstruction: *Inverse Trace Conversion*. In Inverse Trace Conversion, both copies of movement contribute to semantic interpretation, but in a novel way. The quantifier in the higher copy is replaced, and the quantifier in the lower copy is interpreted. This allows for backwards association with the lower determiner, but backwards association of *only* with material in the restrictor will be ungrammatical in the same way that backwards association with *only* is ungrammatical without scope reconstruction.

In contrast to English, German and Dutch allow backwards association of *only* with material in the restrictor of DPs which have moved to clause-initial position. This is the behavior that is predicted under a standard syntactic reconstruction approach, as evidenced by the obligatory narrow scope of the focused constituent in these configurations. This obligatory movement to clause-initial position does not exist in English, making it difficult to directly compare the English and Germanic facts.

These differences highlight the complexity of the problem of backwards association. The availability of backwards association depends on the semantics of the focus-sensitive operator, as well as the nature of the relevant movement chain. I have shown, however, that the distribution of backwards association correlates with independent syntactic processes such as the application of reconstruction and the timing of late-merger. These interactions can be modeled under my Copy Theoretic approach to backwards association.

References

- Anderson, Stephen R. 1972. How to get 'even'. *Language* 48:893–906.
- Aoun, Joseph, and Yen-hui Audrey Li. 1993. Wh-elements in situ: Syntax or LF? *Linguistic Inquiry* 24:199–238.
- Barbiers, Sjef. 1995. The syntax of interpretation. Doctoral Dissertation, Leiden University.
- Bartels, Christine. 1997. Towards a compositional interpretation of English statement and question intonation. Doctoral Dissertation, University of Massachusetts Amherst.
- Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4.
- Bayer, Josef. 1996. *Directionality and logical form: on the scope of focusing particles and wh-in-situ*. Kluwer Academic Publishers.
- Beaver, David Ian, and Brady Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Wiley-Blackwell.
- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.
- van Benthem, Johan. 1983. Determiners and logic. *Linguistics and Philosophy* 6:447–478.
- Bhatt, Rajesh, and Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35:1–45.
- Bhatt, Rajesh, and Roumyana Pancheva. 2007. Degree quantifiers, position of merger effects with their restrictors, and conservativity. In *Direct compositionality*, ed. Chris Barker and Pauline Jacobson, 306–335. Oxford.
- Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language & Linguistic Theory* 22:681–742.
- Bresnan, Joan. 1994. Linear order vs. syntactic rank: Evidence from weak crossover. In *Proceedings of CLS 30*, 57–89.
- Brody, Michael. 1993. θ -theory and arguments. *Linguistic Inquiry* 24:1–23.
- Bruening, Benjamin. 2014. Precede-and-command revisited. *Language* 90:342–388.
- Carlson, Gregory N., and Francis Jeffrey Pelletier. 1995. *The generic book*. University of Chicago Press.

- Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Foris.
- Chomsky, Noam. 1993. A minimalist program for linguistic inquiry. In *The view from Building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 1–52. MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. MIT Press.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36:1–22.
- Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel.
- Cresti, Diana. 1995. Extraction and reconstruction. *Natural Language Semantics* 3:79–122.
- Dayal, Veneeta. 1996. *Locality in wh quantification*. Kluwer Academic Publishers.
- Dayal, Veneeta. 2002. Single-pair versus multiple-pair answers: *wh*-in-situ and scope. *Linguistic Inquiry* 33:512–520.
- Déprez, Viviane. 1994. Questions with floating quantifiers. In *Proceedings of SALT 4*.
- Dowty, David, and Belinda Brodie. 1984. The semantics of “floated” quantifiers in a transformation-less grammar. In *Proceedings of WCCFL 3*, ed. Mark Cobler, Susannah MacKay, and Michael T. Westcoat.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68:81–138.
- Elbourne, Paul. 2002. Situations and individuals. Doctoral Dissertation, Massachusetts Institute of Technology.
- Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy* 6:5–34.
- Erlewine, Michael Yoshitaka. 2014. Explaining leftward focus association with *even* but not *only*. In *Proceedings of Sinn und Bedeutung 18*, ed. Urtzi Etxeberria, Anamaria Fălăuș, Aritz Irurtzun, and Bryan Leferman, 128–145. Bayonne and Vitoria-Gasteiz.
- Fauconnier, Gilles. 1976. *Etude de certains aspects logiques et grammaticaux de la quantification et de l'anaphore en français et en anglais*. Champion, Paris.
- von Stechow, Kai. 1993. Exeptive constructions. *Natural Language Semantics* 1:123–148.
- von Stechow, Kai, and Irene Heim. 2005. Intensional semantics lecture notes. MIT Spring 2005 Edition.
- Fischer, Susan. 1968. On cleft sentences and contrastive stress. Manuscript, MIT.
- Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3:283–341.
- Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.
- Fox, Danny. 2000. *Economy and semantic interpretation: a study of scope and variable binding*. MIT Press.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

- Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: A case for overt QR. In *Proceedings of WCCFL 18*.
- Gajewski, Jon. 2002. On analyticity in natural language. Manuscript, Massachusetts Institute of Technology.
- Geurts, Bart. 1997. Good news about the description theory of names. *Journal of Semantics* 14:319–348.
- Giannakidou, Anastasia. 2007. The landscape of EVEN. *Natural Language & Linguistic Theory* 25:39–81.
- Grice, Herbert Paul. 1975. Logic and conversation. *Syntax and Semantics* 3:41–58.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. Studies on the semantics of questions and the pragmatics of answers. Doctoral Dissertation, University of Amsterdam.
- Guerzoni, Elena. 2004. *Even-NPIs in yes/no questions*. *Natural Language Semantics* 12:319–343.
- Hale, Kenneth. 1978. The structure of English sentences. Handout, Massachusetts Institute of Technology.
- Hamblin, Charles. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.
- Heim, Irene. 1983. On the projection problem for presuppositions. In *Proceedings of WCCFL 2*, ed. M. Barlow, D. Flickinger, and N. Wiegand, 114–125.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell.
- Herburger, Elena. 2000. *What counts: focus and quantification*. Number 36 in *Linguistic Inquiry Monographs*. MIT Press.
- Herdan, Simona. 2005. Who knows the only reading of *only*? *Snippets* 10.
- Higginbotham, James, and Robert Carlen May. 1981. Questions, quantifiers and crossing. *The Linguistic Review* 1:41–80.
- Horn, Laurence Robert. 1969. A presuppositional analysis of *only* and *even*. In *Papers from the Fifth Regional Meeting*, ed. Robert I. Binnick, Alice Davison, Georgia M. Green, and J.L. Morgan, 98–107. Chicago Linguistic Society.
- Hornstein, Norbert. 1995. *Logical Form: from GB to Minimalism*. Oxford: Blackwell.
- Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30:69–96.
- Hornstein, Norbert. 2000. *Move!: a minimalist theory of construal*. Oxford: Blackwell.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. MIT Press.
- Jackendoff, Ray. 1977. *X' syntax: A study of phrase structure*. MIT Press.
- Jacobs, Joachim. 1984. Funktionale satzperspektive und illokutionsemantik. *Linguistische Berichte* 91:25–58.

- Johnson, Kyle, and Satoshi Tomioka. 1998. Lowering and mid-size clauses. In Katz et al. (1998), 185–206.
- Kadmon, Nirit. 1990. Uniqueness. *Linguistics and Philosophy* 13:273–324.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.
- Karttunen, Lauri, and Stanley Peters. 1979. Conventional implicature. In *Syntax and semantics, volume 11: Presupposition*, ed. Choon-Kyu Oh and David A. Dinneen, 1–56. Academic Press.
- Katz, Graham, Shin-Sook Kim, and Heike Winhart, ed. 1998. *Reconstruction: Proceedings of the 1997 Tübingen workshop*. Number 127 in *Arbeitspapiere des Sonderforschungsbereichs 340*. Universities of Tübingen and Stuttgart.
- Katzir, Roni. 2013. A note on contrast. *Natural Language Semantics* 21:333–343.
- Katzir, Roni. 2014. On the roles of markedness and contradiction in the use of alternatives. In *Semantics, pragmatics, and the case of scalar implicatures*, ed. Salvatore Pistoia-Reda. Palgrave.
- Kay, Paul. 1990. Even. *Linguistics and Philosophy* 13:59–111.
- Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1:128–191.
- Keenan, Edward L., and Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9:253–326.
- Kennedy, Christopher. 1994. Argument contained ellipsis. Technical report, Linguistics Research Center Report LRC-94-03, University of California, Santa Cruz.
- Kennedy, Christopher. 1997. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 4:662–688.
- Keshet, Ezra. 2008. Good intensions: paving two roads to a theory of the *de re/de dicto* distinction. Doctoral Dissertation, Massachusetts Institute of Technology.
- Kiparsky, Paul, and Carol Kiparsky. 1970. Fact. In *Progress in linguistics*, ed. M. Bierwisch and K. Heidolph. Mouton.
- Kitagawa, Yoshihisa. 1986. Subjects in Japanese and English. Doctoral Dissertation, University of Massachusetts Amherst.
- König, Ekkehard. 1991. *The meaning of focus particles*. Routledge.
- Kotek, Hadas. 2014. Composing questions. Doctoral Dissertation, Massachusetts Institute of Technology.
- Kratzer, Angelika. 1991. The representation of focus. In von Stechow (1991b), 825–834.
- Krifka, Manfred. 1995. Focus and the interpretation of generic sentences. In Carlson and Pelletier (1995), 238–264.
- Krifka, Manfred. 1998. Additive particles under stress. In *SALT VIII*, ed. Devon Strolovitch and Aaron Lawson, 111–129. Cornell University.

- Krifka, Manfred. 2006. Basic notions of information structure. In *Interdisciplinary studies on information structure 6*, ed. Caroline Féry, Gisbert Fanselow, and Manfred Krifka, 1–50.
- Kripke, Saul. 1990. Presupposition and anaphora: Remarks on the formulation of the projection problem. Manuscript, Princeton University.
- Kuroda, Sige-Yuki. 1988. Whether we agree or not: a comparative syntax of English and Japanese. *Linguisticæ Investigations* 12:1–47.
- Lahiri, Utpal. 2008. The semantics and pragmatics of some scalar expressions in Spanish. *Anuario del Seminario de Filología Vasca Julio de Urquijo* 42:359–389.
- Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral Dissertation, University of Massachusetts Amherst.
- Lebeaux, David. 1998. Where does binding theory apply? Technical report, Princeton, NJ: NEC Research Institute.
- Lebeaux, David. 2009. *Where does binding theory apply?* Number 50 in Linguistic Inquiry Monographs. MIT Press.
- Lycan, William G. 1991. *Even and even if*. *Linguistics and Philosophy* 14:114–150.
- Matushansky, Ora. 2006. Why Rose is the Rose: on the use of definite articles in proper names. In *Empirical issues in syntax and semantics 6*, ed. Olivier Bonami and Patricia Cabredo Hofherr, 285–307. Colloque de syntaxe et sémantique à Paris.
- May, Robert Carlen. 1977. The grammar of quantification. Doctoral Dissertation, Massachusetts Institute of Technology.
- May, Robert Carlen. 1985. *Logical form: Its structure and derivation*. Number 12 in Linguistic Inquiry Monographs. MIT Press.
- McCawley, James D. 1970. English as a VSO language. *Language* 46:286–299.
- McCawley, James D. 1995. The focus and scope of *only*. In *Discourse and meaning*, 171–193. John Benjamins Publishing Company.
- McCloskey, James. 1997. Subjecthood and subject positions. In *Elements of grammar*, ed. Liliane Haegeman, 197–235. Kluwer Academic Publishers.
- McNally, Louise. 2008. DP-internal *only*, amount relatives, and relatives out of existentials. *Linguistic Inquiry* 39:161–169.
- Merchant, Jason. 1999. The syntax of silence: sluicing, islands, and identity in ellipsis. Doctoral Dissertation, University of California Santa Cruz.
- Merchant, Jason. 2001. *The syntax of silence*. Oxford University Press.
- Milsark, Gary. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3:1–30.
- Montague, Richard. 1970. Universal grammar. *Theoria* 36:373–98.

- Nakanishi, Kimiko. 2012. The scope of *even* and quantifier raising. *Natural Language Semantics* 20:115–136.
- Novel, Marc, and Maribel Romero. 2009. Movement, variables, and Hamblin alternatives. In *Proceedings of Sinn und Bedeutung 14*, 322–338.
- Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32:303–344.
- Partee, Barbara Hall. 1975. Deletion and variable binding. In *Formal semantics of natural language*, ed. Edward L. Keenan. Cambridge University Press.
- Partee, Barbara Hall. 2004. *Compositionality in formal semantics: Selected papers by Barbara H. Partee*. Blackwell.
- Pesetsky, David Michael. 2000. *Phrasal movement and its kin*. Number 37 in *Linguistic Inquiry Monographs*. MIT Press.
- Postal, Paul M., and Geoffrey K. Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19:635–670.
- Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford University Press.
- Reinhart, Tanya. 2006. *Interface strategies: optimal and costly computations*. Number 45 in *Linguistic Inquiry Monographs*. MIT Press.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24.
- Romoli, Jacopo. 2009. Towards a structural account of conservativity. In *Proceedings of Sinn und Bedeutung 14*, ed. Martin Prinzhorn, Viola Schmitt, and Sarah Zobel, 373–387.
- Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts, Amherst.
- Rooth, Mats. 1992a. Ellipsis redundancy and reduction redundancy. In *Proceedings of the Stuttgart ellipsis workshop, report series, Sprachtheoretische Grundlagen für die Computerlinguistik, no 29, SFB 340*. IBM Heidelberg.
- Rooth, Mats. 1992b. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Rullmann, Hotze. 1995. Maximality in the semantics of *wh*-constructions. Doctoral Dissertation, University of Massachusetts Amherst.
- Rullmann, Hotze. 1997. *Even*, polarity, and scope. In *Papers in experimental and theoretical linguistics*, ed. Martha Gibson, Grace Wiebe, and Gary Libben, volume 4, 40–64. Department of Linguistics, University of Alberta, Edmonton, Canada.
- Rullmann, Hotze. 2003. Additive particles and polarity. *Journal of Semantics* 20:329–401.
- Rullmann, Hotze, and Sigrid Beck. 1998. Reconstruction and the interpretation of *which*-phrases. In Katz et al. (1998), 223–256.
- Ruys, Eddy G. 2000. Weak crossover as a scope phenomenon. *Linguistic Inquiry* 31.

- Sag, Ivan Andrew. 1976. Deletion and logical form. Doctoral Dissertation, Massachusetts Institute of Technology.
- van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.
- Sauerland, Uli. 1998. The meaning of chains. Doctoral Dissertation, Massachusetts Institute of Technology.
- Sauerland, Uli. 2003. Intermediate adjunction with A-movement. *Linguistic Inquiry* 34:308–314.
- Sauerland, Uli. 2004. The interpretation of traces. *Natural Language Semantics* 12:63–127.
- Sauerland, Uli. 2007. Copying vs. structure sharing: a semantic argument. *Linguistic Variation Yearbook* 7:27–51.
- Sauerland, Uli. 2013. Presuppositions and the alternative tier. In *Proceedings of SALT 23*, 156–173.
- Sauerland, Uli, and Paul Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33:283–319.
- Schlenker, Philippe. 2005. Minimize Restrictors! (notes on definite descriptions, Condition C and epithets). In *Proceedings of Sinn und Bedeutung 9*, 385–416.
- Schwarz, Bernhard. 2005. Scalar additive particles in negative contexts. *Natural Language Semantics* 13:125–168.
- Selkirk, Elisabeth O. 1984. *Phonology and syntax*. MIT Press.
- Sobin, Nicholas. 2010. Echo questions in the minimalist program. *Linguistic Inquiry* 41:131–148.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425–449.
- Sportiche, Dominique, Hilda Koopman, and Edward P. Stabler. 2014. *An introduction to syntactic analysis and theory*. Wiley-Blackwell.
- von Stechow, Arnim. 1991a. Current issues in the theory of focus. In von Stechow (1991b), 804–824.
- von Stechow, Arnim. 1991b. *Semantik: Ein internationales handbuch der zeitgenössischen forschung*. Walter de Gruyter.
- Sudo, Yasutada. 2012. On the semantics of phi features on pronouns. Doctoral Dissertation, Massachusetts Institute of Technology.
- Sudo, Yasutada. 2014. Presupposition projection in quantified sentences and cross-dimensional anaphora. Manuscript, University College London.
- Taglicht, Josef. 1984. *Message and emphasis: on focus and scope in English*. Longman.
- Takahashi, Shoichi, and Danny Fox. 2005. MaxElide and the re-binding problem. In *Proceedings of Semantics and Linguistic Theory 15 (SALT 15)*, ed. Effi Georgala and Jonathan Howell, 223–240. Ithaca, NY: CLC Publications.

- Takahashi, Shoichi, and Sarah Hulseley. 2009. Wholesale late merger: beyond the A/ \bar{A} distinction. *Linguistic Inquiry* 40:387–426.
- Tancredi, Chris. 1990. Not only EVEN, but even ONLY. Manuscript, Massachusetts Institute of Technology.
- Wagner, Michael. 2005. Prosody and recursion. Doctoral Dissertation, Massachusetts Institute of Technology.
- Wagner, Michael. 2006. Association by movement: evidence from NPI-licensing. *Natural Language Semantics* 14:297–324.
- Wagner, Michael. 2013. Additivity and the syntax of *even*. Manuscript, McGill, September 2013.
- Wilkinson, Karina. 1996. The scope of *even*. *Natural Language Semantics* 4:193–215.
- Williams, Edwin S. 1977. Discourse and logical form. *Linguistic Inquiry* 8:101–139.
- Williams, Edwin S. 1982. The NP cycle. *Linguistic Inquiry* 13:277–295.
- Wold, Dag E. 1996. Long-distance selective binding: the case of focus. In *Proceedings of SALT 6*, ed. Teresa Galloway and Justin Spence, 311–328.