The Meaning of Chains

by

Uli Sauerland

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

August 1998

Author .....
Department of Linguistics and Philosophy
August 31, 1998

Certified by .....
Noam Chomsky
Institute Professor of Linguistics
Thesis Supervisor

Certified by .....
Irene Heim
Professor of Linguistics
Thesis Supervisor

Certified by .....
David Pesetsky
Professor of Linguistics
Thesis Supervisor

Accepted by .....
Michael Kenstowicz
Chair of the Linguistics Program

The Meaning of Chains

by

Uli Sauerland

Submitted to the Department of Linguistics and Philosophy
on August 31, 1998, in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

Abstract

This thesis investigates the mechanisms applying in the interpretation of syntactic chains. The theoretical background includes a translation of syntactic forms into semantic forms and a model theoretic explication of the meaning of semantic forms. Simplicity considerations apply to all three stages of the interpretation process: syntactic derivation, translation into semantic forms, interpretation of semantic forms.

Three main results are achieved. The first is that trace positions can have semantic content beyond what is needed for the semantic dependency of trace and binder. This extra content is some or all of the lexical material of the head of the chain, as expected on the copy theory of movement. Two independent arguments support this conclusion. One, discussed in chapter 2, is based on the distribution of Condition C effects, where novel interactions between variable binding, antecedent contained deletion and Condition C are observed. The second, developed in chapter 3, is based on conditions on the identity of traces observed in antecedent contained deletion constructions. Both arguments lead to the same generalizations about what lexical material of the head is interpreted in the trace position.

The second main result is that lambda calculus is superior to both standard predicate logic and combinatorial logic as the mathematical model for the semantic mechanism mediating the dependency of trace (or bound pronoun) and binder. Chapter 4 argues this on the basis of the distribution of focus and destressing in constructions with bound pronouns.

The third main result is that quantification must be allowed to range over pointwise different choice functions. Chapter 5 shows that quantification over individuals is insufficient, and that pointwise different choice functions are required. The result entails that the syntactic difference of A-chains and A-bar chains predicts a semantic difference in the type of the variable involved, which is argued to explain weak crossover phenomena.

Chapters 6 argues that the interpretation procedures developed in the preceeding chapters account for all cases. It is shown that only traces of the type of individuals arise, and that scope reconstruction is a phonological phenomenon. The latter result also supports the T-model of syntax.

Acknowledgments

I thank all the people who helped me on the way towards finishing this thesis. All three members of my committee had a great impact on this thesis and on my thinking more generally. David Pesetsky helped me clear up my thinking in countless meetings and put it into writing. Irene Heim was always quick at finding interesting new challenges for me and had usually more faith that I could meet them than I did. Noam Chomsky expanded my vision of the bigger picture and helped me to sharpen many technical details at the same time.

The intellectual environment this thesis grew in was provided by two, overlapping groups of people, the LF reading group at MIT and the inner hallway of building E39. Danny Fox, especially, was enormously helpful. Also, Jon Nissenbaum, Paul Hagstrom, Martin Hackl, Kai von Fintel, and Kazuko Yatsushiro caused many improvements. I also thank my housemates Susi Wurmbrand, Jonathan Bobaljik, and Meltem Kelepir for sharing many things with me.

When I came to MIT, I was a mathematician. Becoming a linguist was harder than I had expected, and I thank the following people for their help. My classmates, Danny Fox, Paul Hagstrom, Norvin Richards, Martha McGinnis, Rob Pensalfini, Judy Yoon-Kyung Baek, Ingvar Løfsted, David Braun for being my friends. My other teachers at MIT: Alec Marantz, Roger Schwarzschild, Ted Gibson, Morris Halle, Ken Hale, Michael Kenstowicz, Ken Wexler, and Robert Berwick. Hubert Truckenbrodt and Colin Philipps for getting me through first year syntax. Those who contributed to the liveliness of the LF-reading group over the past few years, espe-

cially Renate Musan, Diana Cresti, Lisa Matthewson, Michel DeGraff, Rajesh Bhatt, Roumyana Izvorski, Sabine Iatridou, and Julie Legate. My former housemates Heidi Harley, Hubert Truckenbrodt, Ayumi Ueyama, Danny Fox, Martha McGinnis, and Marie Hélène Côté. And, the many other linguists I was in contact with; in the MIT-community, Pilar Barbosa, Andrew Carnie, Ben Bruening, Vivian Lin, Cheryl Zoll, Marie Claude Boivin, Wayne O'Neill, Maya Honda, Gina Rendon, Maire Noonan, David Embick, Masa Koizumi, Yoonjung Kang, Bridget Copley, Karlos Arregui-Urbina, Jeannette Schaeffer, Calixto Aguero-Bautista, Isabel Oltra-Massuet, Dylan Tsai, Luciana Storto, Jay Rifkin, Cornelia Krause, Dag Wold, Carson Schuetze, Hooi-Ling Soh, Takako Aikawa, Idan Landau, Eleni Anagnostopoulou, Taylor Roberts, Philippe Schlenker, Sonny Vu, Shigeru Miyagawa, San Tunstall, Marlyse Baptista, and Vaijayanthi Sarma; and elsewhere, especially Elizabeth Laurençot, Kyle Johnson, Pauline Jacobson, Winfried Lechner, Sandiway Fong, Daniel Büring, Jason Merchant, Gereon Müller, Sigrid Beck, Artemis Alexiadou, Chris Kennedy, Hotze Rullmann, Dominique Sportiche, Tim Stowell, Mamoru Saito, William Snyder, Angelika Kratzer, Joachim Sabel, Hamida Demirdache, Akira Watanabe, Gennaro Chierchia, Viviane Deprez, Gertjan Postma, John Frampton, Diane Jonas, Matthias Schlesewsky, Laurel Laporte-Grimes, Takeo Kurafuji, Friederike Moltmann, Chris Wilder, Satoshi Oku, Alan Munn, Christina Schmitt, Günther Grewendorf, Howard Lasnik, Yael Sharvit, Satoshi Tomioka, Masao Ochi, Ayumi Matsuoka, Elena Herburger, Piroska Csuri, Gisbert Fanselow, Josef Bayer, Markus Bader, and Michael Meng. I'm also grateful to those who inspired me to pursue a career in linguistics: Peter Eyer, Arnim von Stechow, Urs Egli, Manfred Kupffer, Ulf Friedrichsdorf, Shin-Sook Kim, Uta Schwertel,

and Wolfgang Sternefeld.

Most of section 6.2 was presented at the MIT LingLunch series, at the LF-reading group at MIT, at the postersession of NELS 28 at the University of Toronto, and at WCCFL 18 at the University of British Columbia at Vancouver. Some of chapters 2, 3, and 5 was presented at the LF-reading group at MIT, at SALT 8 at MIT and at the Colloquium Series at the University of California at Los Angeles. I thank the audiences of these presentations for their comments.

During my five years at MIT, I received financial support from the German Academic Exchange Service (DAAD), from the NSF-sponsored RTG grant to David Pesetsky and from the Linguistics Department at MIT.

For being with me through all of this, I thank my family: Vater, Mutter, Stefan, Angela, Kazuko, and Kai, who arrived when it was almost over.

[To match the page numbering of the filed version of the thesis the table of contents starts on page 11.]

Meinen Eltern,

Thankmar und Maria Luise Sauerland

Contents

Chapter 1

Introduction

Chains are dependencies between two or more positions of a syntactic structure which are created by the syntactic computation, and the making of chains is an extensively researched topic in syntax. Why is the meaning of chains interesting? For one, the meaning of chains just like that of bound pronouns is, at least in an intuitive sense, not compositional: the meanings of two disjoint pieces of structure, the head and the tail of a chain, seem intimately related. For this reason, the study of the semantics of chains promises to provide insight into the interpretation processes that apply to syntactic structures. Little research has been done on the semantics of chains, and the semantic dependency in a chain is standardly treated exactly parallel to pronoun binding. This assumption, in particular, this thesis sets out to refute.

Some fundamental questions to ask about the semantics of chains are the following: What aspects of the syntactic representation of a chain are relevant for the semantics? What is the independent contribution of the head of a chain to the meaning of the whole? What is the independent contribution of the tail of the chain?

What are the semantic processes applying to link the two (or more) positions of a chain together? Are there differences between the semantic processes interpreting bound variables and those interpreting chains? Are differences between syntactic types of chains (A/A-bar) reflected in their semantics? This thesis tries to answer the fundamental questions about the interpretation of chains using new diagnostics, in particular focus semantics. While it in some cases provides new arguments for classic assumptions like the use of λ-calculus, it argues in many cases for a substantial revision of the semantics of chains.

The remainder of the introduction briefly clarifies the set of background assumptions in which this thesis is embedded in section 1.1, and then provides an overview of the thesis in section 1.2.

1.1 Background Assumptions

This thesis is intended as a contribution to a growing body of work that attempts to account for the contribution of syntactic structure to sentence meaning by taking both syntactic insights concerning structure forming and modifying processes seriously, and at the same time, providing fully explicit model-theoretic statements of the semantic rules involved. This research project with only minor difference is introduced in two recent semantics textbooks by Larson and Segal (1995) and by Heim and Kratzer (1998). The account of the meaning of structures within this project can involve both (covert) syntactic operations that apply for purely semantic reasons and semantic interpretation rules in whichever combination that leads to the simplest

overall account.

One insight of this research project, which I adopt, is that a clearer statement of both the syntactic computation and the semantic mechanisms is achieved if a third mechanism is hypothesized that translates the output of the syntactic computation at the syntactic level of logical form into another representation to which the semantic mechanisms apply. For example, von Stechow (1993:section 8) discusses this assumption using the term *transparent logical form* for this intermediate level of representation. I will sometimes use the term *semantic form*, but when it's unambiguous, I will often refer to this level of representation as *logical form*, as well. Obviously, simplicity considerations also apply to the translation of logical forms into semantic forms. I assume, in particular, that this translation procedure can delete parts of phrases, but also insert new pieces of structure that are necessary to represent dependencies semantically as discussed in chapter 4.

Within the three step interpretation procedure sketched in the previous paragraph, the last step, the semantic mechanism, is the least tangible. Along with most research on the topic, I content myself with stating semantic rules to define a model-theoretic concept of truth for the semantic form structures. I assume that the notion of truth in a model defined in this manner is related to semantic intuitions, in the intuitive way that is commonly assumed in the field and has proved fruitful. The semantic rules are split into lexical rules and a composition rule $\mathcal{C}$ interpreting complex phrases. The composition procedure $\mathcal{C}$ is defined by recursion over the syntactic structure. As for the basic operation that combines the interpretations of two phrases into the meaning of one branching node, I assume with Heim and Kratzer (1998) that

there are two clauses to $\mathcal{C}$, namely functional application and predicate intersection, and that whichever of the two clauses that is compatible with the semantic types of the two sub-phrases is the one that applies.

I start out with one assumption specific to the first step of the account of chains, their syntactic derivation. Namely, I assume that the syntactic process that creates chains is copying of the lexical material from the tail position of a chain to the head position as endorsed for example by Chomsky (1995). This assumption is strongly supported by reconstruction facts like those discussed in chapters 2 and 3 and in the references cited there. The other two stages of the interpretation of chains are the main subject of the investigation.

The main tool used in this thesis to study the meanings of dependencies is the semantics of focus and destressing/deletion. It is known that the explanation of these phenomena involves sentence internal entailment relationships. Hence, they can be used to test for the meaning of parts of a sentence, in particular a part that contains a dependent element, but not its antecedent. Since the semantics of focus and destressing is not so widely known, I introduce aspects of it as they become relevant: the relationship between destressing and ellipsis in section 3.2, the concept of a presuppositional skeleton and its relevance in section 3.3.2, focus indices and domains in 3.3.3, the contrastiveness requirement in 4.1.1, and the relationship between pitch accent and focus in 4.1.2. For readers who prefer a concise introduction, Kratzer (1991) and Rooth (1996) provide a good overview of the basics of focus semantics. The work of Rooth (1992b) is particularly relevant below for the account of sloppy interpretations and the relationship between focus and ellipsis. Schwarzschild (1998)

presents some ideas that are used and modified in sections 4.1.1 and 4.1.2.

1.2 Overview

The thesis is structured in three parts. The first part, consisting of chapters 2 and 3, concerns the content of a trace position in a chain. The second part, chapters 4 and 5, studies the semantic mechanisms linking a trace and its antecedent. The third part, chapter 6, provides evidence for the completeness of the solution given for the semantics of chains in the other two parts.

The main point of chapters 2 and 3 is that a trace position in a chain can contain lexical material of the head of the chain. Chapters 2 and 3 develop two independent arguments in favor of this conclusion; chapter 2 looks at the distribution of Condition C effects, chapter 3 looks at the identity requirement of two traces in ellipsis constructions. Not only do these two chapters both argue for the presence of lexical material in some cases; it'll also be shown that both argue for the same generalizations about when and what parts of the lexical material of the head of the chain is represented in the trace position.

The discussion in chapter 2 is guided by the proposal of Chomsky (1993) that whenever an R-expression in the head of a chain triggers a Condition C violation with respect to a pronoun that c-commands only the tail position of the chain, this means that the R-expression is lexically represented in the tail position. For example, the R-expression *John* in (1a) is part of a *wh*-movement chain and triggers a Condition C violation in the trace position of this chain. Therefore, I assume a semantic form

representation like (1b), where a part of the moved phrase including the R-expression is lexically represented in the gap position. The correspondence between chapters 2 and 3 argues that lexical representation in the trace position of a chain is the right approach to the distribution of Condition C phenomena.

(1) a. *Which argument of John$_i$'s father did he$_i$ defend.

b. [Which] did he defend [argument of John's father]

Assuming representations like (1b), the distribution of Condition C effects shows that DPs seem to split into independent parts that can be represented in different positions of the chain at the level of semantic form, while the parts themselves cannot be divided. I use the term *segment* for a part of a DP that seems to be always represented in the same position. Segments are the *NP-part*, which I define as the lowest NP-projection of the complement of D (excluding all adjoined modifiers), and each modifier adjoined to the NP-part. The terminology is exemplified in (2), which is a DP with two segments, the NP-part and one modifier.

(2) $\underbrace{\text{which}}_{\text{Det.}}\ \underbrace{\underbrace{\text{argument of John's}}_{\text{NP-part}}\ \underbrace{\text{that Mary had criticized}}_{\text{modifier}}}_{\text{segments}}$

The following factors are shown to affect the presence of segments in the trace position of a DP-chain in chapter 2: the surface position, the A/A-bar status of the chain, the impossibility of self-contained reference in Antecedent Contained Deletion (ACD), and the requirement that bound variables must be in the scope of their binder. In

particular the following two results are new: For one, while an ACD-relative must be represented in a higher position of a chain (Fox 1995b), the NP-part must always be represented in the trace position, as shown in section 2.1. Secondly, extending arguments of Lebeaux (1992, 1995), section 2.2 presents evidence that a segment containing bound variable must be represented in the scope of its binder, while other segments of the same DP can be represented in a higher position. In section 2.3, the first result together with other observations from the literature supports the claim that the distinction of A-bar from A-chains can be reduced to the claim that the NP-part must be represented in the tail position of an A-bar chain. Section 2.4 shows that lexical material of the relative clause head is also present in the relative clause internal trace position, but, depending on the semantic properties of the relative clause, the representation is less direct than in chains. The result of chapter 2 is summarized in section 2.5 as a set of ranked constraints.

The argument from the identity of traces in chapter 3 is based on paradigms like (3), where the interpretation of elided material intended is indicated by a paraphrase in angle brackets. Kennedy (1994) first observed that examples like (3a) are ungrammatical. The observation that (3b) contrasts with (3a) is new, and not predicted by any existing account of (2a).

(3) a. *Polly visited every town that's near the lake Eric did ⟨visit⟩.

b. Polly visited every town that's near the town Eric did ⟨visit⟩.

The examples in (3) involve ACD where the head of the ACD-relative is different from

the DP in the antecedent that corresponds to the trace in the elided VP. Contrasts like (3) show that, for the acceptability of the construction, the NP-parts of the two DPs involved—the head of the relative clause and the correspondent in the antecedent of the trace in the elided VP—must be lexically identical. The account of (3) argued for relies on the semantic form representations in (4). In (4), ACD is resolved by quantifier raising of the DP corresponding to the trace, such that the elided VP and its antecedent both contain a trace position. If the NP-part of the antecedent is lexically represented in both trace positions as in (4), the paradigm (3) follows from the identity requirement of VP-ellipsis. In (4a), the elided VP and its antecedent are not identical and, therefore, (4a) is predicted to be bad. In (4b), however, the elided VP and its antecedent are identical.

(4) a. *[every town that's near the lake Op E. $\overbrace{\text{visited [lake]}}^{\text{elided VP}}$] P. $\overbrace{\text{visited [town]}}^{\text{antecedent}}$

b. [every town that's near the town Op E. $\overbrace{\text{visited [town]}}^{\text{elided VP}}$] P. $\overbrace{\text{visited [town]}}^{\text{antecedent}}$

The account of (3) sketched here is developed in section 3.1. It's also shown there that the amount of lexical representation argued for by this account of Kennedy's observation is the same as that argued for by the distribution of Condition C. Section 3.2 slightly extends the paradigm in (3) and shows that the semantic properties of the material lexically represented in the trace position affect the severity of the ill-formedness in the case of mismatch. This is seen to argue that the lexical content of a trace position at semantic form is indeed interpreted in the trace position. Section

3.3 shows why the effect of the identity requirement on traces is usually not observed in cases of *wh*-movement other than ACD. Furthermore, the account there predicts one exception where the effect is found in examples with *wh*-movement; namely, (5) shows a contrast just like (3).

(5) a. I know which cities Mary visited, and now I would like to know the cities Sue did ⟨visit⟩

b. *I know which cities Mary visited, but I would like to know the lakes she did ⟨visit⟩

Also in section 3.3, I present an argument that the lexical material in the trace position is also represented in the head position of a chain, unless it contains a pronoun that isn't bound in the higher position. That is, I assume the semantic representation of (6a) to be (6b) where the NP-part of the *wh*-phrase is represented in both positions of the *wh*-chain.

(6) a. Which book does he like?

b. [which book] does he like [book]

The second part of the thesis considers the question how both positions of a chain are interpreted together. As mentioned already, the interpretation of a chain, and also that of a bound pronoun, is in a sense not compositional—the interpretation of two positions, the binder and the bound phrase, is intimately connected. A number of mathematical models has been proposed for the semantic mechanism that is at

work in such dependencies and the question is whether and how these views can be empirically distinguished. In chapter 4, I present several arguments in favor of λ-calculus as the model of the semantic mechanism underlying dependencies. The result applies both to chains and to pronoun binding. But, the content of the trace position, which is established in chapter 2 and 3, raises additional questions about the interpretation of chains. Chapter 5 considers the case of a chain headed by a quantificational DP, and develops a complete set of interpretive mechanism for this case. The main claim there is that the quantification ranges not over individuals, but over pointwise different choice functions.

The three mathematical models for the mechanism creating the semantic link in a chain that chapter 4 compares are λ-calculus, combinatorial logic, and predicate logic extended with restricted quantification. Since, as mentioned, pronoun binding seems to involve the same concept of semantic link as chains, the simplest assumption is that the same mechanism is involved in chains and pronoun binding. Because of this assumption, some of the arguments of chapter 4 (which are based on examples involving pronouns) carry over to chains as well.

Section 4.1 presents two arguments in favor of those models that involve variables (λ-calculus and extended predicate logic) and against combinatorial logic. Both arguments revolve around the abstract configuration sketched in (7). Consider the meanings of the two domains A and B, which both include a dependent element, but not the binder of it. I show that the contribution of the dependent elements to the meaning is identical for both domains on the combinatorial logic view. On the views that use variables, however, the two dependents could contribute different variable

names (or indices).

(7) binder ... [domain A: ... dependent ...] ... binder ... [domain B: ... dependent ...]

How can the meaning of such domains that aren't full sentences be investigated? Focus semantics has been argued to involve inference relationships between constituent domains smaller than sentences. Hence, if it can be determined which domains are considered in the licensing of focus and destressing, the meaning of domains like those in (7) can be studied. Section 4.1.1 presents an argument along these lines based on example (8) and on the fact that a focussed phrase must be contrastive to the corresponding phrase in the antecedent of its focus domain. In (8), the pronoun *his* in the second conjunct can optionally be focussed and, therefore, must differ in meaning from the pronoun *his* in the antecedent to be contrastive. If the focus structure of (8) is as indicated, the two pronouns can be different in meaning on the variables view, namely they can be interpreted as different variables. On the combinatorial view, on the other hand, the contrastiveness requirement cannot be satisfied by (8). Since the focus on the pronoun is optional in (8), other focus structures must be possible for (8). Section 4.1.2 presents a second argument for the variables view, based on a case where the focus structure is unambiguous, and therefore the focus on the dependent required.

(8) Every boy [antecedent: called his father] and every TEAcher [focus domain: called HIS father]

Section 4.2 attempts to draw a distinction between the λ-calculus view and the

extended predicate logic view. One argument for λ-calculus shows that the different variables of the tails of two chains are not contrastive when the domains considered are the sisters of the moved constituents. This result is expected if the variables are bound within the sisters of the moved constituents by corresponding λ-operators, but it's unexpected if the two moved constituents themselves bind the variables as on the predicate calculus view. A second argument presented in section 4.2 for λ-calculus comes from the distribution of i-within-i reference. I conclude therefore that the semantic form of (6a) is (9), where the translation from the syntactic logical form to the semantic form might contribute the variable index and the λ-operator.

(9) [which book] λx that he likes [x, book]

Chapter 5 addresses the question how the lexical content of the trace position, for example in (9), contributes to the interpretation of a chain. I consider exclusively the case of a chain headed by a quantificational determiner, and argue later that this might be the only case that arises. The approach developed is guided by the assumption that the semantic mechanisms should apply in the same manner to all chains headed by a DP. Specifically, no difference should be made between interrogative and non-interrogative DPs. Because of this assumption, examples like (10a), where the fronted DP contains a pronoun that's bound in the trace position, are an important case to account for. The semantic representation of (10a), that was argued for in the preceding chapters, is given in (10b). For the semantics, I adopt and extend the choice function approach of Engdahl (1980), which also relies on representations like

(10b).

(10) a. Which friend of her$_i$'s did every student invite?

b. which λx did every student invite [x, friend of her$_i$'s]

To extend the choice function proposal to non-interrogative, non-existential quantifiers, it turns out that it must be modified. The modification that proves most fruitful is a restriction of quantification to only pointwise different choice functions, instead of all choice functions, where the definition of pointwise different is given in (11).

(11) f is *pointwise different* from g if and only if $\forall x$: $f(x) \neq g(x)$

One prediction of the approach that seems desirable is discussed in section 5.2. It predicts that all DP-chains with lexical material in the trace position involve quantification over choice functions, while chains with no lexical material in the trace involve quantification over individuals. Since it was argued in 2.3 that the chains of the former type are most A-bar chains, while all A-chains are of the latter type, a type difference between A-bar chains and A-chains follows. If pronouns are of the type of individuals, it follows that the head of an A-bar chain cannot bind a pronoun, while the head of an A-chain can. In this way, the type difference is seen to predict the distribution of weak crossover effects.

The third part of the thesis is the shortest and most tentative. Chapter 6 presents two results that go some way towards establishing the claim that the semantic mechanisms developed in the previous part can account for the interpretation of all

chains that actually arise. Section 6.1 addresses a limitation of the mechanism of chapter 5, namely its restriction to DP-chains. I present arguments from the literature and one new argument based on facts from quantifier float in Japanese to show that only chains where the type of the trace is the type of individuals occur at LF. This implies that all occurring cases of chains are accounted for by the mechanisms already developed, where the variable ranges over choice functions and the type of the entire trace is that of individuals.

Section 6.2 addresses cases of so-called scope or total reconstruction, where a moved quantificational phrase takes scope in its trace position. In such cases, all lexical material of the chain is interpreted in the trace position, and none in the head. This, however, seems to require me to partially withdraw the claim of section 4.2 that the sister of a moved phrase is interpreted as a predicate. If the head of a chain is semantically empty, nothing would serve as the argument of this predicate. Though the required modification is rather trivial, section 6.2 presents an argument that the modification might not be needed at all. It is argued there that scope reconstruction phenomena should instead be analyzed as cases of movement in the phonological component of the grammar, which therefore doesn't have any semantic effect. Specifically, it's shown that the PF-movement proposal together with the assumption that movement must always target a c-commanding position makes a correct prediction; namely, the generalization that scope reconstruction is blocked in examples like (12) from Barss (1986) where the moved quantifier *some politician* doesn't c-command its trace.

(12) [How likely to t_{QP} address every rally]$_{wh}$ is [some politician]$_{QP}$ t_{wh}? (some$\gg$likely, *likely$\gg$some)

Chapter 2

Binding into Traces

At least since Ross (1967) and Lakoff (1968), it's been known that dislocated phrases can behave as if they were in their base position for the purposes of binding theory as in (1) and (2) below. This phenomenon, *Binding Reconstruction*, still remains only incompletely understood in many ways, but some significant properties of it have been discovered over the years: a correlation with the A/A-bar (or NP-movement/*wh*-movement) distinction (Wasow 1972:66,142,147–57, 1979:157–75, Riemsdijk and Williams 1981:204, Chomsky 1981),[1] a distinction between arguments and adjuncts of the moved phrase (Freidin 1986, Lebeaux 1988 with observations in Riemsdijk and Williams 1981, Chomsky 1981:144), a difference between overt and covert movement (Brody 1979, Chomsky 1981:196–197), a correlation between binding reconstruction for Condition C and for variable binding (Lebeaux 1992, 1995,

[1] In the literature on scrambling (e.g. Tada 1993), differences with respect to binding reconstruction are often the only criterion for the A/A-bar distinction. Hence, it may seem strange to speak only of a correlation. But, the A/A-bar distinction is needed independently for the statement of locality conditions on movement, in particular for overlapping paths (see Chomsky 1977, Rizzi 1990, Takano 1993, Müller 1993, 1998).

1998, Tada 1993:66-68, Chierchia 1995:129–170), and a correlation between narrow scope and binding reconstruction (Heycock 1995, Romero 1996, 1997, Fox 1996, 1997, 1998b, Sportiche 1996).

(1) [Which friend of her$_j$'s]$_i$ did every student$_j$ invite t_i?

(2) *[Which pictures of John$_j$]$_i$ did he$_j$ like t_i?

This chapter discusses facts that demonstrate that some parts of a moved phrase show binding reconstruction effects, but other parts of the moved phrase don't seem to reconstruct. Two kinds of evidence are new: One novelty is evidence that shows that the covert movement that resolves Antecedent Contained Deletion shows some binding reconstruction effects. The other is evidence that, when variable binding forces binding reconstruction of some part of the moved phrase, other parts of the antecedent can still escape binding reconstruction.

In the discussion of binding reconstruction phenomena, I adopt the assumption that binding reconstruction is syntactically represented by lexical material occupying the trace-position at the level where binding theory applies. Throughout, I use the notation exemplified in (3a) and (4): The relation between the lexical material that is represented in the trace position and the lexical material in the top position of the chain is expressed by a variable, x in the examples, which is part of the complex trace and bound in the position marked by the λ-operator, that marks the sister of the moved operator as a derived predicate. The use of λ-calculus as the mechanism mediating the dependency of operator and trace is argued for in chapter 4. For the

moment though, it should be seen as a typographically more convenient version of a notation like (3b) which is agnostic about the semantic mechanism mediating the dependency.

(3) a. $\underbrace{\text{Which}}_{\text{operator}}$ $\underbrace{\lambda x}_{\text{binder}}$ did every student$_j$ invite $\underbrace{[x, \text{friend of her}_j\text{'s}]}_{\text{complex trace}}$?

b. Which did every student$_j$ invite [friend of her$_j$'s]

The bound variable *her's* in (3) must be in the scope of its binder, and therefore I assume that in this case some of the lexical material is represented only in the trace position, as shown in (3). For (2), it's not clear whether to represent the lexical information *picture of John* only in the tail of the chain as in (4a), or to represent it doubly, in the head and the tail of the chain as in (4b) as suggested by Danny Fox (p.c.). To block coreference between *he* an *John* in (2), however, either of the representations in (4) suffices, since in both the pronoun *he* c-commands the R-expression *John*. (4b) may seem redundant, but on the other hand, it's more natural on the assumption that the syntactic operation underlying movement phenomena is copying, since it doesn't require as much deletion as (4a). In section 3.3.3, I present an empirical argument in favor of the latter view.

(4) a. *$\underbrace{\text{Which}}_{\text{operator}}$ λx did he$_j$ like $\underbrace{[x, \text{pictures of John}_j]}_{\text{complex trace}}$?

b. *$\underbrace{[\text{Which picture of John}_j]}_{\text{operator}}$ λx did he$_j$ like $\underbrace{[x, \text{pictures of John}_j]}_{\text{complex trace}}$?

In (3b) and (4b), binding reconstruction is represented by syntactic material that occupies the trace position. It is debated, though, whether the evidence necessitates this syntactic view (cf. Lebeaux 1992, 1995, 1998 Chierchia 1995, Fox 1998b, Romero 1997, Lechner 1998, Sharvit 1998 for discussion). An alternative view developed in detail by Barss (1986) assumes that a the head of a chain can be in the semantic scope of a phrase that c-commands its trace under certain conditions. On this view, the evidence in this section would be relevant towards stating restrictions on this chain-binding strategy. For now, I use lexical material in the trace position in the discussion of binding reconstruction as working hypothesis, though I suspect that it's less elegant to state the generalizations on the semantic reconstruction view. I refer the reader to Fox (1998b) for an overview of arguments in favor of the syntactic approach. The arguments in chapter 3 provide an additional argument for the syntactic approach taken here, and in chapter 5 I make a proposal as to how the lexical material in both the head and the tail position of a chain contributes to its interpretation.

The goal of this chapter is to investigate what parts of the head of a chain undergo binding reconstruction when the chain itself exists at the level of logical form. In other words, since I assume that binding reconstruction is represented by lexical material in the trace position, the goal of this chapter is to find out where what parts of the lexical material of a chain are at the level where binding theory applies. In particular, I present new arguments that the lexical material of a DP-chain can in some cases be split between the top and the bottom position, in ways other then the split between the quantificational D-head and its NP-complement postulated in (3)

and (4). This kind of situation is sketched in (5).

(5) $\underbrace{[\text{D 'some lexical material'}]}_{\text{operator}}\ \lambda x\ \ldots \underbrace{[x, \text{ 'other lexical material'}]}_{\text{trace}}$

The derivation of LF-representations like (5) is straightforward assuming the copy theory of movement (Wasow 1972, Chomsky 1981, Burzio 1986, Chomsky 1995). If syntactic movement creates representations where all positions of a chain contain all the lexical material of the moving phrase, the distribution of lexical material in representations of the kind sketched in (5) can be derived by the application of deletion in the positions of a chain. The goal of this chapter is then to find the conditions under which this deletion operation applies.

Engdahl (1980:131–144) was the first to articulate the claim that binding reconstruction can cooccur with wide scope. The empirical evidence she gives for the LF in (6a) (I discuss her semantics in Chapter 5) is of the type in (1): *wh*-questions where the *wh*-phrase contains a bound variable pronoun. Her argument therefore relies on two assumptions: that bound variable pronouns must be in the scope of their antecedents and that wh-phrases take sentential scope. There's hardly any alternative to the first assumption (Though, see the discussion of Skolem-functions in Engdahl (1986) and Chierchia (1993)). The second assumption is much harder to argue for, and indeed not universally assumed—for example Hamblin (1973), Cresti 1997 and Rullmann and Beck (1997) entertain LF-representations like (6b) for (1) where the *wh*-chain of the overt form is not represented at all. Ultimately, I believe Engdahl's (1980) interpretation of (1) is correct, but it will take some effort to get there.

(6) a. which λx did every student$_i$ invite [x, friend of her$_i$'s]?

b. did every student$_i$ invite $\underbrace{\text{which friend of her}_i\text{'s}}_{\text{wh-phrase}}$?

To sharpen Engdahl's argument, we need better tests for the location of the expression heading the chain that don't reconstruct. In addition to variable binding, the three tests relevant for this chapter are Quantifier Scope, Antecedent Contained Deletion (Sag 1976, May 1985, Larson and May 1990), and Condition C of the Binding Theory (Ross 1967, Langacker 1969, Lasnik 1976, Chomsky 1981).

2.1 Scope, Condition C and Antecedent Contained Deletion

Before applying Condition C to test for the LF-position of syntactic material, I'll summarize an argument from Fox (1995b) that Condition C applies at LF only. His argument relies on the contrast in (7) (from Fiengo and May 1994:296 with some modifications). (In (7) and in the following, I indicate the interpretation of a VP-ellipsis site by a paraphrase given in angle brackets $\langle\rangle$.)

(7) a. You introduced him$_i$ to everyone John$_i$ wanted you to $\langle$introduce him$_i$ to$\rangle$

b. *I introduced him$_i$ to everyone John$_i$ wanted you to dance with.

In both examples in (7), the pronoun *him* c-commands the R-expression *John* in the surface form. While Condition C rules out coreference between *him* and *John* in (7b), Condition C doesn't seem to apply in (7a) despite the surface c-command

of *John* by *him*. In (7a), coreference between the pronoun and the R-expression is acceptable. Fox's (1995b) as well as Fiengo and May's (1994) account of (7) relies on the account of Antecedent Contained Deletion (henceforth ACD) of Sag (1976:73), May (1985) and Larson and May (1990). Sag and Larson and May, in particular, argue, based on an interaction between scope and ACD, that quantifier movement is required for the resolution of ACD. This is displayed by the LF-representation in (8).

(8) $\Big[$everyone $[\lambda y$ John$_i$ wanted you to introduce him$_i$ to $[y]]\Big]$

λx you introduced him$_i$ to $[x]$.

In the LF-representation (8), the R-expression *John* is no longer in the c-domain of the pronoun *him*, and therefore Condition C is not violated. If Condition C doesn't apply to the surface representation (7a), but to the LF-representation (8), it's expected that coreference between *him* and *John* is possible. I leave the question why quantifier movement cannot obviate Condition C in (7b) for the next paragraph. Assuming that there's an answer to this question, the contrast in (7) is a strong argument that Condition C applies only at LF (see Fox 1995b for arguments against the account of (7) in Fiengo and May 1994).

Why quantifier raising doesn't always bleed Condition C, for example not in (7b), is the remaining question. It turns out that (7b) is actually not a good example to raise the question for because (7b) doesn't control for the scope of the universal quantifier and, if Fox (1995a) is right, quantifier movement is blocked in (7b). However, it is well known that even when a universal quantifier takes wide scope,

quantifier movement usually doesn't obviate Condition C (Brody 1979, Chomsky 1981:196–7). This is shown by (9), where even on the $\forall \gg \exists$ wide scope reading, the pronoun *him* cannot be coreferent with the R-expression *John*. The contrast between (9) and (7a) cannot be explained by conditions on whether quantifier movement applies.[2] Rather, the contrast must be explained by an interaction between ACD and Binding Reconstruction, namely that ACD blocks Binding Reconstruction.

(9) *Someone introduced him$_i$ to everyone John$_i$ wanted you to dance with.

As Fox (1995b) points out, the Copy Theory of Binding Reconstruction together with standard theories of ACD predict that ACD blocks Binding Reconstruction: The relevant standard assumption about ACD and VP-ellipsis in general is that an elided VP must be (almost) identical to its antecedent (Sag 1976, Williams 1977, May 1985, Tancredi 1992, Rooth 1992b, Wold 1995, Fiengo and May 1994). For now, I assume that exact lexical identity is required except for tense morphology (see section 3.2 and 3.3 though). Consider (10) which is the LF-representation of (7a) where the elided VP and its antecedent are indicated. In (10) without Binding Reconstruction, the elided VP and its antecedent are identical.

[2]To maintain the view that it is the application of quantifier movement that accounts for the contrast between wide scope as shown by scope and wide scope for the resolution of ACD, the following assumptions would need to be postulated: There is a second scope taking mechanism in addition to quantifier movement which yields wide scope, but doesn't resolve ACD. Furthermore, quantifier movement applies only to resolve ACD, while the other mechanism applies yields wide scope when there's no ACD. A mechanism with exactly these properties has been formulated in the literature, namely Quantifier Storage (Cooper 1983), though not for this particular problem. But, consistent though it is, such an account is quite clearly not more than a restatement of the facts.

(10) $\Big[\overbrace{\text{everyone } [\lambda y \text{ John}_i \text{ wanted you to } \underbrace{\text{introduce him}_i \text{ to } [y]}_{\text{elided VP}}]}^{\text{operator}}\Big]$

λx you $\underbrace{\text{introduced him}_i \text{ to } \overbrace{[x]}^{\text{trace}}}_{\text{antecedent}}$.

But if there's a copy of the lexical material of the relative clause in the bottom position of the chain of quantifier movement as indicated in (11), the elided VP and its antecedent do not satisfy the identity condition. Hence, the representation (11) is not possible for the sentence (7a). This way, the identity requirement on an elided VP and its antecedent blocks Binding Reconstruction into traces that are part of the antecedent in ACD constructions.

(11) $^{*}\overbrace{\text{everyone } [\lambda y \text{ John}_i \text{ wanted you to } \underbrace{\text{introduce him}_i \text{ to } [y]}_{\text{elided VP}}]}^{\text{operator}}$

λx you introduced him$_i$ to $\underbrace{\overbrace{[x, [\lambda y \text{ John}_i \text{ wanted you to } \underbrace{\text{introduce him}_i \text{ to } [y]}_{\text{elided VP (copy)}}]}^{\text{trace}}}_{\text{antecedent}}$.

In (9), on the other hand, Binding Reconstruction is possible. The representation corresponding to Binding Reconstruction is given in (12), which then violates Condition C. If the representation (12) is actually the only one available for (9), the contrast between quantifier raising for ACD resolution and for wide scope alone is accounted for.

(12) $^{*}\Big[\text{everyone } [\lambda y \text{ John}_i \text{ wanted you to dance with } [y]]\Big]$ λx someone introduced him$_i$ to [x, [λy John$_i$ wanted you to dance with [y]]]

To this end, Fox (1995b) adopts the assumption argued for in Chomsky (1993:36-37) that there is a preference for A-bar chains to represent all lexical material except for the D-head in the trace position—in effect, a preference for Binding Reconstruction. After Chomsky (1993), this is often referred to as the *Preference Principle*; it states a preference between two representations: the one with the lexical material interpreted at the top of the chain, and the one with the lexical material interpreted at the bottom of the chain. Fox (1995b) reformulates the Preference Principle as a economy condition and adds that it can be overridden by the requirement of ACD that the relative clause containing the elided VP must be represented outside of the antecedent for deletion. Condition C is different from the requirement of ACD in that it doesn't motivate a violation of the Preference Principle. The interaction between ACD and the preference principle is discussed further at the end of section 3.2.

In addition to the argument that Condition C applies at LF, Fox's (1995b) analysis has implications concerning what are possible LF-representations of A-bar chains, and what preferences exist between them. At least two representations are possible for an A-bar chain consisting of a quantificational operator and lexical material restricting the operator: a representation where the entire restrictor occupies the trace position of the A-bar chain, and a representation where at least a relative clause modifying the restrictor doesn't occupy the trace position, but only occupies the head position. Of these, the former representation is preferred by the grammar and the second one is only used in case ACD blocks the first one.[3] As for the structure of the

[3]As Danny Fox (p.c.) pointed out to me, this view predicts that quantifier movement should always obviate Condition C if the R-expression triggering Condition C is part of the quantificational determiner. It's very difficult to construct relevant examples, but a weak preference in the predicted

representation in the case of ACD, Fox's (1995b) analysis of (7a) gives us information about the LF-position of the R-expression, which is the subject of the ACD relative clause. Namely, it must be in a position higher than its surface position. This follows from the assumption that the entire relative clause must be in a position higher than its surface position, which is required in the ACD-cases for ACD resolution. The evidence, however, leaves it open whether all lexical material of the DP that moves for ACD resolution is represented in the top position of the A-bar chain or only as much as is needed for ACD-resolution. In effect, Fox (1995b) develops both views.

Consider now the contrast in (13), which resolves the open issue. (13a) is comparable to (7a): The R-expression that is in a Condition C configuration in the surface form is the subject of an ACD relative clause just like in Fiengo and May's (1994) example above: Condition C is bled by ACD-resolving Quantifier Movement. In (13b), on the other hand, the R-expression is part of the noun phrase that has to move for ACD-resolution, but it's not inside of the ACD relative clause. The new discovery is that the Condition C effect remains in (13b).

(13) a. In the end, I did ask him$_i$ to teach the book of Irene's that David$_i$ wanted me to ⟨ask him to teach⟩.

direction seems to be detectable in (i). In (ia), it seems that Condition C is obviated on the reading where *John's every* takes scope over *someone*. In (ib), quantifier raising of *John's every* to a position above the subject is blocked because the subject is not a scope bearing element (Fox 1995a). As expected, Condition C cannot be obviated in (ib).

(i) a. Someone must've fed him$_i$ John$_i$'s every move over earphones.
b. *Kasparov must've fed him$_i$ John$_i$'s every move over earphones.

b. *In the end, I did ask him$_i$ to teach the book of David$_i$'s that Irene wanted me to ⟨ask him to teach⟩.

Merchant (1998a) independently discovered similar facts and makes the same point. I give his three examples in (14). In each of them, Condition C applies to block coreference between the object pronoun and the R-expression that occurs in the external head of the ACD-relative clause.

(14) a. *I gave him$_i$ every report on Bob$_i$'s division you did ⟨give him$_i$⟩.

b. *I reported her$_i$ to every cop in Abby$_i$'s neighborhood you did ⟨report her$_i$ to⟩.

c. *I showed her$_i$ every picture from Abby$_i$'s mantlepiece you did ⟨show her$_i$⟩

The examples in (15) show that coreference is possible if Merchant's (1998a) examples are modified such that the R-expression is inside of the ACD-relative clause. In (15b) and (15c), the R-expression is part of the subject of the ACD-relative, like in (7a), while, in (15a), the R-expression is part of the material pied-piped with the relative operator. Note that, while in (13) the two sentences compared differ with respect to the amount of material intervening between the pronoun and the R-expression relevant for Condition C, the contrast among (14b) and (14b) could not be explained in terms of such a difference.

(15) a. I gave him$_i$ every report whose section on Bob$_i$'s division you asked me to ⟨give him$_i$⟩.

b. I reported her$_i$ to every cop Abby's neighbors allowed me to ⟨report her$_i$ to⟩.

c. I showed her$_i$ every picture Abby$_i$'s agent forgot to ⟨show her$_i$⟩

Example (13b), and also Merchant's (1998a) examples in (14), show that ACD doesn't block binding reconstruction of lexical material that is not part of the ACD relative clause. On the Copy Theory of Binding Reconstruction, this lexical material—*book of David's* in the example (13b)—must therefore occupy the trace position, and thereby causes a Condition C violation. In fact, a structure with such a complex trace must be forced in the the examples in (13), since the Condition C violating structure is the only one available for (13b).

As (16), a tentative LF-representation for (13a), illustrates, if we assume that only quantifier movement leaves such a complex trace, the elided VP is not identical to its antecedent anymore, because the relative clause internal trace is a simple variable while the QR-trace is complex. At least if VP-ellipsis requires identity of lexical material (see chapter 3), (16) cannot be the LF representation of (13a).

(16) $\Big[$the book of Irene's λy David$_i$ wanted me to $\underbrace{\text{ask him}_i \text{ to teach } [y]}_{\text{elided VP}}\Big]$

λx I $\underbrace{\text{asked him}_i \text{ to teach } [x, \text{ book of Irene's}]}_{\text{antecedent}}$

Since (13a) allows VP-ellipsis, the relative clause internal trace must be lexically complex, containing lexical material of the noun phrase it attaches to. I present further arguments towards this conclusion in section 2.4. For (13a), this allows the LF representation in (17), which satisfies the identity condition of VP-ellipsis.

(17) $\Big[$the book of Irene's λy David$_i$ wanted me to $\underbrace{\text{ask him}_i \text{ to teach } [y, \text{ book of Irene's}]}_{\text{elided VP}}\Big]$

λx I $\underbrace{\text{asked him}_i \text{ to teach } [x, \text{ book of Irene's}]}_{\text{antecedent}}$

For (13b), these consideration force the LF representation in (18). In (18), the quantifier movement that resolves ACD left the noun phrase part *book of David's* in the trace position, where the pronoun *him* c-commands it. Therefore, Condition C is violated in (18), and coreference between *him* and *David* is correctly ruled out.[4]

(18) $^*\Big[$the book of David$_i$'s λy Irene wanted me to $\underbrace{\text{ask him}_i \text{ to teach } [y, \text{ book of David}_i\text{'s}]}_{\text{elided VP}}\Big]$

λx I $\underbrace{\text{asked him}_i \text{ to teach the } [x, \text{ book of David}_i\text{'s}]}_{\text{antecedent}}$

In the contrast in (13), the offending R-expression was an argument of the head noun the determiner takes as complement. The examples in (19) exhibit a similar contrast to (13), but the offending R-expression is part of an adjunct in (19b). This shows that there's no difference between adjuncts and arguments in the construction

[4]Actually, Condition C seems to actually be violated twice in (17): not only the instance of *David* in the trace of quantifier raising, but also the instance of *David* in the relative clause internal trace is c-commanded by a coreferent pronoun. While this extra violation causes no problem in (17), in general no Condition C reconstruction effect is found in a relative clause, as (i) exemplifies. I address this issue in section 2.4.

(i) The book of Bill$_i$'s that he$_i$ was working on since 1971 finally appeared.

we're considering here.

(19) a. In the end, we did advise him$_i$ to buy the computer compatible with Bev's that Noam$_i$ hoped we would.

b. *In the end, we did advise him$_i$ to buy the computer compatible with Noam$_i$'s that Bev hoped we would.

The two results of this section are worth repeating once again: One, an argument of Fox (1995b) was summarized which shows that Condition C of the Binding Theory applies at LF only. Secondly, an argument was presented that even when some of the material of the covertly A-bar moved phrase is missing from the trace position, other lexical material still seems to occupy the bottom position of the A-bar chain.

2.2 Variable Binding and Condition C

In this section, I'll present more arguments that structures with part of the lexical material of an A-bar chain interpreted in the bottom position and other parts interpreted only in the top position are possible LF-representations. The evidence here will be based on interactions between variable binding and Condition C, which I from now on assume to apply at LF only following Fox (1995b).

In fact, if Condition C applies at LF only, the distribution of Condition C effects alone with overt A-bar movement is evidence for representations where the lexical material is split between two positions. The relevant observation is the well

known contrast between (20a) and (20b) (Riemsdijk and Williams 1981, Freidin 1986, Lebeaux 1988), which is usually seen as an argument-adjunct distinction. The fact is that the R-expression inside an argument of the noun head causes a Condition C violation by Binding Reconstruction into the trace position, making (20a) ungrammatical. The R-expression in (20b), however, which is contained in an adjunct, doesn't cause a Condition C effect in the trace position.

(20) a. *[Which argument that John$_i$ was wrong]$_j$ did he$_i$ accept t_j in the end?

b. [Which argument that John$_i$ had criticized]$_j$ did he$_i$ accept t_j in the end?

The contrast in (21) makes the same point as (20), but controls for structural differences between the *wh*-phrases.[5] The only difference between (21a) and (21b) is that in (21a) the R-expression *John* occupies an argument position of *argument* and *Mary* is the subject of the adjoined clause, but in (21b) the two are switched around. Only when the R-expression is in the argument position in (21a), does it cause a Condition C violation via Binding Reconstruction into the trace position.

(21) a. *[Which argument of John$_i$'s that Mary had criticized] did he$_i$ omit t_j in

[5]However, (20) controls for the depth of embedding better than (21). It's known that the strength of a Condition C violation correlates with the distance and depth of embedding (Chomsky 1981:196–7). Differences in the severity of Condition C as illustrated in (i) are expected on the basis of general processing conditions and in addition the fact that examples like (ia) might also constitute a Condition B violation (Kuno 1997). Therefore, it's important to control for the depth of embedding whenever possible.

(i) a. *He$_i$ liked John$_i$.

b. *?He$_i$ liked that Mary bought a picture of John$_i$.

c. *?He$_i$ liked that John$_i$'s grandfather's stories were popular.

the final version?

b. [Which argument of Mary$_i$'s that John had criticized] did he$_i$ omit t_j in the final version?

The LF-representations the Copy Theory would assign to (21a) and (21b) are given in (22a) and (22b) respectively. Condition C is violated only in (22a), where the lower trace contains the R-expression *John*.[6]

(22) a. $^*\Big[$Which argument of John$_i$'s [λy that Mary had criticized [y]]$\Big]$ λx he$_i$ omit [x, argument of John$_i$'s] in the final version

b. $\Big[$Which argument of Mary's [λy that John$_i$ had criticized [y]]$\Big]$ λx he$_i$ omit [x, argument of Mary's] in the final version

The representations in (22) show that whether a Condition C violation is incurred or not is a function of where in the fronted DP the R-expression appears. If the R-expression is inside the NP that's the complement of *which*, it triggers Condition C in the trace position. If the R-expression occurs inside a relative clause that modifies this NP, it doesn't trigger Condition C in the trace position. In stating this generalization, it is useful to have a term for the part of a DP that is the complement of the determiner excluding relative clauses and other adjuncts that are adjoined to it: Henceforth, I call this the *NP-part* of a DP. In (23), I have marked the determiner, the NP-part and modifiers to it of the *wh*-phrase of (21a). Using this terminology, the contrast in

[6] I'm not representing the lexical content of the relative clause internal trace at this point because it's irrelevant here.

(21) argues that the NP-part of a *wh*-movement chain must reconstruct to the trace position for Condition C. This is stated as a preliminary generalization in (24).

(23) $\underbrace{\text{which}}_{\text{Det.}}\ \underbrace{\text{argument of John's}}_{\text{NP-part}}\ \underbrace{\text{that Mary had criticized}}_{\text{modifier}}$

(24) The NP-part must occupy the position of an A-bar trace. A modifier may, occupy the position of an A-bar trace.[7]

The generalization in (24) assumes that only modifiers adjoined to the NP-part itself, but not modifiers internal to the NP-part, can escape binding reconstruction in *wh*-chain. Based on different assumptions, Tada (1993:65) arrives at this conclusion and presents evidence for it from Japanese. In (25), though the R-expression *John* occurs inside a relative clause and outside the c-domain of the pronoun *kare* in the surface form, coreference is blocked by Condition C. (25) is predicted if only modifiers adjoined to the NP-part of the moving phrase itself can escape Condition C.[8]

(25) $^{*?}$[John$_i$-ni ki-ta tegami-o suteru-yooni]$_j$ kare$_i$-ga tsumani t_j miji-ta
[John$_{\text{DAT}}$ came letter$_{\text{ACC}}$]$_j$ throw-away he$_{i\text{NOM}}$ wife$_{\text{DAT}}$ t_j ordered.

'To throw away the letter that came to John, he told his wife.'

[7]Lebeaux (1988) proposes counter-cyclic adjunction of relative clauses as an explanation for the Condition C obviation of overt movement. Recall though from the previous section, that covert movement also displays obviation of Condition C with relative clauses, if the covert movement is required for Condition C resolution. Hence, Lebeaux's (1988) explanation is at least incomplete. I come back to the question Lebeaux's proposal at the end of this section.

[8]Note that Tada's account predicts that, for a relative clause inside a fronted predicate, overt movement won't obviate Condition C. That this prediction is correct is shown by the examples in (i) from Takano (1995:(12)) (see also Heycock (1995)). Since Tada's explanation of (i) isn't dependent on the VP-internal trace hypothesis, I conclude contrary to Takano (1995) and Heycock (1995) that examples like (i) don't bear on this hypothesis.

(i) a. Criticize a student that John$_i$ taught, he$_i$ said Mary did.
b. How proud of a student that John$_i$ taught did he$_i$ say Mary is?

However, Tada (1993:fn. 25) doubts the existence of a difference between modifiers to the NP-part and modifiers internal to the NP-part for English because of (26a) (attributed to Noam Chomsky, p.c.). In (26a), Condition is obviated even though the relative clause containing the R-expression *John* is adjoined to the lower NP *book*. This argument isn't convincing because the lower NP itself could be part of an modifier if we assume that the *for*-PP is an modifier to *book*. The obviation of Condition C in (27a) shows that the *for*-PP is an adjunct, as does the separability in the copular paraphrase in (27b) (see Schütze 1995).

(26) a. The award for the book that John$_i$ wrote, he$_i$ never received.

b. The award for the book that John$_i$ received, he$_i$ never cashed.

(27) a. Which award for *Titanic*$_i$ did everybody agree it$_i$ deserved.

b. The award was for the book.

The contrast in (28) shows that only modifiers adjoined to the NP-part of the moved phrase can escape binding reconstruction. The R-expression *Bill* occurs inside a relative clause in both, (28a) and (28b). However, there's a contrast depending on whether this relative clause is part of an argument inside the NP-part of the fronted phrase as in (28a), or a modifier to this NP-part.

(28) a. *Which $\underbrace{\text{book of the woman } \overbrace{\text{Bill}_i \text{ admires}}^{\text{rel. clause}}}_{\text{NP-part}}$ did he$_i$ give to his$_i$ parents.

b. Which $\underbrace{\text{book}}_{\text{NP-part}}$ $\underbrace{\text{about the woman } \overbrace{\text{Bill}_i \text{ admires}}^{\text{rel. clause}}}_{\text{modifier}}$ did he$_i$ give to his$_i$ parents.

Another way to enforce binding reconstruction is variable binding, which brings us back to Engdahl's paradigm mentioned at the beginning. As is well known, overt *wh*-movement allows a binding of variable inside the moved material by a quantifier that c-commands the trace position, as in (29a). The ungrammaticality of (29b) shows that c-command of the trace position is indeed necessary.

(29) a. [Which paper of his$_j$]$_i$ did every student$_j$ plan to revise t_i?

b. *[Which paper of his$_j$]$_i$ t_i earned every student$_j$ praise?

On the copy theory of binding reconstruction, the LF-representation this leads us to postulate for (29a) is (30). The variable *his* is interpreted in the bottom position of the A-bar chain where it's c-commanded by the quantifier *every student*.

(30) Which λx every student λy [y] planned to revise [x, paper of his$_y$]

That the representation (30) is correct is shown by the interaction between Binding Reconstruction for Variable Binding and Binding Reconstruction for Condition C Lebeaux (1992) observes.[9] The contrast in (31), which is from Lebeaux (1992) with minor changes, shows this interaction.

(31) a. [Which paper that he$_k$ gave to Mary$_j$]$_i$ did every student$_k$ think t'_i that she$_j$ would like t_i?

[9]Tada (1993:66-68) discusses an interaction between temporal dependencies and Condition C that makes the same point as Lebeaux's data. Chierchia (1995:129-170) shows data with fronted conditionals that supports Lebeaux's conclusion as well.

b. *[Which paper that he$_k$ gave to Mary$_j$]$_i$ did she$_j$ think t'_i that every student$_i$ would like t_i?

In (31a), variable binding can be satisfied via reconstruction in the position t'_i, which is c-commanded by the antecedent of *he*, namely *every student*, but not c-commanded by the pronoun *she*. Therefore, *she* doesn't trigger a Condition C effect in this position, and *she* and the R-expression *Mary* can be coreferent. The interaction observed by Lebeaux is predicted by the Copy Theory view of reconstruction, as is shown by the LF-representation in (32). The relative clause which contains both the bound variable pronoun and the the R-expression is interpreted in the intermediate trace position that is not c-commanded by the R-expression *Mary*.

(32) $\underbrace{[\text{Which paper}]}_{\text{operator}}$ λx every student$_i$ think $\underbrace{[x, \text{ paper}, \lambda z \text{ he}_i \text{ gave } [z] \text{ to Mary}_j]}_{\text{intermediate trace}}$ λy she$_j$ would like $\underbrace{[y, \text{ paper}]}_{\text{lowest trace}}$

In (31b), on the other hand, all reconstruction positions c-commanded by *every student* are also c-commanded by the R-expression *Mary*. If Binding Reconstruction for variable binding always forces Binding Reconstruction for Condition C to take place as well, variable binding is correctly predicted to be blocked in (31b). The Copy Theory captures the interaction observed; namely, that Binding Reconstruction for variable binding forces Binding Reconstruction for Condition C as well. The representation in (33) illustrates that e.g. interpreting the copy of the relative clause in the lowest trace position leads to a Condition C violation.

(33) *$\underbrace{[\text{Which paper}]}_{\text{operator}}$ λx she$_j$ thinks $\underbrace{[x, \text{paper}]}_{\text{intermediate trace}}$ λy every student$_i$ would like $\underbrace{[y, \text{paper } [\lambda z \text{ he}_i \text{ gave } [z] \text{ to Mary}_j]]}_{\text{lowest trace}}$

As we also saw in the interaction of ACD-resolution and Condition C above, Lebeaux's (1992) data demonstrate that a relative clause cannot be split among different positions of a chain. Otherwise it should be possible to interpret the bound variable in a position lower than the the R-expression *Mary* and thereby accomplish variable binding without violating Condition C. To find out whether variable binding can also be satisfied by binding reconstruction of a part of a fronted constituent, we need to test cases where we know the relevant parts of the fronted constituent can be interpreted in different positions of the A-bar chains. Such cases have not been studied in the previous literature.

The contrast in (34) shows that it is possible to accomplish variable binding by reconstructing only parts of a fronted constituent. (34b) has the same structure as Lebeaux example in (31b) and, as above, variable binding into the relative clause brings about a Condition C violation. In (34a), on the other hand, the bound variable is not part of the relative clause, and therefore reconstruction of the relative clause isn't forced. Therefore, (34a) doesn't violate Condition C. Example (35) makes the same point as (34).

(34) a. [Which paper of his$_k$ that Mary$_j$ was given]$_i$ did she$_j$ tell every student$_k$ to revise t_i?

b. *[Which paper that he$_k$ gave to Mary$_j$]$_i$ did she$_j$ tell every student$_k$ to revise t_i?

(35) a. [Which of his$_k$ pictures that Mary$_j$ was shown]$_i$ did she$_j$ return to t_i every student$_k$.

b. *[Which picture that he$_k$ showed to Mary$_j$]$_i$ did she$_j$ return t_i to every student$_k$.

The LF-representation of (34a) is shown in (36). The bound variable *his* is interpreted in the bottom position of the A-bar chain, while the R-expression *Mary* with the relative clause is interpreted in the top position, such that Condition C isn't violated.[10]

(36) $\Big[$Which [λz Mary$_j$ was given [z]]$\Big]$ λx did she$_j$ tell every student$_i$ to revise [x, paper of his$_i$]?

A second case of variable binding taking place in a lower position than the interpretation of the relative clause is found when there are two relative clauses. (37b) shows that it is possible to interpret one relative clause in a low position to achieve variable binding, and at the same time represent the second relative clause only in a higher position, such that Condition C isn't violated.

[10]Notice that here the NP-part *paper of his$_i$* cannot be represented inside the relative clause because the bound variable *his* would not be bound. In section 2.4, I present an analysis of relative clauses that predicts this.

(37) [Which computer $\overbrace{\text{compatible with his}_j}^{\text{inner modifier}}$ $\overbrace{\text{that Mary}_i \text{ knew how to use}}^{\text{outer modifier}}$]$_k$ did she$_i$ tell every boy$_j$ to buy t_k.

While (37) confirms the claim that two modifiers can be represented at LF in different positions of a chain, (38a), where the position of bound variable and R-expression is exchanged, points to a complication. Here, Condition C is violated even though the bound variable and the R-expression relevant for Condition C occur in different relative clauses. Given the contrast to (37), it seems as if reconstruction of the outer relative clause forces reconstruction of the inner relative clause to take place as well.

(38) *[Which computer $\overbrace{\text{compatible with Mary}_i\text{'s}}^{\text{inner modifier}}$ $\overbrace{\text{that he}_j \text{ knew how to use}}^{\text{outer modifier}}$]$_k$ did she$_i$ tell every boy$_j$ to buy t_k?

It is hard to decide whether the presence of a bound variable in the outer modifier in (38) is among the causes of the Condition C effect. Even (39) is not very good though here the outer modifier doesn't contain a bound variable. But, there seems to be a slight contrast between (39) and (38).

(39) ??[Which computer $\overbrace{\text{compatible with Mary}_i\text{'s}}^{\text{inner modifier}}$ $\overbrace{\text{that I knew how to use}}^{\text{outer modifier}}$]$_k$ did she$_i$ tell Tom to buy t_k?

In (40), an example where the R-expression is part of the NP-part is used as an additional item of comparison. It seems that (40b), where the R-expression is part

of an inner modifier and the outer modifier doesn't contain a bound variable, allows coreference more easily than (40a), where the R-expression occurs in the NP-part, or (40c), where the the outer modifier contains a bound variable.

(40) a. *Tell me which descriptions of Kant$_i$'s views that were published every woman said he$_i$ agreed with?

b. ?Tell me which books describing Kant$_i$'s views that were published every woman said he$_i$ agreed with?

c. *Tell me which books describing Kant$_i$'s views that she$_j$ published every woman$_j$ said he$_i$ agreed with?

Therefore, I conclude that the ordering effect between (37) and (38) is real, though it isn't predicted by anything said so far. I think the effect might shed light on the question why relative clauses can escape binding reconstruction with overt movement. Namely, the effect is predicted on a modification of Lebeaux's (1988) proposal Tada (1993:63-70) develops. Lebeaux's (1988) proposal is that relative clauses can adjoin to a *wh*-phrase after it has undergone *wh*-movement and for this reason need not reconstruct for binding to the bottom position of an A-bar chain. Essentially, Lebeaux proposes that adjunction need not obey the syntactic cycle at all. Tada (1993), however, proposes rather than to abandon the cycle, to modify it to accommodate Lebeaux's cases. In effect, Tada proposes that adjunction obeys the cycle, but that adjunction to the specifier of the current cyclic domain is consistent with the cycle. On Tada's proposal, modifiers adjoining to a moved phrase must obey the

cycle with respect to the phrase they are adjoining to. Then, the ordering effect is predicted: The cycle then makes sure that the inner relative clause must be adjoined before the outer relative clause. If reconstruction to the position where a relative clause was first adjoined is forced, the order of adjunction determines that the inner relative clause must reconstruct at least as low as the outer relative clause. Note that this account of (40) supports the central claim of Lebeaux's (1988) account that the reason Condition C can be obviated with overt movement is late adjunction. In the previous section, we saw that also deletion of adjuncts at LF can cause obviation of Condition C. Because of (40), I conclude that both mechanisms are needed.

The main point of this section, however, is that even in cases where variable binding forces binding reconstruction of parts of a chain, others parts of the chain don't have to reconstruct. More specifically, example (34) showed that even the NP-part of a *wh*-phrase can reconstruct while a relative clause adjoined to it can still occupy a higher position. Cases like (34) will be important for the semantics of chains in chapter 5, because in these cases the dependency between the two positions of the chain must be more complex because the meaning of the complex trace depends on the value of the bound variable.

2.3 The A/A-bar Distinction

The previous two sections were concerned with covert quantifier movement chains and overt *wh*-movement chains. While there were differences between overt and covert A-bar movement with respect to relative clause modifiers, the NP-part of the moved

phrase was always represented in the trace position. Recall that an R-expression that is part of the NP-part always triggers a Condition C effect in the trace position, as illustrated by (41a): The pronoun *he* that c-commands the trace position cannot be coreferent with the R-expression *Kai*. It is well known that A-chains differ from A-bar chains in this respect. Namely, an R-expression that is part of the NP-part of an A-moved phrase doesn't trigger a Condition C effect in the trace position. This is illustrated in (41b), where the R-expression *Kai* and the pronoun *him*, which c-commands the A-trace t_i, can be coreferent. (I'm concerned with the interpretation of (41b) where *one* takes scope over *seem*. In case *seem* takes scope over *one*—the case of Scope Reconstruction—, a Condition C effect is found as Fox (1997) and Romero (1997) show. See also section 6.2.)

(41) a. *[Which relative of Kai$_j$'s]$_i$ did he$_j$ say t_i likes Kazuko.

b. [One relative of Kai$_j$'s]$_i$ seemed to him$_j$ to t_i like Kazuko.

On the view that binding reconstruction is represented by lexical material in the trace position, the fact (41b) indicates that in A-chains no lexical material of the head is represented in the trace position. On the copy theory of movement, the behavior of A-chains seems unexpected, since nothing seems to motivate deletion of the lexical material in the trace position. Recall though from the previous section that, while the NP-part was always represented in the trace position in A-bar chains, relative clause modifiers generally weren't required to be represented in the trace position in chains created by overt *wh*-movement. Then, (41b) shows that the NP-part in an A-chain,

which is created by overt movement, behaves in the same way that modifiers behave with overt A-bar chains.

At this point, there are various ways to state the difference between A-chains and A-bar chains. It seems to me that the difference between the NP-part and modifiers in an A-bar chain is unexpected because semantically both the NP-part and the modifiers are alike: they contribute predicates that form the restrictor of the quantificational determiner which is heading the moving DP. Hence, I propose to capture the difference between A-chains and A-bar chains by means of the condition in (42), which stipulates the unexpected behavior of the NP-part in A-bar chains.

(42) In A-bar chains, the NP-part of the moving DP must be represented in the lowest trace position.

Obviously it's desirable to derive (43) from something, but, at this point, I must relegate the issue to future research. I hope to show, however, that the difference between A and A-bar chains at the level of logical form can be reduced to the condition in (42). In the remainder of this section, I present some tentative results that relate to this project concerning the distribution of Condition C effects with modifiers in A-chains. Obviously there are other differences between A-chains and A-bar chain. In section 5.2, I show that differences with respect to weak crossover follow from (42). For the different behavior with respect to the licensing of parasitic gaps, I refer the reader to Nissenbaum (1998). Nissenbaum shows that this difference can be derived from syntactic locality differences between different types of movement,

namely whether intermediate adjunction is required. For the differences with respect to locality, I again refer the reader to the respective literature (Rizzi 1990, Chomsky 1995, Takano 1993, 1994, Müller 1993, 1996), which reduces main differences between different movement types to . The open question remaining, is how the difference with respect to intermediate adjunction sites are captured on this approach. However, this problem doesn't directly relate to the issue of the LF-representation of chains, and hence is not crucial for the following.

The position I take above is that the NP-part of an A-chain is subject to the same principles that determine the distribution of modifiers in all chains. These are discussed in the previous two sections; namely, a preference for the surface position which can be overridden by variable binding or ACD. The interaction with variable binding, leads us to expect cases with A-chains where the determiner of the moving DP is separated from the NP-part. It is difficult to determine whether this expectation is fulfilled, as we see in (43) and (44).

In (43), I tried to force reconstruction of the NP-part of the A-moved phrase. The question we're interested in is whether (43) has the LF-representation in (44a), where *one* takes scope in its surface position, but *himself* is interpreted as bound by *everybody*. However, since (43) definitely allows the representation in (44b), where *one* takes scope below *seem* and *everybody*, it is impossible to discern whether there are also readings with wide scope for *one*. The kind of reading we might expect (44a) to have—and it's not so clear what this might be—could also be a specific or wide-scope reading of (44b) (see Fodor and Sag 1982, Reinhart 1997, Kratzer 1995)

(43) [One picture of himself$_j$]$_i$ seemed to everybody$_j$ to t_i be too small.

(44) a. [One] λx seemed to everybody$_j$ to [x, picture of himself$_j$] to be too small.

b. seemed to everybody$_j$ to [one picture of himself$_j$] to be too small.

A better test are interactions between variable binding and Condition C. The paradigm in (45) resembles that in (34) and the judgment is similar, though the contrast seems to be less sharp.[11]

(45) a. *[A picture that he$_k$ showed to Mary$_j$]$_i$ seemed to her$_j$ to have been given t_i to every student$_k$.

b. [A picture of his$_k$ that Mary$_j$ was shown]$_i$ seemed to her$_j$ to have been t_i given to every student$_k$.

c. *[A picture of his$_k$ meeting with Mary$_j$]$_i$ seemed to her$_j$ to t_i have been given to every student$_k$.

In (46), a slight contrast in the predicted direction is found, though again even the better example (46b) is not perfect. Here the reason might be the complexity of the construction, and the fact that it's generally hard to reconstruct in an A-chain if an overt full DP intervenes.

[11] One problem with the examples in the text might be a minor violation of weak crossover. However, in examples like (i), weak crossover is even weaker than it usually is (Burzio 1986:203, Pesetsky 1994:221-223, Pica and Snyder 1994).

(i) ?A picture of his$_j$ mother seemed to have been given to every student$_j$.

(46) a. *[A letter that his$_k$ mother sent to Mary$_j$]$_i$ seemed to her$_j$ to appear to every student$_k$ to be t_i interesting.

b. [A letter of his$_k$ mother that Mary$_j$ had received]$_i$ seemed to her$_j$ to appear to every student$_k$ to be t_i interesting.

Another prediction of the assumption that the A/A-bar difference reduces to (42) is that covert A and A-bar chains should behave alike (except if ACD is involved) because in covert A chains the preference to represent the NP-part in its surface position also predicts it will be represented there. Most cases discussed as covert A-movement in the older literature, namely movement to replace an expletive, don't exhibit any of the semantic effects associated with movement, therefore aren't regarded as covert A-movement at this point. However, there's one case in Modern Greek which seems to disconfirm my prediction. Namely, Alexiadou and Anagnastopoulou (1997) argue that certain cases of clitic doubling in Greek involve covert A-movement, and are hence similar to overt scrambling in languages like German and Japanese. As we see in (47) (Alexiadou and Anagnastopoulou 1997:147), Condition C is obviated by the covert A-movement in (47b), which indicates that there the NP-part of the A-chain doesn't occupy its surface position, but the top position of the A-chain.[12] This could be a problem for the approach taken here, and definitely deserves further study. The fact alone that this construction in Modern

[12]Alexiadou and Anagnastopoulou (1997) also present a contrast similar to (47), but using weak crossover. Since it's known thought that severity of weak crossover is affected by Pesetsky's (1989) D-linking and since clitic doubling seems to bring about a discourse effect similar to D-linking, I consider Alexiadou and Anagnastopoulou's (1997) weak crossover facts unconvincing.

Greek might be the a case of covert A-movement, the only one known to me, is interesting. However, Sabine Iatridou (p.c.) finds the contrast in (47) less clear than Alexiadou and Anagnastopoulou (1997) indicate, and therefore I ignore (47) for now.

(47) a. *O Janis tis$_i$ epestrepse [to vivlio tis Marias$_i$]$_j$ simiomeno
the John she gave back the book of Mary with notes

b. ?O Janis tis$_i$ to$_j$ epestrepse [to vivlio tis Marias$_i$]$_j$ simiomeno
the John she it gave back the book of Mary with notes

In sum, despite the tentative nature of the evidence presented, it seems feasible to fit A-chains into the picture developed for A-bar chains in the previous two sections. I adopt the assumption that the difference between A and A-bar chain can be reduced to (42). I come back to the A/A-bar distinction in section 5.2 with a discussion of weak crossover.

2.4 Relative Clause Internal Traces

The relationship between the head of a relative clause and the relative clause internal trace position is puzzling, as was first pointed out by Munn (1994): As shown in (48a) and with more examples below, no Condition C effect is triggered in this position. On the other hand, as shown by (48b) and more examples below a variable contained in the head can be bound in the relative clause internal trace position.

(48) a. Which is the picture of John$_i$ that he$_i$ likes?

b. Which is the picture of himself$_i$ that everybody$_i$ likes?

This section is concerned with the absence of Condition C effects in relative clauses. More specifically, only restrictive relative clauses are considered. Relative clause formation obviously involves A-bar movement as the locality restrictions movement show. But, as evidenced by (48) and further examples below, the relation between the relative clause head and the relative clause internal trace is different from that between the head of a *wh*-chain and its trace in a question. The conclusion I argue for in section 2.4.1 is that the proposal of Carlson (1977) is essentially correct: There are two possible LF-structures for relative clauses, a matching structure and a raising structure, and the two can be distinguished by means of their interpretation. In section 2.4.2, I show that the two structures have many things in common and I propose a derivation that can generate both the matching and the raising structure. The result of this section not only solves the puzzle (48), but is also important for chapters 3 and 5 where additional evidence for this analysis of relative clauses will be achieved.

As already mentioned, the relation between the external head and the trace inside the relative clause seems to be less direct than with *wh*-movement in questions with respect to Condition C, as pointed out by (Munn 1994, Safir 1998)[13] In examples like (49a) (repeated from (48a)), (50a), and (50a) no Condition C effect if observed even though the R-expression *John* occurs inside the NP-part of the relative clause head. In the *wh*-questions, in the corresponding b)-examples attest coreference is blocked by Condition C.

[13]In some examples, though, a Condition C effect is observed with relative clauses, as I also show below. In particular, I address examples of this kind that are due to Schachter (1973) in footnote 14 below.

(49) a. Which is the picture of John$_i$ that he$_i$ likes?

b. *Which picture of John$_i$ does he$_i$ like?

(50) a. The pictures of Marsden$_i$ which he$_i$ displays prominently are generally the attractive ones.(Safir 1998:(38a))

b. *Which pictures of Marsden$_i$ does he$_i$ display prominently.

(51) a. I have a report on Bob's division he won't like.(Merchant 1998a:fn.1)

b. *Which report on Bob$_i$'s division will he$_i$ not like.

The Condition C evidence seems to show that there is no material of the external head in the relative clause internal position. In other respects though, the relation between the external head and the relative clause internal trace position seems to be just as tight as that in a *wh*-chain. One such case, first observed by (Jackendoff 1968, Schachter 1973:32-33), are examples where the head of the relative clause contains a variable that is bound by an expression inside the relative clause as in (52). While an example like (52b) might not require c-command for the binding, the fact that in (52c) *her* can be interpreted as a bound variable indicates that in this case *her* must be able to occur in the scope of *every professor*.

(52) a. The interest in each other$_i$ that John and Mary$_i$ showed was fleeting. (Schachter 1973:43a)

b. Une photo de lui$_i$ que Jean$_i$ avait donnée à Marie a été
A photo of him that John has given to Mary has been

retrouvée hier. (Vergnaud 1974:256)
found again yesterday

c. The book on her$_i$ desk that every professor$_i$ liked best concerned model theory.

A second case are examples where the head of the relative clause forms an idiom together with other lexical material inside the relative clause (Brame 1968). The noun *headway* in (53a) cannot appear in any other environment than as part of the idiom *make headway*. This suggest the position where *headway* is interpreted in (53a) is the complement position of *make*. (53b) allows both an idiomatic interpretation (*the pictures John made with a camera*) and a non-idiomatic interpretation (*the pictures John grabbed*). For the idiomatic interpretation, the same point could be made, as for (53a).

(53) a. The headway John made proved insufficient.

b. All the pictures John took showed the baby.

Finally, Irene Heim (p.c.) mentions examples where a part of the head of the relative clause seems to take scope below a relative clause internal scope taking element. The preferred interpretation of example (54a) is one that can be paraphrased as "Gina needs so many books for vet school such that no linguist would read that many books". In this paraphrase, *many* takes scope below *need*. Similarly, (54b) prefers an interpretation paraphrasable as: there is a number such that Mary can take n-many

drinks, but she shouldn't even have *n*-many drinks. In this paraphrase as well, the quantifier *n*-many drinks takes scope below the relative clause internal modal *can*.

(54) a. No linguist would read the many books Gina will need for vet school. (need ≫ many)

b. Mary shouldn't even have the few drinks that she can take. (can ≫ few)

In the evidence so far, the relationship between the relative clause internal trace position and the external head is alike to that between head and trace in an A-chain: While there is usually no reconstruction and hence no Condition C effects, binding and scope can force reconstruction. By contrast, the relationship between the relative clause operator and the trace position is exactly like that in a *wh*-movement chain. Not only the locality restrictions and weak crossover point in this direction, but the A-bar nature of the relative clause internal movement can also be shown using Condition C as a test. As Safir (1998) observes, lexical material that is pied-piped by the movement of the relative clause operator behaves exactly like lexical material in *wh*-chains with respect to Condition C. (55) shows the contrast between material pied-piped with the relative clause operator and the external head. The pronoun *he*, which c-commands the relative clause internal trace position, cannot be coreferent with the R-expression *John* that is part of the material pied-piped with the operator in (55a). In (55b), where the R-expression is part of the external head, on the other hand, coreference is possible.

(55) a. *I respect any writer whose depiction of John$_i$ he$_i$'ll object to. (Safir 1998:34a)

b. I respect any depiction of John$_i$ he$_i$'ll object to.

The pairs in (57) and (58) show that a familiar argument/adjunct contrast with respect to the relative clause internal operator movement. No Condition C effect is found in (56a) and (57a), where the the R-expression is part of an adjoined modifier to the constituent moved in the relative clause. (56b) and (57b) show a Condition C effect just like (55a).

(56) a. There's a singer whose picture in John$_i$'s office he$_i$'s very proud of. (Safir 1998:(34b))

b. *There's a singer whose picture of John$_i$'s office he$_i$'s very proud of.

(57) a. Max is a prince John$_i$'s description of whom he$_i$ varies when spies are around. (Safir 1998:(34c))

b. *Max is a prince whose description of John he$_i$ varies when spies are around.

The well-behaved nature of this chain internal to the relative clause, makes the relationship between the external head and the internal trace all the more interesting.

2.4.1 Two LF-Structures for Relative Clauses

What explains the difference between the examples with Condition C and those involving binding, idioms and scope? The explanation, I pursue is based on the idea of

Carlson (1977) that relative clauses are structurally ambiguous at LF. I'll first consider only the LF-structures. Following Carlson (1977), I call the two LF-structures for relative clauses the *matching analysis* (Lees 1960, 1961, Chomsky 1965) and the *raising analysis* (Schachter 1973, Vergnaud 1974). The differences between the two LF-structures and the main prediction of Carlson's ambiguity view—that Condition C reemerges when the raising analysis is forced—are spelled out in this section. In the section 2.4.2, I then look at the matching analysis in more detail and show how the derivation of the two LF-structures could be unified. The details of the semantic procedures that interpret the structures proposed in this section are left to chapter 5.

The two structures are sketched in (58) for an example that forces the matching analysis and in (59) for and example that forces the raising analysis. On the matching structure in (58b) and (59b), the external head and the internal trace are I assume not related via movement. Therefore, the external head is represented in the relative clause external position at LF, but at least not literally in the relative clause internal position. In (58b) and (59b), none of the lexical material of the external head is represented in the relative clause internal position. Notice that to capture the Condition C evidence, the structures in (58b) and (59b) represent only one possibility. In section 2.4.2, I present an argument that the relative clause external head is represented in some sense in the relative clause internal trace position on the matching analysis, and revise the matching structures accordingly. As shown in (58b), the matching structure assumed here is predicted to obviate Condition C. On the other hand, the position of the external head at LF in (59b) rules out the matching structure in case it contains a variable bound by a quantifier inside the

relative clause (unless we assume that the quantifier can move to a position outside of the relative clause). Similarly, the matching analysis is ruled out in the examples (53) with idioms and (54) with scope.

(58) a. the picture of John$_i$ he$_i$ likes

b. the $\overbrace{\text{picture of John}_i}^{\text{head}}$ λx he$_i$ likes $[x]$ *(matching)*

c. *the picture of John λx he$_i$ likes $[x, \overbrace{\text{picture of John}_i}^{\text{head}}]$ *(raising)*

(59) a. the picture of himself$_i$ everybody$_i$ likes

b. *the $\overbrace{\text{picture of himself}_i}^{\text{head}}$ λx everybody likes $[x]$ *(matching)*

c. the λx everybody$_i$ likes $[x, \overbrace{\text{picture of himself}_i}^{\text{head}}]$ *(raising)*

The raising analysis is sketched in (58c) and (59c). Here, I assume that the relation between the internal trace position and the external head is one of movement. Therefore, the R-expression *John* must be represented in the trace position in (58c), just like in the case of *wh*-movement. Hence, (58c) violates Condition C. On the other hand, it is possible to delete all but the lowest copy of the NP-part of this chain as in (59c), and therefore it's possible to completely delete any relative clause external appearance of the relative clause head. This is in fact required for binding in (59c), as well as for idiom interpretation in examples like (53) and for narrow scope as in the examples in (54).

As shown by (58), Condition C can be used to enforce the matching analysis of

a relative clause, while (59) shows that variable binding enforces the raising analysis of a relative clause. By the same logic as that of (59), idiom interpretations and scope can also be used to ensure that the raising analysis is forced. The ambiguity analysis immediately makes one prediction, namely that factors forcing one analysis are incompatible with factors forcing the other, and raises one question, namely which of the two analyses is chosen when none of the factors seen to choose one analysis is at work. I address the question first, and then demonstrate the prediction.

The question is which analysis of a relative clause is chosen if none of the factors mentioned determines the analysis. Part of the answer can be found in the previous work on relatives (Carlson 1977, Heim 1987, Grosu and Landman 1998), who argue that there is a difference in interpretation between the two analysis. For the raising analysis at least four different interpretations should be entertained: an amount reading (Carlson 1977, Heim 1987, Grosu and Landman 1998) as in (60a), a multiple individual reading (Geach 1964, Sharvit 1996a, Sharvit 1996b) as in (60b), a possibility modal reading (Hackl and Nissenbaum 1998) as in (60c), and maybe also a kind reading in (60d) similar to the one Heim (1987:27–33) observes for *what*-questions. The cases of raising relatives noted above can be subsumed under these four types, namely the idiom cases seem to have either an amount reading as argued by Carlson (1977) or a kind reading, the binding cases clearly have the multiple individual reading, and the scope cases all have an amount reading.

(60) a. It will take us the rest of our lives to drink the champagne they spilled that evening. (Heim 1987:(40))

b. The woman every man invited is waiting in the lobby.

c. Sabine has come up with many problems for us to work on. (Hackl and Nissenbaum 1998:(1))

d. The beer that there was for sale was too expensive for John.

The availability of the four different readings seems to be subject to a number of different constraints: for example, the amount reading and the possibility reading are only available with certain determiners, and the multiple individual reading is most easily possible for the argument of a copular construction. However, in general the restrictions on the four readings are only incompletely understood. Unfortunately, the detailed investigation of the semantics of the different readings and the restrictions on them are beyond the scope of the current investigation (see chapter 5). Despite this lack of precise understanding, I think it's safe to proceed with the assumption that the raising structure is only chosen in cases with one the above four interpretations. Specifically, I assume that the raising analysis is only chosen if the NP-part must be deleted in all positions but the relative clause internal trace position. As shown in section 2.4.2, this allows a fairly uniform derivation of both the matching and raising structures. This is obviously required if it NP-part contains a variable that is only bound in this position, or if it's an idiom chunk that can only be interpreted in this position, or if it takes scope below another relative clause internal quantifier. For the cases in the category kind relative, I suggest that they involve some form of binding as well, for example of an event argument.

The remainder of this section demonstrates two predictions of the analysis of relative clauses pursued here. First, consider the prediction mentioned above: The analysis of relative clauses as structurally ambiguous pursued here makes the clear prediction that the factors forcing one analysis are incompatible with those forcing the other analysis. The one factor that, on this account, definitely forces the matching analysis is obviation of Condition C. In the following examples we see that in all the constructions that motivated the raising analysis, Condition C cannot be obviated. First consider variable binding in (61) and (62). In both (61a) and (62a), the pronoun *her* is interpreted as a variable bound by a quantifier in the relative clause.[14] As discussed above, this forces the raising analysis and therefore the Condition C effects observed in (61a) and (62a) between the R-expression *John* in the external head and the pronoun that c-commands the relative clause internal trace position confirm the analysis.[15]

[14]Schachter (1973:32) discusses the examples in (ia) and (iia), where a Condition C effect is observed. These might fall into place here under the assumption that nouns like *opinion* and *portrait* have an implicit subject argument that in (ia) and (iia) is bound from a relative clause internal position (Jackendoff 1972).The examples in (ib) and (iib), which don't show a Condition C effect, don't have this confound.

(i) a. *The (pro$_j$) opinion of John$_i$ that he$_i$ thinks that Mary$_j$ has is unfavorable. (Schachter 1973:(41b))

b. The opinion of John$_i$ that he$_i$ thinks that Mary$_j$ has refute is described in his$_i$ letter to her.

(ii) a. *The (pro$_i$) portrait of John$_i$ that he$_i$ painted is extremely flattering. (Schachter 1973:(42b))

b. The (pro$_j$) portrait of John$_i$ that he$_i$ ordered two years ago was finally delivered.

[15]The contrast between (62a) and (i) is unexpected so far. It indicates that even when the external head must stand in a movement relationship with an intermediate position of the relative clause internal chain, it can nevertheless stand in the more indirect matching relationship with the lowest trace of the same chain. This might indicate that, in fact, not only the relationship of the lowest trace to the external head, but in fact every link of the relative clause internal chain is ambiguous between a raising and matching analysis.

(i) ?A review of John$_i$'s debate with her$_j$ that every senator$_j$ wanted him$_i$ to read landed in the garbage instead.

(61) a. *The letters by John$_j$ to her$_i$ that he$_j$ told every girl$_i$ to burn were published.

b. The letter by him$_j$ to her$_i$ that John$_j$ told every girl$_i$ to burn were published.

(62) a. *A review of John$_i$'s debate with her$_j$ that he$_i$ wanted every senator$_j$ to read landed in the garbage instead.

b. A review of his$_i$'s debate with her$_j$ that John$_i$ wanted every senator$_j$ to read landed in the garbage instead.

The use of idioms is another way to enforce the raising analysis. As Munn (1994) already observes, the prediction that Condition C effects reemerge is confirmed as shown by the pairs in (63) and (64).

(63) a. *the picture of Bill$_i$ that he$_i$ took (Munn 1994:(15c))

b. the picture of himself$_i$ that Bill$_i$ took

(64) a. *The headway on Mary's project she had made pleased the boss. (Nissenbaum, p.c.)

b. The headway on her project Mary had made pleased the boss.

Also, narrow scope of *many* in (65a) and *few* in (65b) seems to cause a Condition C effect in the expected fashion.

(65) a. *The many books for Gina$_i$'s vet school that she$_i$ needs will be expensive. (need $\gg$ many)

b. *The few coins from Bill$_i$'s pocket he$_i$ could spare weren't enough for all the needy. (could $\gg$ few)

In fact, a Condition C effect is found also with other amount readings, as expected. In (66), the amount reading is forced because the relative clause internal trace occurs in a *there*-existential construction (Carlson 1977). This, as proposed by Carlson (1977) and above, forces the raising analysis, and therefore the Condition C effect in (66a) is expected.

(66) a. *It would have taken us all year to read the letters for John$_j$ he$_j$ expected there would be.

b. It would have taken us all year to read the letters for him$_j$ John$_j$ expected there would be.

The second prediction of the analysis of relative clauses is more intricate. It is made by the position of the lexical material of the head at LF in the raising analysis. What we saw just now, is that the lexical material of the head occupies a raising relative clause internal position, and triggers a Condition C effect there. In section 2.1 above, I showed for examples like (13b), which is repeated in (67a), that the head of the relative clause can occupy a position outside of the relative clause and trigger a Condition C effect there. The LF-representation in (67b), which was argued for above, can obviously only hold for matching relatives. I therefore predict that raising relatives will not show the Condition C effect noticed in (67a).

(67) a. *In the end, I did ask him$_i$ to teach the book of David$_i$'s that Irene wanted me to ⟨ask him to teach⟩.

b. *[the book of David$_i$'s [λy Irene wanted me to ask him$_i$ to teach the [y, book of David$_i$'s]]] λx I asked him$_i$ to teach the [x, book of David$_i$'s]

To verify the prediction, we need to look at examples where covert movement of the DP containing the relative clause is forced by ACD, as it is in (67a), but where the head of the relative clause must occupy a relative clause internal position at LF.[16] The examples in (68a) and (69a) demonstrate that the prediction is correct. Both contrast with (68b) and (69b), where there is no ACD to block binding reconstruction of the relative clause. They also contrast with (68c) and (69c), where there is ACD,

[16] Wold (1995:26) shows that sometimes ACD is incompatible with binding reconstruction into the relative clause, as for example in (ib) and in (iib), where the judgment is actually stronger as Danny Fox (p.c.) observes.

(i) a. Sue likes every picture of himself$_i$ that John$_i$ painted.
b. *Sue likes every picture of himself$_i$ that John$_i$ does.

(ii) a. Sue likes every picture of himself$_i$ that every boy$_i$ painted.
b. *Sue likes every picture of himself$_i$ that every boy$_i$ does.

Wold's (1995) effect can be explained by the lack of identity between the elided VP and its antecedent in the LF-representation (iva) of (iib). For the test in the text, however, we can circumvent it, because Danny Fox (p.c.) also shows that Wold's (1995) effect isn't found if there is a relative clause internal trace position outside of the elided VP, where the variable binding can be satisfied as in (iii). The LF-structure of (iii) is shown in (ivb). The examples in the text have an intermediate position just like (iii). Notice, however, that this analysis conflicts with the main proposal of section 2.4.2 in an interesting way.

(iii) Sue likes every picture of himself$_i$ that every boy$_i$ hoped she would.

(iv) a. *[every λx every boy$_i$ $\underbrace{\text{likes } [x, \text{ picture of himself}_i]}_{\text{elided VP}}$] λy Sue $\underbrace{\text{likes } [y]}_{\text{antecedent}}$

b. [every λx every boy$_i$ hoped [x, picture of himself$_i$] λz Sue would $\underbrace{\text{like } [z]}_{\text{elided VP}}$]

λy Sue $\underbrace{\text{liked } [y]}_{\text{antecedent}}$

but the relative clause isn't forced to have a raising analysis.[17]

(68) a. John asked him$_i$ for the pictures of her$_j$ mother meeting Clinton$_i$ every girl$_j$ wanted him to ⟨ask Clinton$_i$ for⟩.

b. *John asked him$_i$ for the pictures of her$_j$ mother meeting Clinton$_i$ every girl$_j$ had published.

c. *John asked him$_i$ for the picture of the woman meeting Clinton$_i$ every girl$_j$ wanted him to ⟨ask Clinton$_i$ for⟩.

(69) a. The host introduced him$_i$ to the writers of her$_j$ replies to Casanova$_i$ every girl$_j$ refused to ⟨introduce him$_i$ to⟩.

b. *The host introduced him$_i$ to the writers of her$_j$ replies to Casanova$_i$ every girl$_j$ had hired.

c. *The host introduced him$_i$ to the writers of the letters to Casanova$_i$ every girl$_j$ refused to ⟨introduce him$_i$ to⟩.

The LF-representation I propose for the example (68a) is sketched in (70), where irrelevant details about the lowest relative internal trace position are omitted. Since the head of the raising relative clause occupies a relative clause internal position, it escapes Condition C for the same reason that material inside a raising relative was found to do so earlier. Namely, ACD forces deletion of the copy of the relative clause

[17]The judgement in these cases is made easier, if they're put in the context of a little story. For (a), for example, the story might say that John is investigating girls whose mothers had affairs with Clinton. It's known that Clinton maintains photographic records of his affairs, and the girls each would like to see some of the pictures of their mothers with Clinton from his archives, but are afraid to ask him. Therefore, John asks Clinton for the pictures.

in the QR-trace position, and therefore the R-expression occurs only in a position in the head of the QR-chain in (70).

(70) $\Big[$the [λx every girl$_j$ wanted him [x, pictures of her$_j$ mother meeting Clinton$_i$] to $\underbrace{\text{ask Clinton}_i \text{ for } [x]}_{\text{elided VP}}\Big]$ λy John $\underbrace{\text{asked him}_i \text{ for } [y]}_{\text{antecedent}}$

The case in (71) makes the same point as (68) and (69), but the raising analysis of the relative clause is forced by enforcing an amount reading of the relative clause. Again, ACD-resolution in (71a) obviates Condition C even when the R-expression occurs outside of the relative clause on the surface.

(71) a. The company will send her to any fan clubs of Mary there are requesting it ⟨that the company send Mary to them⟩[18].

b. *The company will send her to any fan clubs of Mary there are.

2.4.2 The Internal Head of Matching Relatives

In this section, I look at the matching analysis of relative clauses in more detail. The LF-representation of a matching relative clause assumed in (58b) above is repeated in (72b). In this section, I argue that the representation is instead that in (72c), where the internal position contains an elided NP the antecedent of which is the external head. The argument in this section is based on data from Safir (1998);

[18]Obviously the *it* in this case is not the usual VP-ellipsis of the textbook cases. However, as David Pesetsky (p.c.) pointed out to me, such antecedent contained anaphora are expected to and do indeed behave exactly like ACD.

an additional argument for (72c) is given in section 3.1. Recall that also that the discussion of example (13) lead us to propose the structure (17), which is essentially like (72c). The proposal (72c) takes the term *matching* seriously: At some point of the derivation the internal head must be (almost) identical to the external head. This raises the question at what point of the derivation *matching* must be satisfied. I argue that the point of the derivation where this matching requirement must be satisfied is LF. As I show below, this assumption also allows us to (almost) reduce the raising analysis to a special case of the matching analysis.

(72) a. the picture of John$_i$ he$_i$ likes

b. the $\overbrace{\text{picture of John}_i}^{\text{head}}$ λx he$_i$ likes $[x]$

c. the $\overbrace{\text{picture of John}_i}^{\text{overt NP}}$ λx he$_i$ likes $[x, \overbrace{\text{picture of him}_i}^{\text{elided NP}}]$

Consider first what the absence of Condition C effects tells us about the matching analysis. As we saw already above (examples (55) to (57)), the trace position inside the relative clause must at least contain a representation of the NP-part of the material that is pied-piped with the relative clause operator. The question here is to what extent the external head is represented in the trace position in a matching relative clause. The fact that the external head triggers no Condition C effects inside a matching relative could be explained by various degrees of indirect representation, for example the relationship between the pronoun and its antecedent in (73a) is such that no Condition C effect is obtained, or that between the elided VP and its

antecedent in (73b).

(73) a. John drew a picture of Mary$_i$, but she$_i$ didn't like it ⟨the picture of Mary$_i$⟩.

b. Mary loves John$_i$ and he$_i$ thinks that Sally does ⟨love John$_i$⟩, too. (Fiengo and May 1994:220)

To explain the latter observation, Fiengo and May (1994) propose that the identity relationship between the elided VP and its antecedent is satified, even when an R-expression in the antecedent corresponds to a coreferent pronoun in the elided VP (see also sections 3.2 and 3.3 on the identity relationship). Fiengo and May (1994) introduce the term *Vehicle Change* for such cases where exact identity of syntactic form is violated. The structure (74) is the LF-representation Fiengo and May (1994) propose for (73b). I adopt Fiengo and May's (1994) proposal, as is already indicated in the structure in (72c).

(74) Mary $\underbrace{\text{love John}_i}_{\text{antecedent}}$ and he$_i$ thinks that Sally $\underbrace{\text{loves him}_i}_{\text{elided VP}}$

The argument I present for the claim that the internal trace contains an elided representation of the external head is based on the observation of Safir (1998:(35)) in (75). In (75a), it's impossible for the quantifier *anyone* in the external head to bind the pronoun *he* in the relative clause. In (75b), on the other hand, binding of *him* by *anyone* is possible. As the similar contrast in (76) confirms, the relevant difference is whether the pronoun in the relative clause is c-commanded by the relative clause internal trace.

(75) a. *Pictures of anyone$_i$ which he$_i$/his$_i$ mother displays prominently are likely to be attractive ones.

b. Picture of anyone$_i$ that put him$_i$/his$_i$ mother in a good light are likely to be attractive ones. (Safir 1998:(35))

(76) a. *Mary exhibited the picture of every boy$_i$ that he$_i$/his$_i$ sister brought.

b. Mary exhibited the picture of every boy$_i$ that was brought by him$_i$/his$_i$ sister.

Since the quantifier in all four examples in (75) and (76), doesn't c-command the pronoun it binds in the surface structure, all examples might be expected to be weak crossover violations. But, at least since Gabbay and Moravscik (1974), Hintikka (1974), Reinhart (1976), and May (1977:61-124), it's known that DP-internal quantifiers can bind a pronoun outside quite easily, as long as they take scope over it. This is illustrated in (77). In fact, the status of (75b) and (76b) seems comparable to the examples in (77). On an analysis, where the external head is not represented at all in the relative clause internal position, however, (75a) and (76a) would incorrectly be expected to be as good as the examples in (77), as well.

(77) a. One picture of everyone$_i$ is displayed by him$_i$ prominently.

b. Somebody from every city$_i$ despises it$_i$. (May 1985:68)

Safir (1998) proposes that the bad examples of the contrasts in (75) and (76) should receive the same explanation as the badness of the examples in (78), which

display strong and weak crossover.

(78) a. *He$_i$ is displaying a picture of everyone$_i$.

b. *Which picture of everyone$_i$ is he$_i$ displaying?

c. ??His mother is displaying a picture of everyone$_i$.

d. ??Which picture of everyone$_i$ is his$_i$ mother displaying.

For a raising relative, of course, any account of (78) carries over to the cases under discussion. However, there's no evidence that Safir's examples must receive a raising analysis. Moreover, the examples in (79) where the raising analysis is ruled out by Condition C shows the same contrast as Safir's example.

(79) a. *The Times will generally publish pictures of any woman$_i$ visiting Clinton$_j$ that he$_j$ told her$_i$ about.

b. The Times will generally publish picture of any woman$_i$ visiting Clinton$_j$ that he$_j$ thinks will offend her$_i$.

I believe that any account of Safir's discovery on the matching analysis has to propose a representation of the external head in the internal position, but not one related by movement to the external head. The particular version of this I assume here is that the internal head contains a phonologically deleted version of the external head. Implicit in this proposal is that the external head and the internal head must match at the level of LF, and there only, since this is generally the case for phonological deletion, for example in ACD. If we furthermore assume that the

quantifier in the external head undergoes quantifier raising to position outside of the DP, the LF-representation of (75) can be sketched as in (80), where I assume that the *anyone* leaves the NP-part *one* in its trace position. In (80), the copy of [x, one] in the relative clause internal trace is c-commanded by the pronoun he_x. Therefore, (80) is predicted to be a case of strong crossover.

(80) *anyone λx $\Big[\underbrace{\text{pictures of } [x, \text{ one}]}_{\text{external head}}$ [which $\underbrace{\text{picture of } [x, \text{ one}]}_{\text{internal head}}$] λy he_x displays prominently $\underbrace{[y, \text{ pictures of } [x, \text{ one}]]}_{\text{internal trace}}\Big]$ are likely to be attractive ones.

One more revision of the structure in (80) is required: Though the analysis of (75) in (80) successfully predicts a strong crossover violation, it's not generally the case that an elided correspondent of a trace in the antecedent shows strong crossover effects, as shown by (81a) from Fiengo and May (1994:279). Assuming exact identity of syntactic form between the elided VP and its antecedent, (81b) is the LF-representation of (81a). To resolve ACD, the *to*-object of (81) must undergo quantifier raising to a position outside of the VP. But, since the direct object, *every guy*, binds a varialbe in the relative clause adjoined to the *to*-object, it must undergo quantifier raising as well to a position where it c-commands the raised *to*-object. But, then (81b) violates the strong crossover condition: The antecedent VP contains a trace of quantifier raising in the direct object position, and therefore the elided VP in the relative clause does as well if we assume identity of syntactic form. This trace in the elided VP, however, is c-commanded by a coreferent pronoun he_x. Therefore, (81b) violates strong crossover.

(81) a. Mary introduced every guy to every woman he wanted her to ⟨introduce him to⟩

b. *[every guy] λx [every woman λz he$_x$ wanted her to $\underbrace{\text{introduce } [x, \text{ guy}] \text{ to } [z, \text{ woman}]}_{\text{elided VP}}$]

λy Mary $\underbrace{\text{introduced } [x, \text{ guy}] \text{ to } [y, \text{ woman}]}_{\text{antecedent}}$

But, the obviation of strong crossover in (81a) is not surprising on a view where strong crossover is reduced to Condition C, since Condition C violations disappear under ellipsis as was shown by (73). Extending their notion of vehicle change, Fiengo and May (1994) propose that a trace in the antecedent in the antecedent of VP-ellipsis, just like an R-expression, can correspond to a pronoun in the elided material (see also Merchant 1998b). Adopting this assumption, the LF representation of (81) is given in (82), where the direct object in the ACD-relative clause is a pronoun. Since pronouns are not subject to strong crossover, (82) doesn't violate strong crossover.

(82) [every guy] λx [every woman λz he$_x$ wanted her to $\underbrace{\text{introduce him}_x \text{ to } [z, \text{ woman}]}_{\text{elided VP}}$]

λy Mary $\underbrace{\text{introduced } [x, \text{ guy}] \text{ to } [y, \text{ woman}]}_{\text{antecedent}}$

Fiengo and May (1994) point out the contrast between (83a) and (81a), which lends strong support to their account of (81a). It seems that in (83a) a strong crossover effect is maintained, though the potentially violating trace is also part of an elided VP. As Fiengo and May (1994) argue, the apparent strong crossover effect in (83a) should be analyzed as a Condition B violation. Assuming that the correspondent of the trace in the elided material is a pronoun, the LF-representation of (83a) in (83b)

violates Condition B since the this pronoun him_x is in the local domain of another pronoun him_x. In (82), on the other hand, the pronoun him_x that corresponds to the trace is far enough away from the other pronoun, such that Condition B is satisfied in (82).[19]

(83) a. *Mary introduced every guy to every woman she wanted him to ⟨introduce him to⟩

b. *[every guy] λx [every woman she wanted him$_x$ to introduce him$_x$ to] λy Mary introduced [x, guy] [y, woman]

Condition B suffices to rule out Safir's example (75), which is repeated in (84a). (84b) shows the LF-representation of (84a) assuming that the trace of quantifier raising of *anyone* in the elided occurence of the head is changed to a pronoun. This pronoun is expected to violate Condition B, just like the pronoun in (85) which is part of the fronted *wh*-phrase.

(84) a. *Pictures of anyone$_i$ which he$_i$/his$_i$ mother displays prominently are likely to be attractive ones.

[19]In other examples of strong crossover under ellipsis, like (ia) and (iia) it seems the effect remains even when the distance of the trace and the c-commanding pronoun is big enough to satisfy Condition B. The illformedness of these examples is explained by the parallel dependencies condition (45) on 122 (see also Fox (1998c)).

(i) a. *The man$_i$ who$_i$ Mary said that she likes and who$_i$ he$_i$ did ⟨say that he$_i$ likes⟩ too. (Ristad 1990:144)

b. *The man$_i$ who$_i$ Mary said that Sue likes and who$_i$ he$_i$ did ⟨say that Sue likes⟩ too.

b. *anyone λx $\Big[\underbrace{\text{pictures of } [x, \text{ one}]}_{\text{external head}}$ [which $\underbrace{\text{picture of him}_x}_{\text{internal head}}$] λy he$_x$ displays prominently $\underbrace{[y, \text{ pictures of him}_x]}_{\text{internal trace}}\Big]$ are likely to be attractive ones.

(85) *Which picture of him$_i$ does everyone$_i$ display prominently.

The explanation of Safir's observation (75) as a violation of Condition B, makes new predictions concerning the locality of the the effect. The prediction is that the effect should be obviated if more material intervenes between the trace in the relative clause internal head and the pronoun that triggers the Condition B violation. That this prediction is correct is evidenced by the contrasts in (86) and (87). (86a) (repeated from (76a)) and (87a) display the same degree of illformedness as Safir's observation. (86b) and (87b), where the quantifier is embedded more deeply in the relative clause head, however, are markedly better.

(86) a. *Mary exhibited the picture of every boy$_i$ that he$_i$ bought.

b. Mary exhibited the picture of every boy$_i$'s mother that he$_i$ bought.

(87) a. *John bought a picture of every girl$_i$ that she$_i$ chose.

b. John bought a picture of every girl$_i$'s father that she$_i$ chose.

The improvement exemplified by (86b) and (87b) is predicted by my account of Safir's observation. Because the quantifier is more deeply embedded, the pronoun corresponding to the trace inside the relative clause internal head is not in the local domain of its antecedent, and therefore doesn't violate Condition B. The contrast

between (86a) and (86b) is hence analogous to that between (88a) and (88b).

(88) a. *Which picture of him$_i$ did every boy$_i$ buy.

b. Which picture of his$_i$ mother did every boy$_i$ buy.

As second way to increase the distance between the two positions that give rise to a Condition B violation in Safir's example, is by making the relative clause longer in the way shown in (89b) and (90b). In (89b) and (90b), however, if any, only a very small improvement is found as compared to (89a) and (90a).

(89) a. *Mary exhibited the picture of every boy$_i$ that he$_i$ bought.

b.*?Mary exhibited the picture of every boy$_i$ that he$_i$ thought John bought.

(90) a. *John bought a picture of every girl$_i$ that she$_i$ chose.

b.*?John bought a picture of every girl$_i$ that she$_i$ thought he would choose.

But, the status of (89b) and (90b) should be measured against an example like (91b), where Condition B in an intermediate position of a chain is at issue. Since (91b) seems to be not fully grammatical, Conditon B seems to apply in intermediate positions of a chain. But, if this is the case, (89b) and (90b) are expected to ungrammatical as well.

(91) a. Which picture of herself$_i$ does every girl$_i$ believe Bob likes?

b.*?Which picture of her$_i$ does every girl$_i$ believe Bob likes?

The correlation between Condition B effects and Safir's observation confirms the account of the latter proposed above. Therefore the strong crossover case of (75) supports a version of the matching proposal where the internal trace is identical to the external head modulo vehicle change. Since the weak crossover case of (75) relates to the unsolved problem of why no weak crossover is found in cases like (77), which I also have no solution for, I leave this matter open. Together with the arguments in section 3.1 and that surrounding the structure (17), I believe to have given conclusive evidence for a representation of the external head internal to a matching relative clause.

The next argument concerns the question where matching of the internal and external head applies in matching relatives. The belief expressed above, that the relationship between the two NPs is that of an elided NP and its antecedent presupposes that matching is verified at LF. And the evidence given above for the effects of vehicle change in matching relatives already lend strong support to this conclusion. The following argument provides further evidence that matching applies at LF. Consider (92), which is repeated from (34) above. I argued that, at LF, the NP-part *paper of* his_k of the *wh*-phrase is represented in the trace position t_i, while the relative clause *that* $Mary_j$ *was given* is represented at LF in its surface position. The LF-structure of (92) is shown (93).

(92) [Which paper of his_k that Mary_j was given]$_i$ did she_j tell every student_k to revise t_i?

(93) $\Big[$Which [λz Mary$_j$ was given [z]]$\Big]$ λx did she$_j$ tell every student$_i$ to revise [x, paper of his$_i$]?

Notice that the LF-structure in (93) satisfies matching since both, the internal and the external head, are empty. If, however, matching was applying before the higher copy of *paper of his*$_k$ is deleted, the relative clause in (92) would be expected to contain a copy of it, and specifically the bound variable pronoun *his*$_k$. This pronoun, however, would not be in the scope of its binder in the LF-representation of (92). Hence, matching must apply in (92) after the overt copy of *paper of his* has been deleted.

At this point, the raising analysis proposed above can be analyzed as special case of the matching analysis with one remaining phonological stipulation. Specifically, the argument that the matching applies at LF and that matching is satisfied in case both NP-parts are empty as in (92) suggests that this is the case in the raising analysis as well. Consider again the matching and raising structures in (94) (repeated with modifications from (59)). Empty NPs are indicated in (94c) by empty brackets []. Both the external head position and the complement of the relative clause operator could be occupied by an empty NP, and therefore the structure (94c) satisfies matching.

(94) a. the picture everybody likes

b. the $\overbrace{\text{picture}}^{\text{head}}$ which $\overbrace{\text{picture}}^{\text{elided NP}}$ λx everybody likes [x, picture] *(matching)*

c. the [] which [] λx everybody likes [x, $\overbrace{\text{picture}}^{\text{head}}$] *(raising)*

Since the relative clause operator is related to the trace position by movement, the structure of (94c) before LF-deletion applies must be that in (95). This structure, does however, not directly reflect the facts of English pronounciation: in raising relatives the head is pronounced in front of the relative clause operator, just like in matching relatives. For matching relatives, I have already assumed a pronounciation rule that bans pronounciation of the internal head. To get the pronounciation of (95) right, I suggest that just in case the external head is empty, the internal head is actually pronounced, namely in the position of the external head. This is clearly a stipulation, but at this point it seems to be the best I can do.

(95) the [] which picture λx everybody likes [x, picture]

The assumption that raising relatives are a special case of matching relatives provides a straightforward explanation for the restricted occurence of raising relatives noted in the discussion of (60). Recall from section 2.2 that the lexical material of the top copy of a *wh*-chain can only be deleted if this is required for the interpretation of a bound variable, that is not bound in this high position. If this generalization applies to relative clauses as well (and also encompasses the cases of narrow scope in (54)), it predicts that the copy of the internal head in the position of the relative clause operator can only be deleted if it contains a variable that is bound internal to the relative clause. Furthermore, if the copy of the internal in the operator position

doesn't delete, matching requires the external head to be non-empty as well. Therefore, the external head can only be empty, if the internal head contains a variable that's bound internal to the relative clause. This explains that the raising analysis of relative clauses is restricted to cases with a 'special' interpretation in (60)—only the special interpretation requires that the internal head be deleted in the operator position.

To conclude this section, let me summarize the main points concerning relative clauses. Based on the contrasting behavior with respect to Condition C and other tests for Binding Reconstruction, I concluded that there are two possible LF-structures for relative clauses: a raising and a matching structure. Of these, I claimed the raising analysis to always be associated with a 'special interpretation' as exemplified by (60), whereas the matching analysis I claimed to be the default. Furthermore, I concluded that the relative clause internal trace of a raising relative forms a chain with the external head, whereas on the matching analysis the internal head consists of its own lexical material, but must be phonologically deleted under identity with the external head.

2.5 Summary

In this chapter, I argued for four main generalizations about which material of a moved DP seems to enter binding theory in the trace position. The discussion above has shown that it's useful to distinguish three types of parts of a moved DP, the determiner D, the NP-part, which is the lowest NP-projection (excluding all adjuncts)

of the complement of D, and relative clauses and other modifiers adjoined to the NP-part. For the concise statement of the generalizations, I use the term *segment* to refer to either the NP-part or any modifier of a DP. The terminilogy is exemplified in (96).

(96) $\underbrace{\text{which}}_{\text{Det.}}\ \underbrace{\underbrace{\text{argument of John's}}_{\text{NP-part}}\ \underbrace{\text{that Mary had criticized}}_{\text{modifier}}}_{\text{segments}}$

The generalizations can be stated as in (97), as conditions governing when deletion applies to a copy of the NP-part or a Modifier in a chain. The way the generalizations are stated in (97) reflects a hierarchy between them with (97a) being the highest ranked. Generalizations lower in rank, are only fulfilled up to the an extent such that the higher ranked generalizations are fulfilled.

(97) a. **Recoverability:** At least one copy of every segment of the restrictor must remain represented.

b. **Binding:** Any occurence of a segment that contains a bound pronoun that isn't c-commanded by its antecedent must be deleted.

c. **A-bar:** The lowest position of an A-bar chain must contain a copy of the NP-part.

d. **ACD:** If material inside a modifier is anaphorically related to the constituent surrounding an occurence of this modifier, this occurence of this modifier must be deleted.

e. **Lebeaux's Generalization:** Copies of a segment in positions lower than the copy that is pronounced may be deleted (in a particular order).

f. **Economy of Deletion:** Segments must not be deleted.

Of the six generalizations, (97b) to (97e) have been argued for in detail above, while (97a) and (97f) have been more or less presupposed as background assumptions. Both (97a) and (97f) play an important role in the account. (97f) is, for example, responsible for the fact that quantifier raising doesn't obviate Condition C in examples like (98a) (repeated from (9)), where ACD isn't involved. If it was possible to delete the lower copy of the relative clause modifier in the LF-representation in (98b), Condition C should be obviated by quantifier raising in (98a), contrary to fact.

(98) a. *Someone introduced him$_i$ to everyone John$_i$ wanted you to dance with.

b. *$\Big[$everyone [λy J.$_i$ wanted you to dance with [y]]$\Big]$

λx someone introduced him$_i$ to [x, [λy J.$_i$ wanted you to dance with [y]]]

The Recoverability constraint (97a) is required for examples like (99). Since the every copy of the modifier *who knows her$_i$* in (99) will contain a bound pronoun that isn't c-commanded by its antecedent, (97b) would force deletion of all copies of this modifier. This would incorrectly predict that (99) should be grammatical, namely with the same interpretation as the sentence *The boy thinks that every girl is singing.* (97a) blocks deletion of all copies of the relative clause in the LF-representation of (99) and therefore (99) is correctly predicted to be ungrammatical.

(99) *The boy who knows her$_i$ thinks that every girl$_i$ is singing.

Of the four main generalizations (99b) to (97e), I assume (99b) and (99c) throughout in the form stated. Of course, it would be desirable to derive them from other principles, but at this point such a step seems premature. (97e), I assume involves seemingly countercyclic adjunction because of the discussion of example (37) above. (99d), finally, has curious nature since it seems to involve look-ahead to interpretation in form of the licensing of ACD. Alternatively, (99d) might be a deletion rule that always applies in the case of an VP-deletion dependency where the elided VP is contained in the antecedent it depends. Then, (99d) would require no lookahead, but the dependency of the two VPs would need to be formally represented. At the end of section 3.2, I present one argument for the latter view of (99d).

Chapter 3

Identity of Traces

Certain constructions impose an identity (or parallelism) requirement on two constituents. If these constituents contain traces, we can ask the question under what conditions two traces are identical. This chapter argues that two traces are identical in the relevant sense if the lexical content represented in the trace positions is the same. Therefore, this result provides independent support for the claim of the previous chapter that the lexical content of a moved phrase is partially represented at LF in the bottom position of a chain. More precisely, I show a perfect correspondence; namely, the same parts of the moved phrase are represented in the trace position for the concerns of Binding Theory (previous chapter) and the concerns of the Identity Condition (this chapter). In section 3.2, I argue for the stronger claim that the lexical material in the trace position is not only represented there, but interpreted in the trace position.

A major part of this chapter concerns the analysis of a restriction on ACD that was first studied in detail by Kennedy (1994). The restriction is demonstrated

in (1a), where ACD is blocked. The contrast between (1a) with ellipsis and (1b) without ellipsis shows that the ungrammaticality of (1a) is due to a restriction on VP-ellipsis.

(1) a. *Polly visited every town in every country Eric did ⟨visit⟩.

b. Polly visited every town in every country Eric visited

c. Polly visited every town Eric did.

Kennedy's puzzle is then to explain why ACD is possible in (1c), but not in (1a). Descriptively, the difference between (1a) and (1c) is the following: In (1c), the ACD-relative clause is attached directly to the NP that will undergo quantifier movement for the resolution of ACD. In (1a), on the other hand, the ACD-relative clause is attached to an argument of this NP. In fact, it's marginally possible in (1a) (and much easier with an overt complementizer *that* in the relative clause) to attach the relative clause to the higher noun *town*, in which case ACD is grammatical.

Since I keep referring back to the same example for most of this chapter, it's more convenient to talk about the contrast in (2) instead of (1). For (1), the natural reading is one where *every country* takes scope over *every town*. But, this scope shift is irrelevant for the discussion, and would make the LF-representations more complex than needed. Therefore, I talk about example (2a), where the scope shift is not needed. In (2a) the judgment is more subtle than in (1a), since (2a) is grammatical on a reading of the elided VP clause as *visited every town in* t, which is more marginal in (1a). The fact to explain though is that both (1a) and (2a) don't have the reading

of the elided VP as only *visited*.

(2) a. *Polly visited every town in a country Eric did ⟨visit⟩.

b. Polly visited every town Eric did ⟨visit⟩.

Assuming that traces are interpreted as variables, (3) shows the LF representations for (2a) in (3a) and for (2b) in (3b). In both cases, the quantifier *every town* has moved to resolve ACD. As a result of this movement, the elided VP and the antecedent are identical in both (3a) and (3b). As I argue now, assuming the representations in (3) would make it impossible to account for (2) in a principled way—not surprisingly so, since the VPs in (3a) and (3b) are the same.

(3) a. *[every town, in a country Op_y Eric $\overbrace{\text{visited } [y]}^{\text{elided VP}}$] λx Polly $\overbrace{\text{visited } [x]}^{\text{antecedent}}$

b. [every town, Op_y Eric $\overbrace{\text{visited } [y]}^{\text{elided VP}}$] λx Polly $\overbrace{\text{visited } [x]}^{\text{antecedent}}$

Kennedy (1994) showed that the explanation of (2) must be a constraint on ellipsis, as mentioned above. The only constraint on ellipsis usually assumed is an identity (or parallelism) condition that the elided VP and its antecedent must satisfy, where I for now assume an intuitive concept of identity that is sharpened in sections 3.2 and 3.3 (In the end, the condition I assume is very similar to that of Rooth (1992b).) The only place where a difference could be made between the two VPs in (3b) are the traces. Therefore, I assume that (3) is evidence for a condition on the identity of traces that distinguishes (2a) from (2b). This assumption underlies all approaches to

Kennedy's puzzle I know of, namely those of Kennedy (1994) and Heim (1997a) and the one developed.

The question where the approaches disagree is: What makes the traces in (3a) different, whereas those in (3b) are identical? Sag (1976:66,103) first suggested that traces are only identical for the purposes of VP-ellipsis if their binders are the same. Sag also develops a particular way to implement this suggestion, namely via two restrictions that apply to the indices conventionally used to mark relations of dependence: First, different dependencies, even when they don't overlap, must use different indices, and, second, an elided VP is only identical to its antecedent if the indices on all unbound traces (and other variables) are the same. If relative clause internal traces are viewed as bound by the DP the relative clause is attached to, these considerations yield the LF-representations sketched in (4), where crucially the indices of the traces in (4b) are identical, but not in (4a). Both Kennedy (1994) and Heim (1997a) develop Sag's idea and apply it to cases like (2). I call Sag's approach as well as its descendants the *index identity approach*.

(4) a. *[every town, in a country Op_y Eric $\overbrace{\text{visited } [y]}^{\text{elided VP}}$] λx Polly $\overbrace{\text{visited } [x]}^{\neq \text{ antecedent}}$

b. [every town, Op_x Eric $\overbrace{\text{visited } [x]}^{\text{elided VP}}$] λx Polly $\overbrace{\text{visited } [x]}^{\text{antecedent}}$

The reason I think Sag's index identity is not the right approach to Kennedy's puzzle (2) are contrasts like (5). Both examples in (5) have the same structure. The only difference between (5a) and (5b) is the head of the relative clause. Since this differ-

ence isn't expected to affect the indexation possibilities, the index identity approach predicts (5a) and (5b) to have the same status; namely, both should be ungrammatical. This prediction is wrong: (5a) is clearly better than (5b). Sections 3.1 and 3.2 contain numerous contrasts like (5) which make sure that (5) is representative of a real generalization. The failure of the index identity approach to account for this generalization leads me to reject it and to pursue an alternative approach to Kennedy's puzzle. It should say, though, that while I reject index identity as an approach to Kennedy's puzzle, this doesn't justify provide an argument against the index identity condition *per se*, but only against an account of Kennedy's puzzle (2) based on the index identity condition. In fact, I present empirical support for the index identity requirement in section 4.1 and discuss in section 4.2 which assumption of the index identity approach should be given up. Since I present Heim's (1997a) version of the index identity approach there in more detail and my approach is based on quite different assumptions, I don't discuss it any further in this chapter.

(5) a. Polly visited every town that's near the one Eric did ⟨visit⟩.

b. *Polly visited every town that's near the lake Eric did ⟨visit⟩.

The contrast in (5) shows that lexical properties of the antecedents of the traces affect the acceptability of examples with the structure of Kennedy's puzzle. My approach to Kennedy's puzzle is inspired by the idea of Chomsky (1993) that the trace positions contain copies of the lexical material of their antecedents, which was also discussed in the previous chapter. Hence, I call this approach the *Copy Identity*

Approach. Consider the sketched representations for (2) in (6). In (6), I repeated the head noun of the antecedent in the trace positions. In the sketch (6a) for the bad example, (2a), the antecedent is different from the elided VP. In (6b), on the other hand, the antecedent and the elided VP are identical.

(6) a. *[every town, in a country Op Eric $\overbrace{\text{visited country}}^{\text{elided VP}}$] Polly $\overbrace{\text{visited town}}^{\neq \text{ antecedent}}$

b. [every town, Op Eric $\overbrace{\text{visited town}}^{\text{elided VP}}$] λx Polly $\overbrace{\text{visited town}}^{\text{antecedent}}$

I claim that the lexical material in the trace positions in the way captured by (6) is the right explanation of Kennedy's puzzle. This copy identity approach is developed in section 3.1. It is shown, in particular, that the copy identity approach directly predicts the contrast in (5) and similar such contrast. Another point, section 3.1 discusses that relationship of the copy identity approach to the Condition C evidence discussed in chapter 2. Notice that in (6) only parts of the moved phrases are represented in the trace positions. Section 3.1 shows that the copy identity approach and Condition C converge on the same conclusion as to which parts of a moved phrase are represented in the trace position.

Section 3.2 makes a new argument concerning the lexical material represented in the trace position, that goes beyond what could be tested using Condition C in chapter 2. It argues that the lexical content of the trace position is not only formally represented there, but contributes to interpretation in the trace position. I argue for this based on the observation that the acceptability of examples that test the identity

of traces depends on the semantic relationship of the lexical content of the traces, as well as on general grounds.

Section 3.3 considers facts like (7) where no effect of the copy identity is observed, thought the elided VP and its antecedent contain traces with different lexical content. I show that two mechanisms can circumvent the effect of the copy identity requirement: focus percolation into the trace position and a kind of sloppy reading. The former, I argue in section 3.3.2, applies to example (7a), while the latter applies to (7b) as shown in section 3.3.3.

(7) a. I know which cities Mary visited, but I have no idea which lakes she did ⟨visit⟩. (= (71a))

b. The cities Mary visited are near the lakes Bill did ⟨visit⟩. (= (71b))

3.1 A Copy Identity Account of Kennedy's Puzzle

This section begins to develop an account of Kennedy's observation (2) based on the view that lexical material of the head of a chain is partially represented in the bottom position of the chain. One of the conclusions of the previous chapter was that A-bar traces of a DP always contain the lexical content of the NP-part of the moved DP. (Recall, that I defined *NP-part* as the NP that is the sister of the D-head of the DP minus all adjoined modifiers). In ACD-constructions, specifically, section 2.1 argued based on Condition C that exactly the NP-part is represented in the trace position, whereas all the quantifier and the ACD-relative are represented only in the top position of the QR-chain. Furthermore, the analysis of matching relatives

of section 2.4 argued that the relative clause contains an unpronounced copy of the external head of the relative clause. If this is true of the ungrammatical example from (2), repeated as (8a), the LF-representation must be (8b), which essentially the same as (6a). The elided VP and its antecedent differ in (8b) with respect to the lexical material that appears in the trace position. My proposal is that this difference blocks VP-ellipsis in (8a).

(8) a. *Polly visited every town in a [country Eric did ⟨visit⟩].

b. *[every town in a [country, Op_y Eric $\underbrace{\text{visited } [y, \text{ country}]}_{\text{elided VP}}$]]

λx [Polly $\underbrace{\text{visited } [x, \text{ town}]}_{\neq \text{ antecedent}}$]

Compare (8) with the grammatical example of (2), which is repeated in (9a): The lexical content of the two traces, the one in the elided VP and the one in the antecedent, is identical, namely [x, town]. This is shown by the LF-representation in (9b).

(9) a. Polly visited every town Eric did.

b. every [Op_y Eric $\underbrace{\text{visited } [y, \text{ town}]}_{\text{elided VP}}$] [$\lambda x$ Polly $\underbrace{\text{visited } [x, \text{ town}]}_{\text{antecedent}}$]

In both (8b) and (9b) the names of the variables, x and y, inside the traces differ between antecedent and elided VP. For now, assume that the names of variables are ignored by the identity condition. In section 4.1, I argue contrary that this assumption is in general wrong, but in section 4.2, I argue that in examples like (8b) and (9b) the identity requirement for variable names can be circumvented. For the moment,

it's easiest to assume that variable names generally don't matter.

Now, consider one prediction of the copy identity approach already hinted at in the introduction with (5). This prediction is that if the antecedents of two traces have the same NP-parts, the traces should be considered identical even if the two operators binding the traces are different. Consequently, ACD should be possible. The contrast in (10) shows that this prediction is correct. (10b) is basically the same as Kennedy's example (8a). In (10a), however, the NP to which the relative clause is attached to and the NP-part of the object quantifier are lexically identical. If the second occurrence of *town* is destressed in (10a), the example is fully acceptable.[1]

(10) a. John visited every town near a town Mary did ⟨visit⟩.

b. *John visited every town near a lake Mary did ⟨visit⟩.

The LF-representation for (10a) is shown in (11). The trace-positions in the elided VP in the relative clause and the trace of quantifier raising in the antecedent both have *town* as its lexical content, and therefore the elided VP and its antecedent mean

[1]Some English speakers don't find the improvement in (10a) very strong, but everybody I consulted with found a strong contrast in the examples with *one*-anaphora below. I assume that speakers who find (10a) unacceptable differ from those who do in whether they find it natural to destress the second occurrence of *town*.

The destressing requirement is probably due to contribution stress would make to the meaning in this construction. Consider (i), where also a repeated occurrence of the noun *book* is stressed: The stress indicates a contrast between the book John read and the book Mary read, with respect to their 'bookness'. In effect, (i) entails that what John read wasn't really a book. Therefore, I assume the two nouns *book* in (i), despite having similar phonology, differ in the sense relevant for the identity condition on traces. The destressing requirement argues therefore that the identity required isn't identity of lexical form, but identity of meaning. Section 3.2 presents more arguments for this conclusion.

(i) John read a book and Mary a BOOK.

the same.[2]

(11) $\Big[$every town near a [town, Op_y Mary $\underbrace{\text{visited } [y, \text{ town}]}_{\text{elided VP}}$]$\Big]$ λx John $\underbrace{\text{visited } [x, \text{ town}]}_{\text{antecedent}}$

It's important to go through the argument that (10) makes to see that it's independent support for the result of the previous chapter: (10) shows that for the well-formedness of ACD the head nouns of the antecedents of the two traces involved must be identical, namely of the trace of QR and of the trace internal to the relative clause. Why would there be such a requirement? As already mentioned in the introduction, it's established that an elided VP must be identical to its antecedent. Therefore, I conclude that the head nouns are represented in the elided VP and its antecedent, respectively. The only part of the two VPs related to the head nouns are the traces. Hence, it's natural to assume that, if anywhere, the head nouns are represented in the trace positions. Therefore, (10) argues that the head noun of a

[2]The copy identity approach shares the prediction (10) with—at least a benevolent interpretation of—a proposal of Lappin (1984). Lappin proposes, in effect, that two traces or pronouns are identical if they *can be naturally interpreted as having the same intended range of possible values.* (Lappin 1984:(10)) He, however, doesn't discuss contrasts like (10) and his proposal is too vague to be sure of this prediction. There are other differences between the copy identity approach I'm developing and Lappin's proposal. For one, Lappin doesn't derive the identity condition from properties of the semantic representation in the it's done here, but suggests that the condition is pragmatic which, as far as I can see, he presents no motivation for. Secondly, Lappin's condition applies to all traces and bound pronouns, which isn't true of the copy identity approach pursued here as discussed in section 3.3. The examples in (ib) and (iib) show that this aspect of Lappin's (1984) proposal makes wrong predictions (see also Fiengo and May 1994).

(i) a. Here is the man who Bill saw, and here is the man who he didn't ⟨see⟩. (Lappin 1984:(21b))

b. Here is the man who Bill saw, and here is the woman who he did ⟨see⟩.

(ii) a. [Every friend of John's]$_i$ wants Mary to kiss him$_i$, but [none of the little fellows]$_j$ believes that she will ⟨kiss him$_j$⟩ (Lappin 1984:(10))

b. [Every friend of John's]$_i$ wants Mary to kiss him$_i$, while [every friends of Bill's]$_j$ wants Sue to ⟨kiss him$_j$⟩.

QR-chain is represented in the trace position, and that the head noun of the relative clause external head is represented in the relative clause internal head position.

Notice that the argument is independent of the arguments given in chapter 2 in favor of the same conclusion. In 2.1, I argued with example (12a), repeated from (13b) on page 40, that for Condition C the NP-part phrase that moves for ACD-resolution remains represented in the trace position. The LF-representation proposed for (12a) is repeated in (12b).

(12) a. *In the end, I did ask him$_i$ to teach the book of David$_i$'s that Irene wanted me to ⟨ask him to teach⟩.

b. $\Big[$the book of David's λy Irene$_i$ wanted me to $\underbrace{\text{ask him to teach } [y, \text{ book of David's}]}_{\text{elided VP}}\Big]$
λx I asked him$_i$ to $\underbrace{\text{teach } [x, \text{ book of David's}]}_{\text{antecedent}}$

Furthermore, I argued in section 2.4 based on Safir's (1998) discovery in (13a), repeated from (75) on page 78, that also the relative clause internal trace position contains lexical material. Namely, if the NP-part of the head of the relative clause is represented there as in (13b), (13a) is predicted to violate strong crossover.

(13) a. *Pictures of anyone$_i$ which he$_i$ displays prominently are likely to be attractive ones.

b. *anyone λx $\Big[\underbrace{\text{pictures of } [x, \text{ one}]}_{\text{external head}}$ [which $\underbrace{\text{picture of } [x, \text{ one}]}_{\text{internal head}}$] λy he$_x$ displays prominently $\underbrace{[y, \text{ pictures of } [x, \text{ one}]]}_{\text{internal trace}}\Big]$ are likely to be attractive ones.

The argument based on (10) provides independent confirmation of these two conclusions of chapter 2. In the remainder of this section, I give further evidence for this interpretation of (10) and the parallelism to the arguments of chapter 2. I start by adding some more examples just like those in (10), then I show that it is not just the head noun, but the NP-part that matters for the identity of traces, just like it does for binding theory as argued in chapter 2. Finally, I show a difference between A- and A-bar-movement that parallels the A/A-bar distinction found with respect to binding.

Both examples in (10) marginally allow an interpretation of the elided VP as *visit every town near t*. This is expected because extraction out of a reduced relative clause is marginally possible as in (14a), and on this reading the operator binding both traces is the same.

(14) a. ??Which lake did you visit every town near?

b. *Which lake did you visit every town that's near?

As (14b) shows, extraction out of a full relative clause is impossible. The examples in (15), repeated from (5) in the introduction, and (16) show a similar contrast to (10), but the don't allow a different reading of the elided VP than the indicated one.

(15) a. John visited a town that's near the town Mary did ⟨visit⟩.

b. *John visited a town that's near the lake Mary did ⟨visit⟩.

(16) a. Jon ordered a drink that's more expensive than the drink Sue did ⟨order⟩

b. *Jon ordered a drink that's more expensive than the dish Sue did ⟨order⟩

The repetition of the same noun within one sentence is usually a little unnatural and most speakers prefer to replace the second occurrence with a *one*-anaphor. As the examples in (17) show, the good examples of (14), (15), and (16) are also good with a *one*-anaphor in place of the repeated noun. We can ignore the question whether *one*-anaphora are analyzed as NP-ellipsis (Lakoff 1968) or NP-pronouns (Jackendoff 1977:58-60); on either assumption the facts in (17) are expected: Since on either one the *one* anaphor is, semantically at least, not different from a full NP that could be used to paraphrase it, the examples in (17) are expected to behave just like (10a), (15a), and (16a).

(17) a. John visited every town near the one Mary did ⟨visit⟩.

b. John visited a town that's near the one Mary did ⟨visit⟩.

c. Jon ordered a drink that's more expensive than the one Martin did ⟨order⟩

Consider the LF-representation of (17b) given in (18). In (18), the lexical content of the trace position in the relative clause is indicated as *town*, though the external head of the relative clause is *one*. However, the representation in (17b) is possible if either of the following two assumptions is correct: *one* is a phonologically reduced expression of *town*, or the content of the internal head of a matching relative must have the meaning as that of the external head. I believe that, at the least the latter assumption is correct for the reasons given in section 2.4. Hence, the LF-

representation in (18) is possible, and satisfies the identity requirement of the elided VP straightforwardly.[3]

(18) $\Big[$ a town that's near the one λy Mary did $\underbrace{\text{visit } [y, \text{ town}]}_{\text{elided VP}}\Big]$ λx John $\underbrace{\text{visited } [x, \text{ town}]}_{\text{antecedent}}$

Jacobson (1998a) points out that the ungrammaticality of Kennedy's examples is also found in cases like (19a), where ACD is resolved by overt *wh*-movement rather than by covert movement. In (19a), the antecedent of the trace of *wh*-movement is *town* while the head of the relative clause is *lake*. As we see in (19b) and (19c), if the two nouns are the same or one is anaphoric to the other, the example improves.

(19) a. *Do you know which town near a lake Mary did ⟨visit⟩ John visited?

b. Do you know which town near a town Mary did ⟨visit⟩ John visited?

c. Do you know which town near the one Mary did ⟨visit⟩ John visited?

Example (20) shows that the judgement doesn't change if the other VP is elided—the one that contains the trace of *wh*-movement. As Jacobson (1998a) already notes, (20a), doesn't allow deletion of the VP containing the trace of *wh*-movement. The contrast with (20b) and (20c) shows that, again, the difference of lexical content of the antecedents of the two trace in (20a) causes the ungrammaticality.

(20) a. *Do you know which town near a lake Mary visited John did ⟨visit⟩?

[3]In the next section, I present arguments that the kind of identity required between the elided VP and its antecedent is identity of meaning. This implies that even if *one* occupies the relative clause internal trace position, ACD should be possible as long as *one* means the same as *town*, the content of the QR-trace in the antecedent.

b. Do you know which town near a town Mary visited John did ⟨visit⟩?

c. Do you know which town near the one Mary visited John did ⟨visit⟩?

This argues for the LF-representation of (20a) in (21) (and it would be the one of (19a) if the labels *antecedent* and elided VP were interchanged). In (21), the head noun of the *wh*-phrase is represented in the position of the *wh*-trace, while the head noun of the relative clause head is represented in the relative clause internal trace position. Again, the identity requirement of the elided VP and its antecedent is clearly violated by the content of the traces. That ACD is possible in the b) and c) examples of (19) and (20), on the other hand, shows that the relative clause isn't represented in the position of the *wh*-trace.

(21) $^{*}\Big[$which town near a lake λy Mary $\underbrace{\text{visited } [y, \text{ lake}]}_{\neq \text{ antecedent}}\Big]$ λx John $\underbrace{\text{visited } [x, \text{ town}]}_{\text{elided VP}}$

Note that, again, the conclusion just reached parallels the conclusion pointed out in section 2.2, namely that overt *wh*-movement also represents the lexical content of the NP-part in the trace position. There the argument that the NP-part of a *wh*-phrase must be represented in the trace position, while a relative clause adjoined to it need not be, was the argument/adjunct asymmetry with Condition C of Freidin (1986) and Lebeaux (1988) illustrated in (22) (repeated from (20) on page 44).

(22) a. *[Which argument that John$_i$ was wrong]$_j$ did he$_i$ accept t_j in the end?

b. [Which argument that John$_i$ had criticized]$_j$ did he$_i$ accept t_j in the end?

So far, there is one apparent difference between the conclusions reached here based on identity conditions on traces and the conclusions reached in chapter 2 based on the distribution of Condition C effects. Namely, the discussion in this section has produced arguments that the head noun of an antecedent is represented in the trace position. The distribution of Condition C effects argued that the NP-part, which is the head noun plus its complement, is represented in the trace position. The difference is only apparent: In the examples so far, the NP-parts of the relative clause head and the DP moving for ACD-resolution consisted only out of the head noun. The paradigm in (23) shows that similar contrasts are also found in a case where the NP-parts have the same head-noun but the arguments of the head noun is different. However, repeating a complex NP within the same sentence as in (23b) is so unnatural that the contrast to (23c) is very weak. The contrast between (23a) and (23c) is clear, though. (24) shows the relevant aspect of the LF-representation proposed for (23c)—the elided VP is not identical to its antecedent.

(23) a. Bill gave a description of Mary that's similar to the one John did ⟨give⟩

b.$^{??}$Bill gave a description of Mary that's similar to the description of Mary John did ⟨give⟩

c.$^{*?}$Bill gave a description of Mary that's similar to the description of Sue John did ⟨give⟩

(24) $^{*}\Big[$a description of Mary that's similar to the description of Sue λy John did $\underbrace{\text{give } [y, \text{ a description of Sue}]}_{\text{elided VP}}\Big]$ λx Bill $\underbrace{\text{gave a } [x, \text{ a description of Mary}]}_{\neq \text{ antecedent}}$

The lack of a contrast in (25) below shows that different adjuncts of the antecedents of the two traces don't block ACD. One possible explanation could be that the difference between (23c) and (25b) mirrors the argument/adjunct distinction of binding theory. A conclusive judgement on this issue, however, would need to take into account the considerations brought up in the next section 3.2. (See footnote 6 below).

(25) a. John visited a town near Madrid that had signs for the one Bill did ⟨visit⟩.

b. John visited a town near Madrid that had signs for the one near Rome Bill did ⟨visit⟩.

Example (26) is another place where the predictions of an identity of NP-parts requirement differ from an identity of head noun requirement. In (26), the head of the relative clause is an argument of the noun heading the NP-part of the DP that moves for ACD-resolution: Even though the head-nouns of the two NP-parts involved in (26) are identical, the examples are ungrammatical. The two NP-parts itself aren't identical in (26), because one contains the other. For example, in (26a), the NP-part of the relative clause head is only *picture*, but the NP-part in the antecedent is *picture of a picture*. This is captured in the representation in (27).[4]

(26) a. *Susi produced a picture of a picture Meltem did ⟨produce⟩.

[4]The LF-representation I actually assume for (26a) must contain the relative clause in the trace position as in (i), and therefore doesn't allow ACD (See the discussion of example (28) on page 47). With respect to question of whether the NP-part of the noun head is represented (i) leads to the same conclusion.

(i) [a picture of a picture Meltem did produce] λx Susi produced [x, picture of a picture Meltem did produce]

b. *Jonathan visited every relative of the relative Danny did ⟨visit⟩.

(27) *[a picture of a picture λy Meltem $\underbrace{\text{produce } [y, \text{ picture}]}_{\text{elided VP}}$]

λx Susi $\underbrace{\text{produced } [x, \text{ picture of a picture}]}_{\neq \text{ antecedent}}$

The contrast in (28) brings out the difference between arguments and adjuncts as a minimal pair. For the judgement, imagine that John's art is painting pictures of Dali's pictures. One day, John meets Dali and Dali tells him about his plan for a new great painting. John likes the plan a lot, and immediately makes his own plans based on Dali's plan. In this context, (28b) is an acceptable sentence, but (28a) remains unacceptable.

(28) a. *John is planning to paint many pictures of the one Dali is ⟨planning to paint.⟩

b. John is planning to paint many pictures showing the one Dali is ⟨planning to paint.⟩

Since the head of the relative clause *one* is an argument of the higher NP in (28a) and therefore inside the NP-part of the DP that moves for ACD-resolution, (28a) is expected to be bad. (28b), on the other hand, is expected to have the same status as (10) because the head of the relative clause is contained in an adjunct to the higher NP-part.

Given the parallelism of chapter 2 and the conclusions here, it's expected that a difference between A- and A-bar-chains is also found with the trace identity requirement. Recall from section 2.3 that A-chains and A-bar-chains differ with respect to Condition C as illustrated in (29) (repeated from (41) on page 55): While *Kai* in the A-bar moved phrase behaves as if in the trace position with respect to Condition C, the R-expression *Kai* in (29b) can be coreferent with the pronoun *him*. Hence, section 2.3 concluded that the requirement that the NP-part must be represented in the trace position of a chain, only applies to A-bar chains.

(29) a. *[Which relative of Kai$_j$'s]$_i$ did he$_j$ say t_i likes Kazuko.

b. [One relative of Kai$_j$'s]$_i$ seemed to him$_j$ to t_i like Kazuko.

The examples in (30) and (31) show a contrast between topicalization (A-bar-movement) and passivization (A-movement) that argues that the requirement on trace identity is sensitive to the A/A-bar-distinction as well. Namely, the passive examples in (30a) and (31a) are acceptable, while the topicalization cases in (30b) and (31b) are ungrammatical.

(30) a. The town near the lake that was ⟨visited by vandals⟩ seems to have been visited by vandals, as well.

b. *The town near the lake they did ⟨visit⟩, the vandals seem to have visited, as well.

(31) a. The town near the lake that was visited by vandals seems to have been ⟨visited by vandals⟩, as well.

b. *The town near the lake they visited, the vandals seem to have ⟨visited⟩, as well.

These contrasts argue, based on the identity criterion, that the trace in an A-chain need not contain lexical material of the antecedent. Consider the LF-representation of (30a) in (32). If both the elided VP and its antecedent contain only a variable in the object position, but not the lexical material of the antecedent, the identity condition of VP-ellipsis is satisfied.

(32) [The town near the lake λz [z, lake] λy was $\underbrace{\text{visited by vandals } [y]}_{\text{elided VP}}$] λx seems to $\underbrace{\text{have been visited } [x] \text{ by vandals}}_{\text{antecedent}}$, as well.

Therefore, the contrasts in (30) and (31) provide independent support for the A/A-bar distinction as stated in 2.3. There are, however, examples like (33) where deletion of a VP containing an A-trace is blocked, even though the identity condition is predicted to be satisfied. (34) is the LF-representation of (33a) assuming that the VP internal subject hypothesis, which claims that the subject A-moves from a VP-internal position to its surface position (see Webelhuth 1995:60-64 and references therein). In fact, most of the examples discussed in the papers of Kennedy (1994) and Heim (1997a) are examples with A-traces, and Kennedy and Heim both view examples like (33) as support for the index identity view.

(33) a. *A proof that God exists does ⟨exist⟩ (Wasow 1972:93)

b. *Every man who said George would buy some salmon did ⟨buy some salmon⟩ (Kennedy 1994:(2b))

(34) A proof that God λy $\underbrace{\text{exists}}_{\text{antecedent}}$ λx does $\underbrace{[x] \text{ exist}}_{\text{elided VP}}$

As Kennedy (1994:fn. 3) notes and Heim (1997a) discusses in detail, the grammaticality of examples like in (33) improves for many speakers with the addition of focus particles like *too*, *as well*, or *instead*, which is not the case for examples with A-bar movement like (2). Hence, I reject the conclusion that the examples in (33) should receive a similar explanation as Kennedy's A-bar movement cases. Example (33a) is probably ill-formed because it requires scope reconstruction of the subject into a VP-internal position (see for example Diesing 1992). In section 4.1, I provide an account that predicts that examples like (33b), while not ungrammatical, are difficult to parse.

3.2 Semantic Content of the Trace

Two independent lines of argumentation established that parts of the antecedent of trace are represented in the trace position at the LF-level. In chapter 2, the argument was based on the distribution of Condition C effects. In the previous section 3.1, I presented an argument based on the identity condition between traces imposed by VP-ellipsis. While this correspondence is quite remarkable, it still leaves it open what the contribution of the lexical material in the trace is to interpretation. Up to now,

it's conceivable that the material in the trace doesn't contribute to interpretation at all, except in the cases of variable binding in section 2.2. In this section, I argue that in the lexical material in the trace position is also interpreted there—it constitutes the semantic content of the trace. In particular, I show that the range of entailments drawn from a constituent containing a trace, but not its antecedent, is affected by the semantic content of the trace.

The alternative position I'm arguing against here doesn't, at least at this point, look very attractive. The assumption that the lexical material in a trace position isn't interpreted there, but in the position of the antecedent, would necessitate the following additional assumptions: To begin with, it requires the assumption that the lexical material in the trace position is also represented in the antecedent, so that no information is lost if the material in the trace position is ignored. This assumption is unproblematic (In fact, I have been making this assumption throughout and give an argument for it in section 3.3.3 below), except when the fronted material contains a variable in (35a) (repeated from (34) on page 50). Here, I have been assuming a representation like that in (35b), where the part of the fronted constituent that contains the bound variable is only represented in the trace position of the *wh*-chain. On the assumption that normally, lexical material in the trace position is ignored, either the case of a bound variable must constitute an exception to this, or additional semantic mechanisms that allow the interpretation of a bound variable in a position outside the c-domain of its binder must be postulated. However, such mechanisms have been postulated; for example Skolem-functions in Engdahl (1986) and Chierchia (1993) and abstraction over assignments as in Sternefeld (1998) (technically a generalization

of Skolem-functions).

(35) a. [Which paper of his$_k$ that Mary$_j$ was given]$_i$ did she$_j$ tell every student$_k$ to revise t_i?

b. $\Big[$Which $[\lambda z$ Mary$_j$ was given $[z]]\Big]$ λx did she$_j$ tell every student$_i$ to revise $[x$, paper of his$_i]$?

A second consequence of the assumption that the lexical content of the trace is semantically vacuous, is the existence of two kinds of deletion at the LF-level. This is clear in examples like (36a) (repeated from (2)), and the same point could be made for (13b) in section 2.1. In the QR-chain of the LF-representation of (36a), as repeated in (36b), the ACD-relative clause is represented only in the operator position of the chain, while the NP-part of the QR-chain is represented in both the operator and the trace position of this chain. However, if the lexical material in the trace position is also not entering interpretation in that position, (36) in effect involves two steps of deletion: Before Condition C and the identity condition of ellipsis apply, the relative clause is deleted in the position of the QR-trace. After the two conditions applied, the NP-part is deleted in the trace position. While this position isn't incoherent, it's also not particularly attractive from my point of view since the second step of deletion operations seems unmotivated.

(36) a. *Polly visited every town in a country Eric did ⟨visit⟩.

b. *every [town, in a country Op$_y$ Eric $\underbrace{\text{visited } [y, \text{ country}]}_{\text{elided VP}}$]

$$\lambda x \text{ Polly } \underbrace{\text{visited } [x, \text{ town}]}_{\neq \text{ antecedent}}$$

The argument I give now in favor of the trace actually having semantic content is quite a bit stronger than the preceeding two arguments. The argument comes from a closer look at Kennedy's restriction on ACD. Consider the data in (37): (37a) and (37b) are repeated from (15) above, however (37c) is new. Surprisingly, (37c) is almost as good as (37a), though the lexical material of the QR-trace is predicted to be *town* while that of the relative clause internal trace is *city*. As in (37a), the judgement requires leaving *city* unstressed; if *city* is stressed, (37c) is unacceptable (cf. footnote 1).[5]

(37) a. John visited a town that's near the town Mary did ⟨visit⟩.

b. *John visited a town that's near the lake Mary did ⟨visit⟩.

c. ?John visited a $\underbrace{\text{town}}_{\text{QRNP}}$ that's near the $\underbrace{\text{city}}_{\text{RCNP}}$ Mary did ⟨visit⟩

For the following discussion it is convenient to have the following two terms at our disposal: The NP-part of the DP that moves covertly for ACD-resolution I call *QRNP*, and the NP-part of the head of the ACD-relative I call the *RCNP*. The empirical generalization argued for in section 3.1 could then be stated as follows: ACD is possible if and only if QRNP is equal to RCNP. (37c) shows that this generalization is not exactly correct. While ACD is always possible when QRNP and RCNP are identical, there seem to be more cases where ACD is possible. (38c), (38d), (39c),

[5]In fact, for some people, it's possible to leave *lake* in (37b) unstressed and then (37b) becomes acceptable. The explanation given below for (37c) carries over to this case as well.

and (39d) show that (37c) isn't the only exception while (38a) and (38b), as well as (39a) and (39b) (repeated from (16)) display the contrast familiar from the previous section.

(38) a. John lives in a city that's close to a city Mary used to ⟨live in⟩.

b. *John lives in a city that's close to a castle Mary used to ⟨live in⟩.

c. ?John lives in a city that's close to a town Mary used to ⟨live in⟩.

d. ?John lives in a city that's close to where Mary used to ⟨live⟩.

(39) a. Jon ordered a drink that's more expensive than the drink Sue did ⟨order⟩

b. *Jon ordered a drink that's more expensive than the dish Sue did ⟨order⟩

c. ??Jon ordered a cocktail that's more expensive than the beer Sue did ⟨order⟩

d. ?Jon ordered a drink that's more expensive than what Sue did ⟨order⟩

It seems that the relationship between the RCNP and the QRNP that is required for the ACD to be acceptable has a semantic character. While I found a great amount of speaker variation with respect to examples like the above ones, every speaker I consulted found that the semantic relationship of RCNP and QRNP was the deciding factor in the judgements. This semantic character supports strongly that it's the semantic contribution of the trace content which determines the identity of traces. The account I develop now makes the semantic character precise, and specifically aims to accounts for the facts in (38) and (39). The generalization I end up with in this section is that ACD is possible if and only if QRNP denotes a subset of RCNP as in (38d) and (39d) or QRNP and RCNP are from the semantic field as in (38c) and

(39c). It remains to be seen whether this generalization is correct exactly as stated here; I do believe, though, that the general point concerning the semantic character is.[6]

This generalization I argue is implied by the conjunction of two things: the assumption that the lexical content of the trace contributes to interpretation and the right account of the identity requirement. As already argued, the identity requirement imposed on RCNP and QRNP in ACD is part of the identity requirement of VP-ellipsis in general. Therefore, I now present an account of the identity requirement of VP-ellipsis. I, then, come back to ACD, and show that this account predicts the generalization just mentioned.

It is a well known observation that VP-ellipsis requires some form of semantic identity or sameness of meaning between the elided VP and its antecedent (see Sag 1976:92-95 and references therein). For example, in (40a), the first conjunct is ambiguous between a volitional reading and an idiomatic, non-volitional reading, but the interpretation of the elided VP in the second conjunct has to correspond to that of the first VP. Similarly, the first VP in (40b) can receive an interpretation like *put paint on the bike* or one like *made a picture of a bike*, but the elided VP has to correspond in meaning to the first.

[6]In the present context, reconsider the example in (i), repeated from (25b). The well-formedness of (i) could be due to the fact that the RCNP and the QRNP are from the same semantic field even if the modifiers *near Madrid* and *near Rome* are represented in the trace positions. Therefore, it seems at this point impossible to test using examples with the structure of Kennedy's puzzle whether modifiers are represented in the trace position.

(i) John visited a town near Madrid that had signs for the one near Rome Bill did ⟨visit⟩.

(40) a. John hit the wall and then Pete did ⟨hit the wall⟩

b. John painted a bike after Mary did ⟨paint a bike⟩

An important observation that plays a role in the following is that destressing of a VP displays a semantic requirement very similar to that of VP-deletion (Tancredi 1992, Rooth 1992b, Wold 1995, Fox 1998a). For example, the examples in (41) show the same disambiguation as those in (40). (I represent destressing in (41) and in the following by italics and a reduced character size.)

(41) a. John hit the wall and then Pete *hit the wall*

b. John painted the bicycle after Mary *paint the bicycle*

The first to discuss in detail the claim that VP-ellipsis requires identity of meaning is Sag (1976). Sag, however, rejects this claim, and opts instead for a requirement that LF-representations must be identical. His only reason is the example in (42), where *a child* is interpreted generically in the first VP, but existentially in the elided VP. Sag's conclusion isn't forced by example (42), at least not on a quantificational variability account of generic interpretations of indefinites (Wilkinson 1991). On such an account, the first conjunct contains a covert generic quantifier *usually* that lends its quantificational force to the indefinite *a child* and could take scope outside of VP. The indefinite *a child* can, on this account, receive an existential interpretation in both the elided VP and its antecedent.

(42) They caned a child severely when I was a child, but not like Miss Grundy did ⟨cane a child⟩ yesterday. (Sag 1976:(2.0.13))

Since there are examples like (40) where I know of no argument in favor of an LF-difference between the two interpretations that VP-ellipsis draws a distinction between, I assume that VP-ellipsis requires identity of meaning to an antecedent. This is also assumed by Tancredi (1992), Rooth (1992b), Wold (1995), and Fox (1998a).

The nature of the identity condition affects the question what the nature of the content of the trace is because the identity condition is sensitive to it. If there is only an identity of meaning requirement on elided VPs, the lexical material in the trace position must be interpreted there since it's relevant to the identity condition. However, exact semantic identity wouldn't explain the examples like (37c). Hence, the identity requirement must be a more complicated technical condition, not the intuitive notion of *identical* I have made appeal to up to this point. Whether this more complicated condition makes references to the form or to the meaning is what needs to understood. Before coming back to (37c), I summarize the literature on this question. In the literature, the main disagreement is whether there is also a requirement of identity of form in addition to the identity of meaning requirement. If there's a requirement of identity of form, the requirement that traces must have identical lexical content might be a purely formal requirement. In that case, there's no evidence that traces have semantic content other than being variables.

Rooth (1992b), in particular, argues that there is also a requirement of identity

of form. One of his argument has the following structure:[7] As seen above, destressing and deletion seem to share the semantic identity requirement. There are cases, though, where destressing and also VP-ellipsis can be licensed by satisfying a weaker requirement of *indirect identity*. Therefore, the semantic requirement allows indirect identity. This, however, overgenerates possible interpretations for VP-ellipsis. Therefore, there must be an additional requirement, identity of form, for elided material. I now present Rooth's (1992b) argument for identity of form in more detail. I then summarize a different way to draw the distinction between destressing and deletion argued for by Fox (1998a), which doesn't require identity of form. Finally, I argue that the facts from ACD above argue for Fox's (1998a) statement of the condition and also show that the trace has semantic content.

Strict semantic identity is too strong in cases of destressing like those in (43). As Tancredi (1992) and Rooth (1992b) argue, such examples argue that destressing can be licensed under identity of meaning with a sentence that is not part of the discourse itself, but rather, entailed by the discourse. For example, the first conjunct in (43b) entails that Mary is having a drink, and it's semantic identity to the VP of the second conjunct that can license destressing in (43a). I call this relationship, where identity is satisfied by an entailment of the antecedent, *indirect identity*.

(43) a. John enjoyed one Russian novel, and even Bill *read a book.*

b. Mary ordered a beer and Sue *is having a drink,* too.

[7]The other argument concerns the requirement that sloppy readings must have a same dependency in the antecedent. See the dicussion of the parallel dependencies requirement at the end of section 4.1.2.

While Tancredi (1992) claims that indirect identity is only found with destressed VPs, Rooth (1992b) shows a case where an elided VP seems to be only indirectly identical to its antecedent. The argument relies on an observation illustrated in (44a) and (44b). If *her* in the first conjunct refers to Mary, (44a) allows a so called *sloppy* interpretation: the pronoun *her* in the elided VP need not refer to Mary, but can also refer to Jane instead. However, as shown in (44b), the pronoun *her* in the elided VP cannot be taken to refer to Sue. (44) argues for a constraint on sloppy readings such as that given in (45) (see Ristad 1990, Fiengo and May 1994:96–117, Rooth 1992b, Fox 1998c for further evidence for (45)). I assume (45) for the rest of this chapter as an empirical generalization—I discuss briefly at the end of section 4.1.2 how Rooth (1992b) actually derives most cases of the requirement (44).

(44) a. First, John told Mary$_i$ I was bad-mouthing her$_i$,
and then Sue told Jane$_j$ I was ⟨bad-mouthing her$_j$⟩

b. *First, John told Mary$_i$ I was bad-mouthing her$_i$,
and then Sue$_j$ told Jane I was ⟨bad-mouthing her$_j$⟩

(45) *Parallel Dependencies:* If a pronoun isn't identical in reference to the corresponding pronoun in the antecedent, it must stand in the same structural relationship to its binder as the corresponding pronoun in the antecedent.

Assuming (45), Rooth's (1992b) argument is based on (46). In (46), as indicated the pronoun *her* in the elided VP can refer to Sue, even when the the corresponding pronoun in the antecedent refers to Mary. In (46), the condition (45) seems to be

violated, because the antecedent of the pronoun in the second conjunct is the subject, whereas that of the corresponding pronoun in the first conjunct is the object.[8] Rooth proposes that the violation of (45) in (46) is only apparent; ellipsis in (46) isn't licensed by direct identity with the first conjunct, but by indirect identity where the relevant entailment of the first conjunct is the sentence (47).

(46) First, John told Mary$_i$ I was bad-mouthing her$_i$,
and then Sue$_j$ heard I was ⟨bad-mouthing her$_j$⟩ (Rooth 1992b:(30))

(47) Mary$_i$ heard I was bad-mouthing her$_i$.

Therefore, (46) argues that indirect identity can license ellipsis as well. At the end of section 3.3.2, I summarize an additional argument from Jacobson (1998a) that indirect identity can license ellipsis. However, there are many cases where indirect identity is lenient: Rooth (1992b) and Tancredi (1992) show with examples likes those in (48) that ellipsis must require more than indirect identity. In (48a), which contrasts with destressing in (43a), the elided VP cannot receive the interpretation indicated though it would satisfy indirect identity. Similarly, (48b) and (48c) clearly contrast with (46), since the entailment argued to be involved in the licensing of deletion in (46) cannot license deletion in (48a) and (48b).

[8]Fiengo and May (1994:100) claim that the surface subject of the second conjunct of (45) is in fact an object, and therefore (46) doesn't violate the condition (45). But, as Danny Fox (p.c.) points out, Fiengo and May's (1994) account predicts that sloppy ellipsis should also be possible if the order of the conjuncts is reversed, which isn't the case as (i) shows. Rooth's (1992b) analysis of (46) makes the right prediction for (i).

(i) *First, Sue$_j$ heard I was bad-mouthing her$_j$, and then John told Mary$_i$ I was ⟨bad-mouthing her$_i$⟩

(48) a. *John enjoyed one Russian novel, and even Bill did ⟨read a book.⟩

b. *First someone told Mary about the budget cuts and then Sue did ⟨hear about the budget cuts⟩ (Rooth 1992b:(15))

c. *First John told Mary$_i$ I was bad-mouthing her$_i$ and then Sue$_j$ did ⟨hear I was bad-mouthing her$_j$⟩

The argument shows that VP-ellipsis has an additional requirement which distinguishes it from destressing. Rooth (1992b) proposes that this is an identity of form requirement; specifically, he refers to the *reconstruction* relation of Fiengo and May (1994). While this is a possible account of (48) it could also be the case that the semantic identity requirement VP-ellipsis imposes is slightly stricter than that of destressing, in a way similar to the proposal of Tancredi (1992). Fox (1998a:ch. 3) develops such a proposal, which I adopt. The arguments of Fox (1998a) for his proposal, and against that of Rooth (1992b), are too intricate to summarize here; instead I point out some of the problems an identity of form proposal faces, before I summarize Fox's (1998a) proposal. The account of facts like (37c) I present then is an additional argument for Fox's (1998a) approach.

Consider the examples in (49) in the context of an identity of form requirement. In each of them, identity of form must be compromised because, if the elided VP would have to be identical in the choice of lexical items to the antecedent VP, all four sentences are predicted to be ungrammatical. Therefore, as Johnson (1996:7) argues, examples like these are a significant challenge for any version of an identity

of form requirement, and the best developed proposal of this kind I'm aware of is that of Fiengo and May (1994:220). According to Fiengo and May, the examples require essentially of list expressions that satisfy identity of form despite being lexically different, for which Fiengo and May (1994) introduce the term *vehicle change* as mentioned in section 2.4.

(49) a. John doesn't see anyone, but Bill does. (Sag 1976:(2.3.39))

b. Jonathan didn't have a red cent, but Susi did ⟨have money⟩.

c. John won't leave until midnight, but Bill will ⟨leave before midnight⟩ (Chomsky 1972a:(75))[9]

d. Because Sue didn't want to buy Bill$_i$'s dinner, he$_i$ had to ⟨buy his$_i$ dinner⟩.

Based on other (stronger) arguments, Fox (1998a) proposes to replace the concept of identity of form that Rooth (1992b) appealed to with a stricter condition on the semantic relation between an elided VP and its overt antecedent. Specifically, he proposes a restriction on indirect identity that amounts to the recursive condition in (50):[10]

(50) The antecedent VP$_{antecedent}$ and the elided VP$_{elided}$, which is part of a sentence S, can satisfy indirect identity only if there's no VP$_{elided}'$ such that

[9]Grinder and Postal (1971) judge the sentence (49c) ungrammatical. However, Chomsky (1972a) and my informants do find it acceptable with the appropriate contrastive foci on the subjects *John* and *Bill*. See also Sag (1976:158-60) for discussion.

[10]The difference is mainly that Fox states the condition for focus domains, rather than for elided VPs specifically. This presupposes the semantics of focus which are only introduced in section 3.3. Moreover, this aspect of Fox's proposal would not be useful at this point.

a. replacing VP_{elided} with VP_{elided}' in S yield a grammatical sentence S′

b. S′ is logically stronger than S

c. VP_{elided}' is (directly or indirectly) identical with $\text{VP}_{antecedent}$

The restriction (50) can also be seen as a constraint on the parsing (or recovery) of elided material. Then it could be stated as follows: For an elided VP site, choose a parse with an interpretation as strong as possible, but entailed by the antecedent, and that's compatible with the overt material surrounding the deletion site.

Fox's condition accounts for Rooth's problem that VP-ellipsis of indirectly identical material was licensed in (46), but blocked in (48). Consider (48a), as repeated in (51), first. In this case, there is an interpretation of the elided VP site is possible that results in a stronger statement and that is also indirectly (as well as directly) identical to the first conjunct: *Bill enjoyed one Russian novel.* Therefore, the interpretation indicated in (51) is blocked. In general, it will be the case that the interpretation of the elided VP directly identical to the antecedent is the one chosen, unless there is a requirement imposed by the material surrounding the elided VP which blocks direct identity.

(51) *John enjoyed one Russian novel, and even Bill did ⟨read a book.⟩

The difference between (51) and (46), repeated in (52), is that in (52) material outside of the elided VP together with the constraint (45) forces indirect identity. The only parse which satisfies (45) is the strict reading. But, the strict reading stands in no

entailment relation to the sloppy reading indicated in (52), and therefore doesn't block (52). Among the sloppy readings, condition (45) forces a parse of the elided VP as a verb with an object pronoun. Of these, the one in (52) is the logically strongest, and hence, is the one possible for the elided VP.

(52) First, John told Mary$_i$ I was bad-mouthing her$_i$, and then Sue$_j$ heard I was ⟨bad-mouthing her$_j$⟩

In sum, Fox (1998a) claims that there's no requirement of identity of form between the elided VP and its antecedent, but only a slightly stronger semantic identity condition than the one for destressing. Going back to the question whether the lexical material in the trace position is interpreted there, Fox's (1998a) account is only compatible with one answer. Namely if Fox is right, it predicts that traces must have semantic content beyond being a variable, since it must be their semantic contribution that's blocking ACD in cases like (2). Additional support for this conclusion and Fox's (1998a) account comes from the facts in (37) to (39), as I show now.

Consider first the examples where RCNP denotes a subset of RCQR, like those in (53) (repeated from (38d) and (39d)). The contrast in (54) shows that RCNP must denote a subset of RCQR, and not the other way round.

(53) a. ?John lives in a city that's close to where Mary used to ⟨live⟩.

b. ?Jon ordered a drink that's more expensive than what Sue did ⟨order⟩

(54) a. Last night, I talked to a bachelor who looked like the guy you did ⟨talk to⟩

b. *Last night, I talked to a guy who looked like the bachelor you did ⟨talk to⟩

I propose that in these cases indirect identity is involved. Consider the elided VP and its antecedent in the LF-representation of (54a) in (55). The matching requirement on relative clauses doesn't allow a lexical content of the trace in the elided VP other than *guy* (see (56) below). Therefore, if the antecedent in (55) has an entailment where the lexical content of the trace is [x, guy], the elided VP satisfies the requirements for indirect identity that no logically stronger replacement of the elided VP is grammatically possible.

(55) $\Big[$a bachelor who looked like the guy λy you $\underbrace{\text{talked to } [y, \text{ guy}]}_{\text{elided VP}}\Big]$

λx I $\underbrace{\text{talked to } [x, \text{ bachelor}]}_{\text{antecedent}}$

Treating (53) and (54) as cases of indirect identity, therefore, requires that the contribution of [x, bachelor] to the meaning of the antecedent is such that it allows an entailment to the VP where it's replaced by [x, guy]. For now, let us assume that the meaning of the trace [x, bachelor] can be paraphrased as the indefinite *a bachelor*. Then, it's indeed predicted that ellipsis is licensed in (55), since 'I talked to a bachelor' entails 'I talked to a guy'.

The possibility of indirect identity makes it necessary to briefly talk about the analysis of relative clauses again. In section 2.4, I argued one class of relative clauses, the matching relatives, involve both an internal and external head NP, and I suggested that the internal head NP is obligatorily phonologically deleted and the antecedent

of it is the external head. If indirect identity could be satisfied in the relationship of the internal and external head, this would have consequences on the account of Kennedy's puzzle. Consider Kennedy's example in (55a), repeated from (2). Since, being a country entails being something, an empty internal head is entailed by the external head in (56a). But, as (56b) illustrates, (56a) is predicted to be grammatical if a representation with an empty internal head was possible. However, entailment isn't the only requirement of indirect identity. If we assume that condition (50) carries over to directly to the case of NPs, direct identity of the semantic content of internal and external head is required in cases like (56a).

(56) a. *Polly visited every town in a country Eric did ⟨visit⟩.

b. *[every town in a [country, Op_y Eric $\underbrace{\text{visited } [y, \,]}_{\text{elided VP}}$]]

λx [Polly $\underbrace{\text{visited } [x, \text{ town}]}_{\text{antecedent}}$]

Now, consider the cases where NPRC and NPQR are from the same semantic field, but no subset relation holds between them. In (57), (38c) and (39c) from above are repeated. In contrast to the subset cases, in these cases NPRC and NPQR seem interchangeable as (58) aims to show ((58a) repeated from (37c)).

(57) a. ?John lives in a city that's close to a town Mary used to ⟨live in⟩.

b.??Jon ordered a cocktail that's more expensive than the beer Sue did ⟨order⟩

(58) a. ?John visited a city that's near a town Mary did ⟨visit⟩.

b. ?John visited a town that's near a city Mary did ⟨visit⟩

Notice that for the examples to be acceptable, a particular intonation is required. Even then, they still remain marginal for most speakers if compared to the cases where NPQR and NPRC are identical. On the required intonation, NPQR is stressed, while NPRC is unstressed. Though I haven't been able to verify whether the pitch on NPQR confirms this claim, I claim that NPQR bears a topic accent in the sense of Büring (1996, 1998) since this helps in explaining the relative acceptability of (57) and (58).

One of the functions of topic accents Büring discusses is that they signal the presence of an alternative question to the assertion made. One of Büring's (1995) examples is (59a), where the topic accent on *female* is phonetically realized as a falling pitch. With this intonation, (59a) indicates that the question *What did the male pop stars wear?* is still open. (59b) answers this open question, and in (59b) *male* cannot bear a topic accent. This is because, with the assertion of (59b), it is known what all the pop stars wear. In (59b) *male* can though optionally bear focus. (In (59) and the following, I indicate pitch accents with capital letters, semantic focus with an F-subscript, and semantic topic with a T-subscript. With Büring (1995), I assume that a pitch accent inside a semantic topic is phonetically realized as falling pitch, whereas a pitch accent inside a semantic focus is realized by a rise in pitch.)

(59) a. The [FEmale]$_T$ pop stars wore [KAFtans]$_F$.

b. The male pop stars wore [tuXEdos]$_F$.

The examples in (60) and (61) indicate that the implicit question raised in-

dicated by a topic accent can license destressing and deletion, if it's not ambiguous which implicit question raised is the licensing one. Hence, (60) and (61) are acceptable in a situation where only German and American beer or red and green gummibears are under consideration.

(60) a. $[\text{John}]_F$ bought $[\text{German}]_T$ beer. $[\text{Mary}]_F$ bought the American one.

b. $[\text{John}]_F$ eats only $[\text{red}]_T$ gummibears. $[\text{Mary}]_F$ must've eaten the green one.

(61) a. $[\text{John}]_F$ bought $[\text{German}]_T$ beer. As for the American one, $[\text{Mary}]_F$ did.

b. $[\text{John}]_F$ eats only $[\text{red}]_T$ gummibears. As for the green ones, $[\text{Mary}]_F$ does.

What I claim is that the examples of (37) to (39) where RCNP and RCQR are from the same semantic field satisfy the licensing condition on VP-ellipsis for the reason the examples in (61) do. Assume that (37c) has the focus structure in (62). The alternative needed to license VP-deletion is a question like *Who visited a town?*. Alhough the matter is far from clear, the topic accent might make this antecedent available. The requirement to be in the same semantic field, I believe, is a general requirement on alternative questions raised by topic accents.

(62) $[\text{John}]_F$ visited a $[\text{city}]_T$ that's near a town $[\text{Mary}]_F$ did $\langle$visit$\rangle$.

This concludes the discussion of the examples (37) to (39). I showed that a class of such examples, where NPRC denotes a subset of NPQR, is actually predicted by the assumption of Fox (1998a) that VP-deletion can be licensed indirectly if this is forced by the overt material surrounding the deletion site. This is part of a theory

where all the licensing conditions for VP-ellipsis are only sensitive to meaning. As for the second class of examples where NPRC and NPRC are from the same semantic field, I offered a suggestion of how to incorporate these cases into the theory of ellipsis. Both accounts relied on the assumption that the lexical material represented in the trace position contributes to the interpretation of the trace. In particular, for the examples above, the interpretation of an unbound trace [x, NP-part] could be paraphrased as an indefinite 'a NP-part'.

The remaining pages of this section contain a digression. The question it addresses is raised at the end of chapter 2 and concerns deletion of ACD-relatives in a trace position. I state this as the requirement in (63) (repeated from (97d)). The question raised above is whether this requirement involves look-ahead of the mechanism that deletes parts of a chain to the level where the licensing conditions for ellipsis apply, or whether chain deletion applies in an ACD configuration independently of whether this will license deletion or not. In chapter 2, it was impossible to distinguish these possibilities empirically because it seemed that all examples with an ACD configuration required deletion of the modifier in the trace position for the licensing of VP-ellipsis. But, at this point, it becomes possible to draw a distinction, and as I will show, the results lead me to conclude that no look-ahead is involved with (63).

(63) **ACD** If material inside a modifier is anaphorically related to the constituent surrounding an occurence of this modifier, this occurence of this modifier must be deleted.

Consider the following prediction of the idea that VP-deletion can be licensed indirectly via an entailment. In (64a), the head of the ACD-relative is introduced by an upward entailing quantifier. In this case, VP-ellipsis in the relative clause is predicted to be licensed even if the relative clause remains in its surface position. Consider the LF-representation in (64b). For the antecedent and elided VP as indicated, indirect identity is satisfied by the structure in (64b): Since *a* is an upward entailing quantifier, the clause (64a) entails that John read a book. But, this provides a suitable antecedent for the elided VP in (64b) if, as I argued above, an indefinite in the antecedent is sufficient to license deletion of a VP where a trace with the same NP-part corresponds to the indefinite.

(64) a. John read a book Mary did ⟨read⟩

b. John $\underbrace{\text{read [a book } \lambda x \text{ Mary did } \overbrace{\text{read } [x, \text{ book}]}^{\text{elided VP}}}_{\text{antecedent}}$

This prediction gives us way to test whether the condition in (63) is looking ahead to see whether ACD can be licensed without deletion, or whether ACD applies whenever the formal configuration of ACD arises. Namely, if (63) is looking ahead the configuration in (64b) should arise in the licensing of ACD with indefinites. On the other hand, if (63) doesn't look ahead, it should apply in (64b), and therefore ACD should require QR for its resolution even with upward entailing quantifiers. How can we test this prediction? The first test that comes to mind is to use quantifier scope as a diagnostic of whether QR applied. As we'll see, the result seems to favor the no look-ahead position, but is ultimately not decisive. A second diagnostic is Condition

C obviation effects discussed in section 2.1. Here again, the result argues for the no look-ahead position and, in this case, I find the argument convincing.

First, the quantifier scope test. As Sag (1976:72-74) and Larson and May (1990:112-15) observe, ACD forces the DP that the ACD-relative is part of to take scope outside of the antecedent of the elided VP. This follows from—and to be more precise, strongly supports—the assumption that covert movement to a position outside of the antecedent is required for the resolution of ACD. This is shown in (65), where only the case where the elided VP is interpreted as indicated is relevant. The fact observed by Sag is that then the quantifier *every* cannot take scope below *want*. The explanation of this fact is that QR to a position outside of the antecedents is required which forces an LF-representation like (65b). In (65b), however, *every* is outside of the c-domain of *want*.

(65) a. Betsy's father wants her to read everything her boss wants ⟨her to read⟩ (everything ≫ want, *want ≫ everything) (Sag 1976:(1.3.38))

b. $\Big[$everything λy her boss wants her to read $[y,$ thing$]\Big]$ λx Betsy's father wants her to read

Sag's example (65a), as well as the example of Larson and May (1990), demonstrate the correlation between scope and ACD using the universal quantifier *every*. The question at hand is whether upward entailing quantifier behave differently. The examples in (66), based on suggestions by Irene Heim (p.c.) and Danny Fox (p.c.), show that indefinites behave like universal quantifiers. Both involve the scope of the negative

polarity item *anything* which is upward entailing in its NPI-meaning. In this meaning it must occur in the scope of an NPI-licensing expression, which in (66a) is negation and in (66b) the verb *refuse*. In both examples, consider only an interpretation where the NPI-licenser is part of the antecedent of the deletion in ACD, as indicated. If the correlation between scope and ACD wouldn't hold for upward entailing quantifiers, the VP-deletions indicated could be licensed without moving the NPI to a position outside the scope of its licenser. However, this prediction seems to be wrong—the interpretations of the elided VPs indicated seem unavailable. This shows that with upward entailing quantifier, Sag's observation also holds: the ACD-containing DP must take scope outside of the antecedent.

(66) a. *John plans never to be anywhere Mary did ⟨plan never to be⟩ (*any ≫ not, *not ≫ any)

b.??John is refusing to read anything Mary is ⟨refusing to read⟩ (??any ≫ refuse, *refuse ≫ any)

The result in (66) seems to favor the idea that the deletion in ACD involves no look-ahead. In fact, though, the look-ahead view probably predicts the lack of a narrow scope reading for the examples in (66) in the following way. Consider the LF-representation for (66a) in (67). While it's true that "John is in a place that Mary is" entail that "John is in a place", this entailment is not sufficient to license indirect identity in (67). The entailment required for (67) would be from "John plans never to be anywhere Mary plans never to be" to "John plans never to be anywhere",

which obviously doesn't hold. Therefore, (67) doesn't satisfy indirect identity.

(67) John $\underbrace{\text{plans never to be anywhere } \lambda x \text{ Mary did } \overbrace{\text{plan never to be } [x, \text{ place}]}^{\text{elided VP}}}_{\text{antecedent}}$

The failure of (67) to license indirect identity is not an accident, but inherent to the logic of the scope argument: An example without this flaw would be one where an upward entailing quantifier with an ACD-relative, entails the wide scope reading while taking narrow scope. As Abusch (1994) argues, it is for pragmatic reasons impossible to test whether a certain reading is present, if it's entailed by another reading whose existence is established. Hence, I believe the argument based on scope argument doesn't decide whether there's look-ahead in ACD-resolution.

The second test for the look-ahead question is based the obviation of Condition C discussed in section 2.1. Recall that in an example like (68) (repeated from (7a) on page 34) Condition C was obviated by ACD, because ACD requires deletion of the ACD-relative clause in the position of the QR-trace, as shown by the LF-representation in (68b).

(68) a. You introduced him$_i$ to everyone John$_i$ wanted you to ⟨introduce him$_i$ to⟩

b. $\Big[$everyone $[\lambda y$ John$_i$ wanted you to $\underbrace{\text{introduce him}_i \text{ to } [y]}_{\text{elided VP}}]\Big]$

λx you $\underbrace{\text{introduced him}_i \text{ to } [x]}_{\text{antecedent}}$.

Since indirect identity could license ACD without movement with upward entailing quantifiers, the look-ahead view would predict that in these cases we shouldn't find Condition C obviation. The contrast in (69), which seems as strong as in the case

of the universal quantifier in (68a), falsifies this prediction. This argues that the LF-representation in (70), though it would satisfy indirect identity, is ruled out by a formal requirement that rules out any configuration where an elided VP occurs inside of its antecedent. Therefore, I conclude that the no look-ahead position is correct.

(69) a. ?John introduced her$_i$ to a man Mary$_i$ wanted him to ⟨introduce her to⟩.

b. *John introduced her$_i$ to a man Mary$_i$ wanted Bill to like.

(70) John $\underbrace{\text{introduced her}_i \text{ to a man } \lambda x \text{ Mary}_i \text{ wanted him to } \overbrace{\text{introduce her to } [x, \text{ man}]}^{\text{elided VP}}}_{\text{antecedent}}$

3.3 *Wh*-Traces and Focus in Chains

This sections starts out with a number of apparent problems for the assumption that traces have content that matters for the licensing of VP-ellipsis. The goal of the section is to show that the right understanding of how focus works in chains, which is actually mostly drawn from the literature, yields a natural solution to these problems, and in fact provides new support for the main claim of this chapter.

The problems are examples like (71). Apparently, in constructions other than ACD, the lexical content of a trace, if it's there, doesn't block VP-ellipsis (Evans 1988, Jacobson 1992)[11]

[11]Sag (1976:63–67) and Williams (1977:130–31) claim based on examples like those in (i) that *wh*-extraction from an elided VP is impossible. The examples in the text falsify this general claim and the examples in (i) are probably ruled out for irrelevant reasons: In (ia), since the verb moves to Comp, the elided VP is in the complement position of an empty head, which is generally impossible (Lobeck 1992). (ib) and (ic), as Fiengo and May (1994:244) suggest, indicate a preference to delete as much material as possible once material is deleted (See also 2 on Lappin 1984).

(i) a. What did Harry take a picture of?
*What did Bill? (Sag 1976:(1.3.18))

(71) a. I know which cities Mary visited, but I have no idea which lakes she did ⟨visit⟩.

b. The cities Mary visited are near the lakes Bill did ⟨visit⟩.

Let me briefly sketch the solution before I spell it out in detail below. The first point to note is that the examples in (71) require two different solutions. This is shown by the surprising contrast in (72): (72a) shows that a VP containing a relative clause internal trace can license deletion of a VP containing a trace of *wh*-movement with different lexical content. (72b) shows that it's impossible to license deletion in the other direction—a VP containing the trace of *wh*-movement cannot be the antecedent for an elided VP that contains a relative clause internal trace with different lexical content.

(72) a. I know the cities Mary visited, but I would like to know which lakes she did ⟨visit⟩.

b. *I know which cities Mary visited, but I would like to know the lakes she did ⟨visit⟩

In section 2.4, I discussed another difference between relative clauses and questions, namely with respect to Condition C as illustrated in (73)(repeated from (49) on page 62). The account presented in section 2.4 essentially claimed that in a question, the

b. *John who Bill saw and who Bob did, too. (Williams 1977:(93))

c. ??We finally got in touch with John, who my brother Al tried to visit, but who he couldn't ⟨visit⟩ (Sag 1976:(1.3.22))

fronted material is directly related to the trace position. The relationship between the head of a relative clause and the relative clause internal trace position, on the other hand, is less direct, and therefore allows the minor change in the lexical content of the NP-part that's represented in the trace position, needed to circumvent Condition C.

(73) a. Which is the picture of John$_i$ that he$_i$ likes?

b. *Which picture of John$_i$ does he$_i$ like?

My solution for (71a) makes use of this distinction between relative clauses and questions. I claim that material that's directly related to the dislocated material, as in a question, is essentially overt material, and therefore isn't subject to the identity requirement on elided material. Indirectly related material, on the other hand, is subject to the requirement. This is the essence of my solution for (71a) presented in more detail below.

The solution of (71a) just sketched doesn't carry over to (71b), because here the relative clause internal traces depend only indirectly on the external heads, which are different. Because of the contrast in (72), this is a desirable result. The solution for (71b) I propose draws an analogy between them and sloppy readings. Consider the sloppy reading in example (74): While I have so far assumed that the semantic contribution of a variable is ignored by the identity condition, this is usually assumed to be incorrect—in fact, I argue that it's incorrect in section 4.1. But, then the two VPs are different in meaning, namely *bribed John* and *bribed Bill*. This problem looks

similar to the problem in (71b), since there as well the lexical content of the trace depends on a phrase outside of the elided VP.

(74) John$_i$ admitted that Mary had bribed him$_i$.

Bill$_j$ admitted that she had ⟨bribed him$_j$⟩, too. (Hardt 1992:(27))

I elaborate this analogy by extending Rooth's (1992b) account for sloppy readings, which I argue for in section 4.1, to the case of (71b). In particular, I show that the explanation of the parallel dependencies requirement (45) also accounts for the difference between the well-formed(71b) and the ill-formed (72b), as well as the ACD-cases like (2): In (71b) the material the traces are related to is the external head of the relative clause for both, the elided VP and the antecedent. In the ill-formed examples, this isn't the case, and therefore (71b) seems to be satisfactorily explained.

The remainder of this section is divided in three subsections, each of which considers one possible way to circumvent the identity requirement of VP-ellipsis. First, I consider pseudo-gapping but only to conclude that it doesn't account for cases like (71a) and (71b). Second, I consider Focus, and will argue that it provides the solution for (71a), but not for (71b). Finally, I consider Domain Extension and argue that it solves (71b).

3.3.1 Pseudogapping and Traces

Pseudogapping, illustrated in (75), as a potential way to circumvent the identity requirement of VP-ellipsis, was brought to my attention by Fox and Nissenbaum

(1998). In pseudogapping, a VP is elided, except for the object which is pronounced and can be different from the antecedent. In fact, the object must be different.

(75) a. While some visited cities, others did ⟨visit⟩ lakes.

b. While some people advised Mary to visit cities, others did ⟨advise Mary to visit⟩ lakes.

It seems possible that most examples of apparent VP-ellipsis with a trace in object position are really instances of pseudogapping, where the trace isn't part of the elided material (Cormack 1984, Jacobson 1992). Consider, for example, the LF-representation for (71a) in (76). If pseudogapping only imposes an identity requirement on the elided material, and the trace in object position can be not part of the elided material, (76) is predicted to satisfy the identity requirement. This is the right prediction for (71a), but the account overgenerates massively: It predicts that all examples of VP-deletion with an object trace should allow an analysis like (76), unless the trace is too deeply embedded in the VP. This would incorrectly predict all examples like Kennedy's puzzle (2) to be good.

(76) I know which cities λx Mary $\underbrace{\text{visited}}_{\text{antecedent}}$ $[x$, cities], but I have no idea which lakes λy she did $\underbrace{\text{visit}}_{\text{elided VP-part}}$ $[y$, lakes].

There is reason to doubt that pseudogapping involves deletion of parts of a VP. For one, deletion on this account would be an unusual operation since it does target non-constituents. For example, but the VP-parts hypothetically deleted in (75b) don't

form a constituent. Based on this and additional arguments, all recent analyses of pseudogapping (Jayaseelan 1990, Lasnik 1995, Johnson 1996) conclude that pseudo-gapping is actually VP-ellipsis preceeded by movement of the object to a position outside of VP. This means that, even in pseudo-gapping, the identity condition always applies to a trace.

The remaining question is whether the lexical content of the antecedent is represented in this trace or not. Or in other words, whether the object movement in pseudogapping is A- or A-bar-movement. If it is A-movement, as shown in section 2.3 and at the end of section 3.1, the trace usually doesn't contain material of the antecedent, and pseudogapping would help to circumvent the identity requirement. If it's A-bar-movement, the content of the trace in pseudo-gapping would not be different from content of the trace of *wh*-movement.

The recent analysis of pseudogapping by Jayaseelan, Lasnik and Johnson disagree about the type of movement involved in pseudogapping. Jayaseelan (1990) proposes heavy NP shift, Lasnik (1995) advocates object shift analogous to what is found in Scandinavian languages, and Johnson (1996) opts for a movement analogous to Dutch scrambling. At least Lasnik's and Johnson's point in the direction of A-movement; object shift is always A-movement and scrambling in Dutch is A-movement in many cases (Déprez 1990). However, the objective of all three papers is to account for the locality restrictions of the movement and restrictions on the type of object that can occur. Therefore the argument for A-movement is only indirect: the movement in pseudogapping shows restrictions reminiscent of restrictions on A-movement some other languages exhibits.

We can test for whether A- or A-bar-movement is involved in pseudo-gapping by considering Condition C obviation. As seen in section 2.3, only A-movement obviates Condition C regardless of where in the moving phrase the R-expression occurs. The examples in (77) show that pseudo-gapping must involve A-bar-movement by the Condition C test. Neither in (77a) nor (77b) is it possible for *her* to corefer with an R-expression that's part of the NP-part of the moving DP.

(77) a. *I gave her$_i$ a book and you did ⟨give her$_i$⟩ a picture of Mary$_i$.

b. *While some told her$_i$ to paint a portrait of John, others did ⟨tell her$_i$ to paint⟩ a picture of Sue$_i$.

The contrast in (78) is parallel to the well-known argument/adjunct contrast of Freidin (1986) and Lebeaux (1988) discussed in section 2.2 above. If the R-expression occurs in a modifier adjoined to the object-DP that moves in pseudogapping as in (78b), it doesn't cause a Condition C effect.

(78) a. *While some believed him$_i$ everything, others did ⟨believe him$_i$⟩ only the story that John$_i$ had met aliens.

b. While some believed him$_i$ everything, others did ⟨believe him$_i$⟩ only the story that John$_i$ had evidence for.

Because of (78) and (79), I conclude that the object in pseudo-gapping undergoes A-bar movement, and leaving a trace with the same lexical content as the trace of *wh*-movement. Hence, pseudo-gapping can never obviate the identity requirement that's

imposed on the lexical content of A-bar traces and isn't involved in the explanation of (71a).

3.3.2 Focus and *Wh*-Traces

Focus is another possibility to escape the semantic identity requirement. This is illustrated in (79) using the identity requirement of destressing. Recall that (79a) (repeated from (41b)) is ambiguous with respect to whether John produces a picture of the bicycle or puts paint on the bicycle, but the interpretation of the destressed VP in the second conjunct must correspond to that of the first conjunct. This was used to show that the destressed VP has to satisfy the identity of meaning requirement, as well. Now consider (79b): it exhibits a similar requirement of sameness of interpretation, but the object *car* cannot be subject to this requirement. Since *car* must be focussed in (79b), this argues that focus is required for material in the scope of an identity requirement that isn't identical to the antecedent. In a sense, focus must make material invisible to the identity of meaning requirement.

(79) a. John painted the bicycle and Mary *painted the bicycle*

b. John painted the bicycle and Mary *painted the* $[\text{CAR}]_F$

Examples like (79b) show that the identity requirement is only imposed on the meaning of the non-focussed parts of a destressed VP. Following Jackendoff (1972), Rooth (1985), and Kratzer (1991), I refer to the meaning of the non-focussed parts of a phrase as the *presuppositional skeleton*. It isn't always intuitively obvious what the

meaning of the non-focussed parts is; for example, if the non-focussed parts aren't a constituent. To make the notion *presuppositional skeleton* more precise, I adopt (with one minor difference) a formalization argued for by Kratzer (1991), which was inspired by Jackendoff (1972) and Rooth (1985:12). Kratzer defines a function that assigns to a phrase-marker its presuppositional skeleton, which I use the notation $[\![—]\!]$ for. Informally, the value of $[\![\text{XP}]\!]$ is the meaning of an XP′ that is derived from XP by replacing all focussed subconstituent of XP with designated variables of the corresponding semantic type. Examples like (80a) with multiple foci, argue that different focussed constituents must correspond to different variables in the presuppositional skeleton. Otherwise, interpretation (80c), would be incorrectly predicted to be possible for (80a).

(80) a. At the party, John only introduce $[\text{MAry}]_F{}^1$ to $[\text{GRANDma}]_F{}^2$,

b. 'For any x_1 different from 'Mary', and any x_2 different from 'Grandma', John didn't introduce x_1 to x_2.

c. *'For any x_1 different from 'Mary', and any x_1 different from 'Grandma', John didn't introduce x_1 to x_1.

I indicate the variables a focussed constituent translates as by a superscripting it to the focussed constituent in the syntactic representation. This is exemplified by (80a). Assuming a similar notation, Kratzer defines $[\![—]\!]$ by recursion over the syntactic structure, where G is the assignment function for the focus variables:

(81) a. $[\![[\mathrm{X}]_F{}^n]\!]^G = \mathrm{G(n)}$

b. $[\![\mathrm{X}_0]\!] = \{[\![\mathrm{X}_0]\!]\}$, if X_0 is a terminal node

c. otherwise $[\![[\mathrm{X\ Y}]]\!] = \mathcal{C}([\![\mathrm{X}]\!]^G, [\![\mathrm{Y}]\!]^G)$, where $\mathcal{C}$ represents the function that assigns to $[\![\mathrm{X}']\!]$ and $[\![\mathrm{Y}']\!]$ the semantic value $[\![\mathrm{X}'\ \mathrm{Y}']\!]$ for any X' of the same semantic type as X and any Y' of the same semantic type of Y

Assuming Kratzer's definition of presuppositional skeleton, the identity of meaning requirement can be restated as a relationship of the meaning of the antecedent to the presuppositional skeleton to take the role of focus into account. Namely, the antecedent must be identical to the value of the presuppositional skeleton under G for at least one choice of assignment function G.

(82) There is an assignment G such that $[\![\text{antecedent}]\!] = [\![\mathrm{VP}]\!]^G$

For (80), (82) requires that the meaning of the antecedent VP, *painted the bicycle*, must be an element of the presuppositional skeleton given in (81). For this to be the case, the interpretation of *paint* in the antecedent and the partially destressed VP in (80) must be identical.

The example in (83a) from Fox (1995a) is another case where focus marks material that escapes the identity requirement. This case involves VP-ellipsis. (83a) allows an interpretation where the subject in both conjuncts takes scope below *seem*. As we see in (83b), on this interpretation, the elided VP doesn't seem to be identical to its antecedent, unless we exempt focussed material from the identity requirement.

Notice that the example in (83) provides an additional argument against the identity of form requirement discussed in the previous section: If there was an identity of form requirement on elided VP, it would have to apply at the level of LF to allow ACD. But, to allow (83a), it would need to be sensitive to focus. Therefore, an identity of form requirement would have to redundantly replicate the identity of meaning requirement ellipsis and destressing have in common.

(83) a. An American athlete seemed to Bill to have won a Gold Medal, and a [RUSsian RUNner]$_F$ did to. (Fox 1995a)

b. $\underbrace{\text{seemed to Bill to an American athlete have won a Gold Medal}}_{\text{antecedent}}$ and
$\underbrace{\text{seemed to Bill to a [Russian runner]}_F \text{ have won a Gold medal}}_{\text{elided VP}}$

Focus is relevant for the cases we're interested in, if traces or their lexical content can be focussed. Clearly traces cannot bear the pitch accent that usually indicates focus phonetically. But, it has been argued by Selkirk (1995) and references therein that traces can inherit the F-marking of their antecedents: *F-marking of a constituent licenses the F-marking of its trace* (Selkirk 1995:559). One of Selkirk's (1995) arguments based on the work of Bresnan (1971, 1972) starts with the observation that while usually, as we saw in (79b) above, material that is not identical in meaning to preceding material must be focus marked and receive a pitch accent, this doesn't hold for the verb in case the object is also focussed. *Reviewed* in (84a) is new information, but doesn't need to bear a pitch accent. For comparison, when the object in (84b) isn't focussed, pitch accent on the verb is required.

(84) a. Bill read the article and Helen [reviewed]$_F$ [the BOOK]$_F$.

b. Bill read the article and Helen [reVIEWed]$_F$ the article.

From (84), Selkirk (1995) concludes that an F-marked verb doesn't need to receive pitch accent if its complement is F-marked. Based on this generalization, she argues that in (85b) the object trace must be F-marked. Consider (85b) in the context of (85a). Again, both the object and the verb must be F-marked. However, a pitch accent on the fronted object suffices to phonetically realize this F-marking. This is explained by the same phonological principle that applied in (84a), if the trace of the fronted phrase is F-marked. Otherwise though, (85b) not only requires a new phonological principle for pitch placement, but one that refers to a syntactic notion such as the antecedent of a trace. Therefore, (85b) argues that the trace in (85b) can obtain F-marking from its antecedent.

(85) a. Bill read an article, but ...

b. [Which [BOOK]$_F$]$_i$ did Helen [review]$_F$ [t_i]$_F$? (Selkirk 1995:(24))

Consider now example (71a) again, which is repeated in (86a). If Selkirk (1995) is right, the trace of a *wh*-moved phrase is F-marked in case the moved phrase is F-marked. On a copy theory of movement this can be restated as follows: Copies of F-marked phrases are F-marked. Assuming this for (86a), yields the LF-representation in (86b). To satisfy the focus-sensitive identity requirement in (82), there must be a replacement of the focussed material inside the elided VP, such that the antecedent VP

and the elided VP mean the same. In (86b), the focus-sensitive identity requirement is obviously satisfied because it's possible to replace the *lakes*, the lexical content of the trace in the elided VP, with *cities* to achieve identity.[12]

(86) a. I know which cities Mary visited, but I have no idea which $[\text{lakes}]_F$ she did ⟨visit⟩.

b. I know which cities λx Mary $\underbrace{\text{visited } [x, \text{ cities}]}_{\text{antecedent}}$,

but I have no idea which $[\text{lakes}]_F$ λy she did $\underbrace{\text{visit } [y, [\text{lakes}]_F]}_{\text{elided VP}}$

This concludes the account of (71a). As shown the standard claim that focus in the head of a chain is also represented in the trace position together with the equally standard claim that focussed material isn't seen by the identity requirement on elided material, yields a straightforward explanation of this case. The next question is under which circumstances a focus that is phonetically expressed on a one constituent is also represented as F-marking in a dependent position. For the discussion, I refer to this

[12]The account makes a prediction for examples like (ia) (repeated from (20a)). Since the elided VP contains a trace of *wh*-movement in (ia), focus percolation should from the *wh*-phrase should be possible, and the focus structure in (ib) should be result. But, then (ia) should be acceptable with focus on *town*-phrase, since *lake* is a focus-alternative to *town* that would satisfy direct parallelism. This prediction seems factually incorrect.

(i) a. *Do you know which town near a lake Mary visited John did ⟨visit⟩?

b. which $[\text{town}]_F$ near a lake λy Mary visited $[y, \text{lake}]]$ λx John did visit $[x, [\text{town}]_F]$

The structure of (iib) resembles that of the A-movement examples like (33) discussed at the end of section 3.1 and in section 4.1, where the ellipsis site also contained a trace of overt movement. However, in contrast to the A-movement cases, the addition of focus particles doesn't seem to lead to an improvement of (ia) as (ii) attests. While I still hope that a better understanding of what are possible focus structures will provide an explanation for (ia), at this point, I have to leave the matter open.

(ii) *Do you know which town near a lake that Mary visited John did instead.

situation as *percolation of the F-marking* from one position to another. I want to argue that F-marking can only percolate if the two positions are related via movement.[13]

One piece of support for the claim that F-marking can only percolate to a dependent within a chain comes from the sluicing paradigm in (87) and (88).[14] I assume an analysis of sluicing as IP-ellipsis (Ross 1969b, Chung *et al.* 1995). It's quite well known that the fronted *wh*-phrase in sluicing can be related to the trace position in the elided IP either via movement, or via a different process that doesn't create a syntactic chain. For example, Chung *et al.* (1995:279) distinguish between *sprouting* (involving a chain) and *sluicing* (not involving a chain). I cannot discuss the process invoked by sluicing (in the narrow sense of Chung *et al.* 1995) in detail (see also Reinhart 1994). Two properties of sluicing matter: that it isn't sensitive to syntactic islands and that it doesn't involve formation of a syntactic chain. Therefore, in the examples in (87), both sprouting and sluicing are possible, but in (88) only sluicing is possible because an island intervenes between the antecedent and the trace in the elided material, as can be seen from the paraphrases in both (88a) and (88b). Then the facts in (87) and (88) yield the following conclusion: Sprouting, as in (87a) is possible even if the NP-part of the *wh*-phrase is different from the NP-part of the corresponding indefinite in the antecedent. Sluicing, however, as shown, in (88), requires that the NP-part of the *wh*-phrase be identical to the NP-part of the corresponding indefinite in the antecedent.

[13]As is expected from the absence of lexical content in the trace position, Focus in A-chains cannot percolate to the trace position. Hence, the only way for a F-marked DP in an A-chain can contribute an F-mark in the trace position is by scope reconstruction (see section 6.2). Diesing (1992) and Selkirk (1995) argue based on data from Berman and Szamosi (1972) that this prediction is correct.

[14]Examples like (87) came to my attention during a discussion with Chris Kennedy.

(87) a. An astronomer needs to find a lot of new supernovae for her Ph.D., but I don't know how many galaxies ⟨an astronomer needs to a find for her Ph.D.⟩

b. An astronomer needs to find a lot of new supernovae for her Ph.D., but I don't know exactly how many ⟨new supernovae an astronomer need to a find for her Ph.D.⟩

(88) a. *An astronomer needs to find a quadrant that contains a lot of new supernovae for her Ph.D., but I don't know how many galaxies ⟨an astronomer needs to find a quadrant that contains for her Ph.D.⟩

b. An astronomer needs to find a quadrant that contains a lot of new supernovae for her Ph.D., but I don't know exactly how many ⟨new supernovae an astronomer needs to find a quadrant that contains for her Ph.D.⟩

For the account of sluicing, I adopt the assumption that an indefinite in the antecedent can correspond to a trace with the same NP-part in the elided IP, which is in the spirit of Reinhart (1994). Then the facts in (87) and (88) follow directly from the assumption that focus can only percolate from the antecedent to its dependent if the two are linked by a syntactic chain. Notice that in (87a), the NP-part of the *wh*-phrase *how many galaxies* must be focussed. Since in (87a) the *wh*-phrase can be linked to the trace position by a syntactic chain, the focus of the *wh*-phrase can percolate to the trace position, as shown in (89a). Therefore, the elided IP in (89a) is identical modulo its focussed parts to the antecedent. In (88a), on the other hand, no syntactic

chain can be formed between the antecedent and the trace position. I claim that, as shown in (89b), F-marking cannot percolate to the NP-part of the trace because it's not linked to its focussed antecedent by a chain. But, if the NP-part of the trace isn't F-marked in (89b), it doesn't satisfy the identity requirement. Therefore, (88a) is ill-formed on the assumption that F-marking can only percolate in a chain.

(89) a. how many $[\text{galaxies}]_F$ λx an astronomer needs to find $[x, [\text{galaxies}]_F]$

b. how many $[\text{galaxies}]_F$ λx an astronomer needs to find a quadrant that contains $[x$, galaxies$]$

In relative clauses, the question whether F-marking can percolate from the external head to the relative clause internal trace is much harder to investigate. Recall from section 2.4 that there are two possible structures for a relative clause, the matching and the raising structure, and that these are quite hard to distinguish based on their interpretation. The main discussion of focus percolation in relative clauses I'm aware of is found in (Bresnan 1971) and the replies to Bresnan's paper (Lakoff 1972, Berman and Szamosi 1972, and Bresnan 1972). Especially, the discussion of a correlation between a difference in interpretation and stress placement in Bresnan (1972:337-40) is quite interesting for our current purposes. The discussion is based on the notion of *normal stress* which unfortunately isn't made very precise in the paper itself. For the following, I assume that normal stress can be characterized as bearing F-marking on the rightmost 'most embedded' constituent that can be F-marked (Höhle 1979 1982, Cinque 1993, Schwarzschild 1998, Zubizaretta 1998). This assump-

tion predicts a difference in normal stress between the matching and raising analysis, in examples like (90), where an object relative clause is attached to the object of the main verb. On the raising analysis, normal stress should require pitch accent on the relative clause head so that the trace in the object position of the relative clause is F-marked. On the matching analysis, however, it should be impossible to F-mark the trace in the relative clause, and therefore the verb in the relative clause should be F-marked.

(90) a. I gave John the [BOOKs]$_F$ he wanted. (Bresnan 1972:(43))

'I gave John the number of books that he wanted.'

'Of books, I gave John the ones he wanted.'

b. I gave John the books he [WANted]$_F$ (Bresnan 1972:(44))

'Of the books, I gave John the ones he wanted.'

*'I gave the John the number of books that he wanted.'

In her discussion of (90), Bresnan rejects the claim of Lakoff (1972) that the normal stress in an example with an object attached relative like in (90) can be freely assigned to either the head of the relative clause as in (90a) or the verb inside the relative clause as in (90b). Instead, she argues that the apparent optionality correlates with a difference in interpretation. The two interpretations Bresnan characterizes are one where *where the relative clause applies to the entire head* in (90a) and *a concealed partitive* interpretation for (90b). The prediction seems therefore at least partially borne out as (90b) doesn't allow an amount interpretation. I find it impossible to

assess whether (90a) has only an amount or kind interpretation, or whether it also has a restrictive interpretation.

Accent placement in (90) might also be affected by the implicature that if John was given books, he probably wanted them. For the example (91), most of my informants agree that (91b) prefers an interpretation where *those* is used to refer to the same tokens of chips as those that used to be ours. (91a), on the other hand, could be used when the chips are different tokens, but the amount of chips is the amount of chips that we lost. Again, (91) confirms a part of the prediction, while it leaves it open whether pitch accent on the head noun can be the 'normal' stress for a matching analysis of the relative clause.

(91) a. Those are the $[\text{CHIPs}]_F$ we lost

b. Those are the chips we $[\text{LOST}]_F$

One argument for the prediction concerning examples with pitch accent on the head noun comes from the example in (92), where a raising analysis is ruled out by Condition C. As predicted, the pitch placement in (92b) seems to be preferred in (92).

(92) a.$^{??}$Those are the $[\text{AUNTs of Mary}_i]_F$ she_i likes

b. Those are the aunts of Mary_i she_i $[\text{LIKes}]_F$

With non-finite relatives in (93) the intuitions are sharper. In a neutral context, Hackl and Nissenbaum (1998) argue that pitch accent on the head noun, as in (93a), forces an interpretation where the relative clause has possibility modal force.

Pitch accent on the verb, on the other hand, forces an interpretation paraphrasable only with a necessity modal. Hackl and Nissenbaum (1998) present arguments based on Condition C that (93a) has a raising analysis, while (93b) has a matching analysis. Therefore the prediction mentioned above is fully confirmed by (93).

(93) a. Sabine came up with many [PROblems]$_F$ for us to work on
'Sabine came up with many problems we could work on.'
*'Sabine came up with many problems we should work on.'

b. Sabine came up with many problems for us to [WORK]$_F$ on
*'Sabine came up with many problems we could work on.'
'Sabine came up with many problems we should work on.'

Based on these arguments, I conclude that F-marking can only percolate to a dependent within a chain. This predicts that focus on a fronted phrase can obviate the effect of the identity condition on a trace position, only if the trace is related to the antecedent directly via movement as in the case of *wh*-movement, but not in the case of matching relatives. Though I regard the arguments based on normal stress above as tentative, this conclusion must be correct: Otherwise, all the examples of Kennedy (1994) like (2) would be predicted to be acceptable with the right pitch placement, which isn't the case as shown by (94a) and (95a).[15]

[15]One person in the audience at the SALT 8 conference at MIT reported a general improvement of Kennedy's examples if the head of the relative clause is stressed. However, none of my informants share this intuition, and the person in question was not a native speaker of English. In fact, as discussed in 1 and also observed with the paradigm in (39), destressing the head of the relative clause is required even in the good examples.

In contrast to matching relative clauses, raising relatives are predicted to pattern with *wh*-questions, because here the relationship of the external head to the relative clause internal trace position was argued to be created by movement (section 2.4). The contrasts in (94) and (95) seem to confirm this prediction for the kind reading of raising relatives. In (94a) (repeated from (10)) and (95a) (repeated from (16)), placing pitch accent on the head of the ACD-relative *lakes* doesn't improve the example. But in (94b) and (95b), where a kind reading is possible, accenting the head of the ACD-relative improves the example.

(94) a. *John visited towns that are near the LAKes Mary did ⟨visit⟩.

b. John visits towns that are much nicer than the LAKes Mary does ⟨visit⟩.

(95) a. *Jon ordered a drink that's more expensive than the DISH Martin did ⟨order⟩.

b. John orders drinks that are more expensive than the DISHes Martin does ⟨order⟩

To see that the contrast in (94) is predicted, consider the LF-representations in (96). In (96a), the lexical content of the trace in the elided VP is not F-marked, and therefore blocks identity between the elided VP and the potential antecedent. In (96b), on the other hand, the lexical content of the trace in the elided VP is F-marked, and therefore irrelevant for the identity condition (82). Therefore, the antecedent and the elided VP are considered identical in (96b).

(96) a. *[towns that are near the [lakes]$_F$ λy Mary did $\underbrace{\text{visit } [y, \text{ lakes}]}_{\text{elided VP}}$] λx John $\underbrace{\text{visited } [x, \text{ towns}]}_{\neq \text{ antecedent}}$

b. [towns that are much nicer than the λy Mary did $\underbrace{\text{visit } [y, [\text{lakes}]_F]}_{\text{elided VP}}$] λx John $\underbrace{\text{visited } [x]}_{\text{antecedent}}$

In matching relative clauses, focus percolation is predicted to be possible for overt material that is part of the internal head. This is pied-piped material surrounding the *wh*-word in a matching relative clause. Jacobson (1998a) points out examples like those in (97) which confirm this prediction.

(97) a. Mary visited every country the [EMbassy]$_F$ of which [BILL]$_F$ did ⟨visit⟩.

b. John greeted every boy whose [MOther] [Sue]$_F$ did ⟨like⟩

c. Sue voted for every candidate the [FAther] of whom [BILL]$_F$ had ⟨voted for⟩ (Jacobson 1998a:(15))

Consider the LF-representation of (97a) in (98). The pied-piped material is represented in the relative clause internal trace position, but is focussed. Therefore, the identity requirement is satisfied in (98).

(98) [every country λy Bill $\underbrace{\text{visited the [embassy]}_F \text{ of } [y, \text{ country}]}_{\text{elided VP}}$]

λx Mary $\underbrace{\text{visited } [x, \text{ country}]}_{\text{antecedent}}$

Jacobson (1998a) makes a very interesting observation about (97) that is not relevant to the current discussion, but strongly supports Rooth's (1992b) view of the parallel dependencies requirement and indirect identity. Notice that the parallel dependencies requirement seems to be violated in (98) because the relationship of y in the elided VP to its antecedent is different from that of x to its antecedent. Jacobson (1998a) proposes therefore that (98) requires indirect identity, and the entailment of the sentence that contains the antecedent in (98) that satisfies the parallel dependencies requirement is one like that paraphrased in (99), where *territory of* is inserted to make the dependency parallel to the y-dependency in the elided VP.

(99) λx Mary visited the territory of [x, country]

Jacobson points out that the contrasts in (100) and (101) corroborate the view that (97) requires indirect identity. The examples (100a) and (101a) are very similar to those in (97), except that the DP that moves covertly for ACD-resolution in (97) is moved overtly by topicalization in (100) and (101). Because the elided VP precedes its antecedent linearly, (100a) and (101a) are slightly harder to parse than (97), but both acceptable. This is expected because the account just given for (97) carries over to (100a) and (101a). The examples (100b) and (101b), however, are ill-formed.

(100) a. ?Every country the embassy of which Bill did ⟨visit⟩, Mary visited.

b. *Every country the embassy of which Bill visited, Mary did ⟨visit⟩.

(101) a. ?Every candidate the father of whom Bill had ⟨voted for⟩, Sue voted for.

b. *Every candidate the father of whom Bill had voted for, Sue had ⟨voted for⟩

(Jacobson 1998a:(18))

Jacobson's account correctly predicts (100b) and (101b) to be ill-formed. Consider the LF-representation of (100b) in (102). The antecedent isn't directly identical to the elided VP in (102). The entailment that would be required to license indirect identity would be one from "Bill visited the embassy of the country" to "Bill visited the country". Since this entailment isn't valid, no antecedent for the elided VP is available in (102).

(102) *[every country λy $\underbrace{\text{Bill visited the [embassy]}_F \text{ of [}y\text{, country]}}_{\neq \text{ antecedent}}$]

λx Mary $\underbrace{\text{visited [}x\text{, country]}}_{\text{elided VP}}$

In this way, the asymmetry between deletion of the VP with the complex trace and deletion of the VP with the simpler trace seen in (101) and (102) follows directly from the fact that an entailment from a simple DP like *country* to a more complex DP *territory of the country* seems always possible, whereas an entailment in the other direction is usually impossible. Notice that (103), where the entailment from the complex DP to the simpler one is licit, is better than (100b), as predicted. Therefore, Jacobson (1998a) concludes that her paradigm involving pied-piping in the relative clause lends strong support to the claim the VP-ellipsis can be licensed by indirect identity.

(103) ?Every country the capital of which Bill visited, Mary did ⟨visit⟩, too.

In sum, Focus percolation in a chain explains the puzzle (71a) from the beginning of this section, while still predicting the ACD examples like (2) to be bad. Since focus percolation doesn't apply to matching relative clauses, it also doesn't predict (71b) (repeated in (104a)) to be good, unless it could be argued to allow a raising analysis. As (104b) shows, examples like (104a) are acceptable even when a raising analysis of the relative clause is ruled out by Condition C. Hence, the explanation of (104) cannot be focus percolation into the trace position.

(104) a. The cities Mary visited are near the lakes Bill did ⟨visit⟩.

b. The aunt of Mary$_i$ she$_i$ visited live near the uncle of Bill$_i$ he$_i$ did ⟨visit⟩.

3.3.3 Domain Expansion and Focus Index Sloppiness

Example (71b) (just repeated in (104a)) still needs an account. What I suggest, is that (71b) is similar to examples that have a sloppy reading. Consider (105a), repeated from (74) above, on a sloppy reading of the second VP. The LF-representation of (104a) is shown in (105), where the elided VP and a potential antecedent are indicated. On the sloppy reading the name and value of the variable x in the antecedent and that of y in the elided VP differ. So far I have been assuming that this difference is irrelevant for the identity condition of VP-ellipsis. It is, however, widely assumed that this assumption is incorrect and I present an argument against this assumption in chapter 4. That means, though, that the elided VP and the antecedent in (105b)

are predicted to be not identical, and it looks surprising that ellipsis is possible in (105a).

(105) a. John$_i$ admitted that Mary had bribed him$_i$.

Bill$_j$ admitted that she had ⟨bribed him$_j$⟩, too.

b. John $\lambda x\ x$ admitted that Mary had $\underbrace{\text{bribed } x}_{\neq \text{ antecedent}}$

Bill $\lambda y\ y$ admitted that Mary had $\underbrace{\text{bribed } y}_{\text{elided VP}}$

Rooth (1992b), incorporating an idea of Ristad (1990) and Fiengo and May (1994) into the focus semantics framework of Rooth (1992a), presents a solution for this problem. He proposes that the identity requirement (82) need not be verified for the elided VPs, but can instead be verified for any bigger constituent that contains the elided VP. If the bigger domain includes the binder of the variable, the names and values of the variables effectively aren't visible to the identity condition since, once bound, variables names have no semantic effect. For example, in the LF-representation in (106b) the domain of identity indicated and the antecedent are semantically identical: Both denote a function from individuals into truth values which yields true if and only if the individual admitted that Mary had bribed it. Therefore, (106) satisfies the identity condition in the expanded domain of identity.

(106) John $\underbrace{\lambda x\ x \text{ admitted that Mary had bribed } x}_{\text{antecedent}}$

Bill $\underbrace{\lambda y\ y \text{ admitted that Mary had } \underbrace{\text{bribed } y}_{\text{elided VP}}}_{\text{domain of identity } (\sim\text{P})}$

In (106), the expanded domains were exactly identical because the material between the binder and the elided VP was identical. There are examples of sloppy readings like (107) where this isn't the case. But, the solution to this problem is the focus sensitivity of the identity condition already argued for in (82) above: The material that is different in (107) between the antecedent and ∼P must be focussed, as Hardt (1992) notices, and therefore doesn't block identity on Rooth's (1992b) account.

(107) a. John$_i$ admitted that Mary had bribed him$_i$.

Bill$_j$ didn't admit that MAry had ⟨bribed him$_j$⟩

But he$_j$ admitted that SOMEbody had ⟨bribed him$_j$⟩ (Hardt 1992:(31))

b. he $\underbrace{\lambda x\ x \text{ admitted that } [\text{somebody}]_F \text{ had } \underbrace{\text{bribed him}_x}_{\text{elided VP}}}_{\text{domain of identity } (\sim\text{P})}$

I claim that (71b), repeated in (108a), is analogous to a sloppy reading. Recall that on the matching analysis of relative clauses, which I argued for in section 2.4, there are two heads, an internal head and an external head. Furthermore, the lexical content of the internal head depends on that of the external head. One could say that the external head binds the internal head, forgetting for the moment the argument in 2.4 that the internal head has an internal syntactic structure. Under this assumption, the LF-representation in (108b) is possible, where N and M are variables over NP-

part meanings that express the dependency of the internal head on the external head. In (108b), the antecedent is identical to ∼P, if *Bill* is focussed.

(108) a. The cities Mary visited are near the lakes Bill did ⟨visit⟩.

b. the cities $\underbrace{\lambda N\ \lambda x\ \text{Mary visited}\ [x,\ \text{N}]}_{\text{antecedent}}$ are near

the lakes $\underbrace{\lambda M\ \lambda y\ [\text{Bill}]_F\ \underbrace{\text{visited}\ [y,\ \text{M}]}_{\text{elided VP}}}_{\text{domain of identity (}\sim\text{P)}}$

That's essentially the solution of (71b): to capture the dependency of the non-focus correspondent of a focussed phrase by a semantic mechanism. The one last difficulty to consider is of a more technical nature: Based on Safir's observation, I argued in section 2.4 that the internal head of a matching relative is not just a variable, as in (108b), but a copy of the lexical material of the external head. Whereas a variable can be bound in the familiar way, for a copy of lexical material that's unexpected. The dependency of the internal head on the external head is not variable binding, but a deletion-antecedent relationship since the internal head is obligatorily phonologically deleted and therefore has to satisfy an identity requirement similar to that of VP-ellipsis. In the following paragraphs, I first show that the difference between the variable-binding relationship of the NP-parts assumed in (109b) and the deletion-antecedent relationship of the NP-parts argued for by Safir's fact causes a technical problem for the account of (71b). I then show that the same property of a deletion-antecedent relationship that's needed to solve this technical problem is also needed in examples of VP-deletion, one of them being Kratzer's (1991) 'Tanglewood-example'.

I then basically adopt the account of Kratzer (1991), but argue for one minor change. After that, I discuss some new predictions of the solution of (71b).

To begin with, it's necessary to distinguish the antecedents in examples with more than phrase that imposes an identity requirement. Following Rooth (1992a), I use numerical indices to mark the antecedents and the corresponding ∼Ps, the phrases that need to satisfy the identity requirement. Using this notation, the focus structure of (108a) is given in (109).

(109) $\underbrace{\text{the } \underbrace{[\text{cities}]_F}_{\text{antecedent}_2} \quad \lambda x \text{ Bill visited } [x, \underbrace{\text{cities}}_{\sim P_2}] \text{ are near}}_{\text{antecedent}_1}$

$\underbrace{\text{the } \underbrace{[\text{lakes}]_F}_{\text{antecedent}_3} \quad \lambda y \text{ Bill visited } [y, \underbrace{\text{lakes}}_{\sim P_3}]}_{\sim P_1}$

The question is why *lakes* in the trace position in the second relative clause is identical to *cities* in the relative clause trace in the first antecedent. The properties of (109) that are going to play a role in the a solution, are: one, that *lakes* stands in an deletion-antecedent relationship to another copy of *lakes*; two, that the other copy of *lakes* is focussed; and three, that the other copy is also part of the domain of identity under consideration.

Notice that examples that raise the same issue as (109) can be given with VP-destressing and deletion. Consider the VP in the *while*-clauses in (110). Except for the focussed verb, the VP can be destressed under identity to the preceeding material. But, part of the destressed material as well as the antecedent is a elided

VP that can be interpreted as indicated. The elided material is neither focussed nor identical to the elided VP in the first clause. Hence, it seems to violate the identity condition just like the lexical content of the the trace in (108a) does.

(110) a. John bought a book from the guy Bill did ⟨buy a book from⟩, while Sue [STOle]$_F$ *a book from the guy Bill did* ⟨*steal a book from*⟩

b. Every girl$_i$ writes like her$_i$ teacher does ⟨write⟩, while every boy$_j$ [TALks]$_F$ *like his$_j$ teacher does* ⟨*talk*⟩

The LF-representation of (110a) in (111) brings out the parallelism to (109) (To save space, the lexical content of the traces is not represented in (111)). The bold-faced instance of *steal* in ∼P$_3$ isn't identical to the corresponding verb *buy* in the antecedent, but isn't focussed. Therefore, it should block identity, on our assumptions so far.

(111) John $\underbrace{\lambda x \left[\text{the guy } \lambda y \text{ Bill } \underbrace{\text{did buy a book from } y}_{\sim P_2}\right] \lambda z\ x\ \underbrace{\text{bought a book from } z}_{\text{antecedent}_2}}_{\text{antecedent}_1}$

while

Sue $\underbrace{\lambda x \left[\text{the guy } \lambda y \text{ Bill } \underbrace{\text{did } \textbf{steal} \text{ a book from } y}_{\sim P_3}\right] \lambda z\ x\ \underbrace{[\text{steal}]_F \text{ a book from } z}_{\text{antecedent}_3}}_{\sim P_1}$

In both (109) and (111), the non-identical material of ∼P stood itself in an identity relationship to another instance of the same material within ∼P. It seems that in computing the identity requirement of ∼P such internal relationships need to be taken into account. More generally the problem can be characterized in the following

way: If an unfocussed XP is related by an identity condition to a focussed YP and the focus value of a domain that includes both XP and YP is computed, then the value of XP should covary with the focus-alternatives of YP. The examples in (110) also show that the solution given above, binding of the destressed XP by YP is insufficient because in (111) the destressed XP isn't c-commanded by YP.

A third place where the same problem comes up is (112a) from Kratzer (1991). The interpretation of (112a) paraphrased in (112b) shows that the relationship between the focussed instance of *Tanglewood* and the instance of Tanglewood in the elided VP must be visible for the interpretation of the focus particle *only*.

(112) a. I only went to [TANglewood]$_F$ because YOU did ⟨go to Tanglewood⟩ (Kratzer 1991:(15))

b. 'The only place such that I went there because you went there is Tanglewood.'

Recall, at this point, the assumption of Kratzer (1991) introduced in (81) that an F-marked constituent corresponds to a variable in the presuppositional skeleton. The argument Kratzer gives for this assumption is based on (112a). Her point is for the interpretation (112b) the elided instance of *Tanglewood* must also correspond to a variable in the presuppositional skeleton and more precisely, it must correspond to the variable as the focussed instance of *Tanglewood*.

Kratzer (1991) executes her proposal assuming on an identity of form requirement for VP-ellipsis, namely LF-copying of syntactic material. If it's required that

the F-mark and the focus index of the overt instance of *Tanglewood* are also present on the elided instance of *Tanglewood*, both instances are translated as the same variable in the presuppositional skeleton by the procedure (81).

(113) a. I only went to [TANglewood]$_F{}^1$ because YOU did ⟨go to Tanglewood$_F{}^1$⟩
(Kratzer 1991:(15))

As shown by (113), Kratzer's (1991) proposal predicts that *Tanglewood* in the elided VP is focussed. This seems counterintuitive on a deletion view of VP-ellipsis and is in fact ruled out by most analysis of ellipsis. Furthermore, Kratzer's example can also be created with VP-destressing as in (114a) and does allow the same reading, though most people prefer *there* in (114). In this case, it would be probably contradictory to assume either LF-copying or that the destressed *there* or occurence *Tanglewood* be focussed in (114a).

(114) I only went to [TANglewood]$_F$ because YOU *went to there*/$^{??}$ *Tanglewood*[16]

Finally, the claim that F-marking is present on material in a ∼P that corresponds to an F-marked constituent in the antecedent would interfere with my account Kennedy's example (2a) repeated in (115a). Under Kratzer's (1991) assumption, the relative clause internal copy of *country* is F-marked as shown in (115b). But, (115b) satisfies the focus sensitive identity condition, and hence (115a) would incorrectly be predicted to be acceptable.

[16]While (114) is clearly acceptable with *there*, it seems to be considerably degraded when the lexical item *Tanglewood* is reapeated. This is at present unexplained.

(115) a. *Polly visited every town in a country Eric did ⟨visit⟩.

b. *[every town, in a [country]$_F$ λy Eric $\underbrace{\text{visited } [y, [\text{country}]_F]}_{\text{elided VP}}$]

λx Polly $\underbrace{\text{visited } [x, \text{town}]}_{\text{antecedent}}$

The problem with Kratzer's account is that the correspondent of a focussed phrase is always interpreted as a variable in the presuppositional skeleton—that's after all what it means ot be focussed. Rather, it seems that the correspondent of a focussed phrase is only translated as a variable if the constituent the presuppositional skeleton is computed for includes the focussed constituent itself. Therefore, I propose to represent the correspondent of a focussed constituent not as focussed itself, but as bearing focus index, specifically the same focus index as the focussed phrase it corresponds to. Using this notation, consider the two representations in (116a) and (117a). For the computation of the presuppositional skeleton of (116a), the F-index on the instance of *Tanglewood* that corresponds to the focussed instance must be ignored. The result is (116b). But, for the computation of the presuppositional skeleton of (117b) the F-index on the same instance of *Tanglewood* must cause it to be interpreted as a variable, such that (117b) is the result. Moreover, the identity relationship internal to (117a) is probably the reason the focus index is present on elided instance of *Tanglewood*.

(116) a. ⟦went to [Tanglewood]^1⟧

b. 'went to Tanglewood'

(117) a. $[\![\underbrace{\text{went to [Tanglewood]}_{F}{}^{1}}_{\text{antecedent}} \text{ because [you]}^{2} \underbrace{\text{went to [Tanglewood]}^{1}}_{\sim\text{P}}]\!]$

b. 'went to x_1 because x_2 went to x_1'

Wold (1996, 1998) develops for independent purposes a formalism where a focus can be sometimes interpreted as a variable and at other times as its lexical content, depending on the constituent the presuppositional content is computed for. He calls this device a *Switch Strategy*. It's an interesting question whether his motivation for introducing this device and the use of it I make here are related in a more relevant way. For the moment, I just apply Wold's device to the case at hand.

Wold's idea is to only use partial assignment functions for the computation of the presuppositional skeleton and interpret a focus as a variable only if this variable is defined. In difference to Wold, I assume a distinction between F-marked constituents bearing a focus index and constituents that aren't F-marked, but bear a focus index. This difference is that the F-marked constituents must, in some sense, be able to enforce an interpretation as variable not only of themselves, but also of the the non-F-marked constituents bearing the same focus index.

Therefore, I postulate the interpretation rules of constituents bearing an F-index as in (118). What F-marking adds to a focus index x is that it forces an evaluation under an assignment G that assigns a value to x. Since, this effect influences the evaluation of the entire constituent under consideration, it can be used to trigger that the non-F-marked constituents with the same focus index are interpreted as variables.

(118) a. $[\![[\text{XP}]_F{}^x]\!]^G$ is only defined if x is in the Domain(G)

$[\![[\text{XP}]_F{}^x]\!]^G = G(x)$ where defined

b. $[\![[\text{XP}]^x]\!]^G = \begin{cases} G(x) & \text{if } x \text{ is in the domain of } G \\ [\![\text{XP}]\!]^G & \text{otherwise} \end{cases}$

The identity requirement must now be revised to take advantage of the flexibility the switch strategy provides. Moreover, it needs to ensure that the focus index of the correspondent of a focussed phrase is that of the focussed phrase. Therefore, the identity condition I propose has the two clauses in (119). The first clause is almost the same as (82) except for the domain minimality requirement. This requirement makes sure that a constituent bearing a focus index is only interpreted as a variable, if either itself or another instance of the same focus index in ∼P also bears an F-mark. The second clause of (119) makes sure that the correspondents in ∼P of a focussed phrase in the antecedent bear the same focus index.

(119) $[\![\text{antecedent}]\!] = [\![\sim\text{P}]\!]^G$ for a G with a minimal domain such that $[\![\sim\text{P}]\!]^G$ is defined and for this G:

$[\![\text{antecedent}]\!]^H = [\![\sim\text{P}]\!]^H$ for any H that expands[17] G

Coming back to (71b), repeated in (120a), consider the LF-representation in (120b). As indicated by the raised focus indices in (120b), the identity condition applying between the external and internal head of the relative clause, has ensured

[17] An (assignment) function H *expands* an (assignment) function G if the domain of H is a superset of the domain of G and for any x in the domain of G it's the case that $G(x) = H(x)$.

that the internal head bears the same focus index as the external head, though the internal head isn't F-marked. For the antecedent and ∼P indicated, the definition of the identity condition just given is satisfied. The reason is that ∼P contains an F-marked instance of the focus index 2, and therefore even the phrase [lakes]^2 in the relative clause internal trace position is interpreted as a variable when the presuppositional skeleton of ∼P is considered.

(120) a. The cities Mary visited are near the lakes Bill did ⟨visit⟩.

b. The $\underbrace{\text{[CIties]}_F{}^1\ \text{[MAry]}_F{}^3\ \text{visited}\ [x,\ \text{[cities]}^1]}_{\text{antecedent}}$ are near the $\underbrace{\text{[LAkes]}_F{}^2\ \text{[BIll]}_F{}^4\ \text{did visit}\ [y,\ \text{[lakes]}^2]}_{\sim\text{P}}$

In this way, the modification of Kratzer's (1991) approach developed above accounts for (71b). The status of the focus indices probably needs to be thought about more, if they occur independently of focus as I have claimed. Nevertheless, I believe the essential idea of the account is right; that there is a dependency between a focussed constituent in the antecedent and its correspondent in ∼P that is relevant when the presuppositional skeleton of a phrase containing both of them is computed. In particular, the account make one interesting prediction; namely, it's expected that the dependency it introduces is subject to the parallel dependencies requirement (45).[18] I argue now that two classes of examples show that the prediction is correct. The first are the examples of Kennedy's puzzle like (2a). The second one are cases

[18]This might intuitively seem surprising, but it follows from Rooth's (1992b) account of the parallel dependencies requirement as I show at the end of section 4.1.2. Note in this context that 'pseudo-sloppy' readings in (i) obey parallel dependencies, where the effect of a sloppy pronoun arises despite that lack of c-command.

like (72b) from above, where the only potential antecedent for a relative clause is a question with a different NP-part.

Consider first (2), repeated in (121a) with the LF-representation in (121b), where the relevant focus indices are indicated. The dependencies that are at issue in (121b) are: the dependency between *country*^{2} in the relative clause internal trace position and the focussed instance *country*$_F{}^2$ in the external head; and the dependency between the copy of *town*$_F{}^1$ in the QR-trace and the copy of *town*$_F{}^1$ in the phrase that moved by QR for ACD-resolution.

(121) a. *Polly visited every town in a country Eric did ⟨visit⟩.

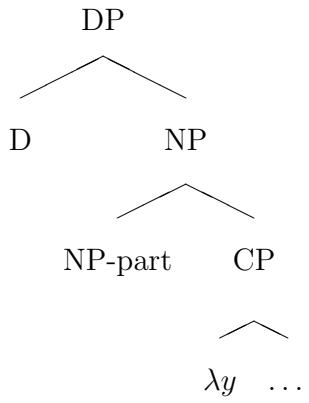
b. *$\Big[$every [town]$_F{}^1$, in a [country]$_F{}^2$ λy Eric visited [y, [country]^2]$\Big]$ λx Polly visited [x, town]$^1{}_F$

The two dependencies aren't parallel: Compare the structural relationship of the relative clause head to the relative clause in (122a), with that of the NP-part of the moved phrase to its complement in (122b). The NP-part in (122a) c-commands its dependent in the relative clause, while the NP-part in (122b) doesn't.

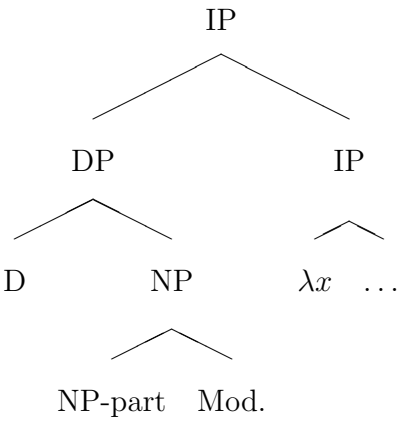
(i) a. The policeman who arrested John$_i$ read him$_i$ his$_i$ rights and the policeman who arrested Bill$_j$ did ⟨read him$_j$ his$_j$ rights⟩ too.

b. *The policeman who John$_i$ talked to read him$_i$ his$_i$ rights and the policeman who arrested Bill$_j$ did ⟨read him$_j$ his$_j$ rights⟩, too.

(122) a.



b.



Now consider the second case: an example where the only potential antecedent for a relative clause is a question with a different NP-part. The contrast in (123) ((123a) and (123b) repeated from (72)) illustrates that this case is also ill-formed. (123a) is an example where a relative clause serves as the antecedent for a question, while (123b), where the question is the potential antecedent for a relative clause, is ill-formed. The controls in (123c) and (123d) show that the ill-formedness of (123b) is due to the difference in lexical content of the antecedent of the relative clause internal trace and the the *wh*-word in the question.

(123) a. I know the cities Mary visited, but I would like to know which lakes she did ⟨visit⟩.

b. *I know which cities Mary visited, but I would like to know the lakes she did ⟨visit⟩

c. I know the cities Mary visited, but I would like to know which cities Bill did ⟨visit⟩.

d. I know which cities Mary visited, but I would like to know the cities Bill did ⟨visit⟩

The paradigm in (124) illustrates the same point. Again, the NP-parts of the trace antecedents differ in (124a) and (124b). (124a), where the antecedent is a relative clause, and the elided VP appears in a question, is acceptable just like (123a). (124b), however, where the question is the antecedent and the relative clause contains the elided VP is ungrammatical. In (124c) and (124c) the NP-parts of the two trace antecedents are identical, and there is no contrast between the relative clause antecedent, question with deletion case in (124c) and the question antecedent, relative clause with deletion in (124d).

(124) a. We know which is the house Marlyse bought, but not which car Paul did ⟨buy⟩

b. *We know which house Marlyse bought, but not which car is the one Paul did ⟨buy⟩

c. We know which is the house Marlyse bought, but not which house Paul did ⟨buy⟩

d. We know which house Marlyse bought, but not which is the house Paul did ⟨buy⟩

The facts in (123) and (124) seem to be a remarkable discovery, since they essentially recreate Kennedy's puzzle without antecedent containment. It is, I believe, no small achievement of the account developed here, that is predicts the entire paradigms in (123) and (124) correctly. The examples with identical NP-parts ((123c), (123d), (124c), and (124d)) are acceptable, because the content of the traces is identical, just like the examples (10). Consider the LF-representation in for (123d) in (125). For the ∼P indicated, the antecedent is suitable.

(125) I know which cities $\underbrace{\lambda x \text{ Mary visited } [x, \text{ cities}]}_{\text{antecedent}}$,
but I would like to know the cities $\underbrace{\lambda y \text{ [Bill]}_F \text{ did visit } [x, \text{ cities}]}_{\sim\text{P}}$

Next, consider the examples (123a) and (124a) where the NP-parts are different and the antecedent of a question is a relative clause. In the question the focus of the head of *wh*-chain can percolate to the trace position, as argued in the previous subsection. This is indicated in the LF-representation in (126). Because the lexical content of the trace in ∼P is focussed, it doesn't block identity, and the antecedent and ∼P satisfy the identity requirement.

(126) I know the cities $\underbrace{\lambda x \text{ Mary visited } [x, \text{ cities}]}_{\text{antecedent}}$,

but I would like to know [which [lakes]$_F$] $\underbrace{\lambda y \text{ she did visit } [y, [\text{lakes}]_F]}_{\sim\text{P}}$

Finally, consider (123b) and (124b). The LF-representation of (123b) is given in (127). Because focus cannot percolate to the trace position in a matching relative clause, the ∼P indicated in (127) isn't identical to the antecedent, even if the overt occurrence of *lakes* is focussed.

(127) *I know [which cities] $\underbrace{\lambda x \text{ Mary visited } [x, \text{ cities}]}_{\text{antecedent}}$,

but I would like to know the [lakes]$_F$ $\underbrace{\lambda y \text{ she did visit } [y, \text{ lakes}]}_{\neq \text{ antecedent}}$

If the domain the identity condition is applied to is expanded as in (128), the focus index dependency between the external and internal head of the matching relative is part of ∼P, and now effectively counts as a focus on *lakes*. But, (128) violates the parallel dependencies requirement: the dependency of the focus index 2 isn't parallel to that of focus index 1 in the antecedent. The difference between the two dependencies is due to the fact that the relative clause is a sister of NP, while the predicate created by *wh*-movement is a sister of DP. This is same structural difference as that shown in (122). Therefore, the examples like (123b) are in all respects relevant to the account pursued here alike to Kennedy's puzzle (2).

(128) *I know $\underbrace{[\text{which cities}_F{}^1]\ \lambda x\ \text{Mary visited}\ [x,\ \text{cities}_F{}^1]}_{\text{antecedent}}$,

but I would like to know $\underbrace{\text{the}\ [\text{lakes}]_F{}^2\ \lambda y\ \text{she did visit}\ [y,\ \text{lakes}^2]}_{\neq\ \text{antecedent}}$

The account of (71b) is hence corroborated by these non-trivial prediction. In the remainder of this section I show that, based on the account of (71b), an argument can be made for the assumption that the NP-part of a chain is not only represented in the trace position, but also in the higher positions of the chain unless binding is involved. I introduced this assumption in the introduction of chapter 2 without presenting any arguments in favor of it. At this point, an empirical argument for it can be made.

Recall that the lexical material inside a relative clause trace can become invisible to the identity condition if the focus index dependency to the external head is part of the domain considered. In addition the focus index dependency of the relative clause trace material is subject to the parallel dependencies requirement as attested by the ungrammaticality of (2) and (72) as just discussed. With this in mind, consider the examples in (129).

(129) a. Which city that Mary did ⟨visit⟩ is near the lake that John visited.

b. After I saw a lake that John visited, I was wondering which city that Mary did ⟨visit⟩ is more enjoyable.

The examples in (129) are like (71b), except that the relative clause that contains the elided VP is part of a *wh*-chain. The previous discussion of *wh*-chains in sections

2.2 and 3.1, argued that the NP-part of the *wh*-chain and the relative clause can be represented in different positions of the chain. However, the ellipsis in (129) requires parallel focus index dependencies of the head of the relative clause and the relative clause internal trace for the two relative clauses in both (130a) and (130b). If the lexical content of the trace is only represented in the trace position, even if the relative clause occurs in the top position of the chain the LF representation of (129a) is (130), which doesn't satisfy parallel dependencies for the focus indices 1 and 2.

(130) $^{*}\big[$Which λy that Mary visited $[y,$ city$^{2}]\big]$ λx is $[x,$ [city]$_{F}{}^{2}]$
near the lake$_{F}{}^{1}$ λz that John visited $[z,$ lake$^{1}]$

Therefore, if the NP-part in a chain is represented in only one position, it follows that the examples in (129) require reconstruction of the relative clause to the position that the NP-part must reconstruct to. Reconstruction results in the LF-representation in (131), where the focus index dependencies are parallel.

(131) $\big[$Which$\big]$ λx is $[x,$ [city]$_{F}{}^{2}$ λy that Mary visited $[y,$ city$^{2}]]$
near the lake$_{F}{}^{1}$ λz that John visited $[z,$ lake$^{1}]$

If, on the other hand, the NP-part of a *wh*-chain is represented not only in the trace position, but in every position of a chain, reconstruction of the relative clause in (129) is not required. Consider the LF-representation in (132), which is just like (130) except for the additional, seemingly redundant, instance of *city*$_{F}{}^{2}$ in the head position of the *wh*-chain. This extra instance of *city* satisfies focus index parallelism

in the domains indicated.

(132) $\Big[\text{Which } \underbrace{\text{city}_F{}^2\ \lambda y \text{ that Mary visited } [y, \text{city}^2]}_{\sim\text{P}}\Big]\ \lambda x \text{ is } [x, [\text{city}]_F{}^2]$

near the $\underbrace{\text{lake}_F{}^1\ \lambda z \text{ that John visited } [z, \text{lake}^1]}_{\text{antecedent}}$

A test for reconstruction is Condition C. The examples in (133a) and (133b) are structurally like the examples in (129), but reconstruction of the relative clause attached to the *wh*-word is blocked by Condition C. It seems that coreference is possible in (133a) and (133b) as indicated. Hence, I conclude that the examples in (133) argue that the NP-part is represented in all positions of a *wh*-chain.

(133) a. After I saw the lake that John visited, I was wondering which city that Mary$_i$ did ⟨visit⟩ she$_i$ would prefer to that lake.

b. The person John met at the party knows which girl that Mary$_i$ did ⟨meet at the party⟩ she$_i$ stayed in touch with afterwards.

3.4 Summary

In this chapter, I looked at ellipsis constructions where the elided constituent and the antecedent contain a trace. The question I asked is when the two traces are considered identical in the sense that is relevant for the licensing of ellipsis. It turns out that the most intricate pattern of data is found in the examples with the structure of Kennedy's puzzle (2), which sections 3.1 and 3.2 are concerned with. Section 3.3 looks at other constructions with traces in ellipsis and shows why their properties

differ from those of Kennedy's puzzle.

The main empirical discovery of section 3.1 are contrasts like (134) (repeated from (5)). The only difference between (134a) and (134b) is the lexical content of the NP the relative clause is attached to. This difference, however, affects the possibility of ellipsis: If the head of the relative clause is identical in meaning to the NP-part of object of the matrix verb visit, as in (134a) ellipsis is possible. Otherwise, ellipsis is impossible.

(134) a. Polly visited every town that's near the one Eric did ⟨visit⟩.

b. *Polly visited every town that's near the lake Eric did ⟨visit⟩.

Section 3.1 shows that contrasts like (134) argue that parts of the moved constituent are represented in the trace position of movement. If traces have lexical content, it's expected that whether two traces are identical for the licensing of ellipsis is affected by the lexical content of their antecedent. I argue that this is precisely the explanation of (134). Furthermore, section 3.1 shows that the amount of lexical material in a trace position that the identity criterion establishes is the same as that argued for based on the distribution of Condition C in chapter 2. Specifically, I show three cases where the two criteria lead to the same result: the effect of ACD, the integrity of the NP-part, and the the A/A-bar distinction. The results of chapter 2 and section 3.1 together are therefore a much stronger argument for the lexical content of a trace, then each result individually.

Section 3.2 establishes that the lexical content of a trace makes a semantic

contribution to the constituent containing the trace. I show that the semantic contribution of the content of the trace to the elided constituent and its antecedent is important for the licensing of ellipsis. One way I argue for this claim is to argue that the licensing of ellipsis only looks at the semantic content of the elided constituent and its antecedent—it requires identity of meaning, not identity of form. To make this point I summarize the account of Fox (1998a) for the few cases that were thought to require an identity of form requirement in addition to an identity of meaning requirement in Rooth (1992b), and conclude that Fox's (1998a) account renders the identity of form requirement redundant. The second argument for the claim that the semantic content of the traces determines the possibility of ellipsis comes from facts like (135)(repeated from (39)). (135) shows that the acceptability of examples with the structure of Kennedy's example is affected by the semantic relationship of the two NPs involved, the head of the relative clause, and the NP-part of the matrix object. The relevance of this semantic relationship cannot be explained if the two NPs, which constitute the content of the traces at logical form, don't make a semantic contribution in the trace positions. On the view that the content of a trace makes a semantic contribution, on the other hand, the effect of the semantic relationship can be predicted in the was section 3.2 discusses.

(135) a. Jon ordered a drink that's more expensive than the drink Sue did ⟨order⟩

b. *Jon ordered a drink that's more expensive than the dish Sue did ⟨order⟩

c. ??Jon ordered a cocktail that's more expensive than the beer Sue did ⟨order⟩

d. ?Jon ordered a drink that's more expensive than what Sue did ⟨order⟩

While sections 3.1 and 3.2 look at the identity criterion in examples with the structure of Kennedy's puzzle, section 3.3 looks at other cases. If explain why the effect of the identity requirement on traces is usually not observed in cases of *wh*-movement other than ACD, as for example in (136) (repeated from (71)).

(136) a. I know which cities Mary visited, but I have no idea which lakes she did ⟨visit⟩.

b. The cities Mary visited are near the lakes Bill did ⟨visit⟩.

Section 3.3.2 argues that (137a) should be explained based on the assumption a focus on the head of a *wh*-chain is also represented in the trace position. Then, the general observation that focussed material is not relevant for the identity condition of ellipsis also explains why ellipsis is possible in (137a). This account, however, doesn't carry over to (136b), because I argue that the focus on the head of a relative clause is not represented in the relative clause internal trace position in (136b). Section 3.3.2 argues that (136b) should be analyzed as a kind of sloppy reading of the dependency of the internal head of the relative clause and the external head. In particular, I show that the notion of sloppiness required for (136b) independently required following in one case the argumentation of Kratzer (1991). Interestingly the account developed in section 3.3.3 predicts that the identity criterion should apply in (137) (repeated from (123)) in the same way as in examples that have the structure of Kennedy's puzzle. This prediction is seen to support the approach developed in section 3.3.3.

(137) a. I know which cities Mary visited, and now I would like to know the cities Sue did ⟨visit⟩

b. *I know which cities Mary visited, but I would like to know the lakes she did ⟨visit⟩

Chapter 4

Linking Trace and Antecedent

In this chapter, I present several new results concerning the semantic mechanism that links a trace to its antecedent. I call this mechanism the *Dependency Mechanism*. This mechanism is a central piece of semantic competence, but it is unfortunately quite difficult to study. The results I present compare three widely used mathematical models of the dependency mechanism: one, the variable free model of combinatorial logic; two, variable binding in the form of first order logic (with restrictors added); and three, variable binding combined with λ-calculus. The three models are mathematically equivalent, as far as I know, and the choice between them is a matter of convenience for mathematic purposes. But, I present arguments that for linguistic purposes the third mechanism is the most appropriate. In the conclusions, I try to isolate the factors that make model three more successful in accounting for the data I present, in the hope that these properties are properties of the dependency mechanism.

In the organization of this thesis, this chapter starts a new topic. The previous

two chapters have concerned the content of the trace position in the LF-representation and its contribution to interpretation. Almost no attention was paid to the fact, that ultimately the interpretation of the antecedent of the chain must be able to, in some sense, involve the trace position. In the notation I used, a dependency was represented by the λx next to the antecedent phrase and the x that's part of the representation of the trace. This chapter and the following concern the mechanism(s) that accomplishes the semantic relationship of trace and antecedent.

(1) [Which book] λx did John read [x, book]

This chapter, in particular, concerns the aspects of the interpretation mechanism(s) that chains have in common with the relationship of a bound pronominal and its binder. For example the bound pronoun *his* in (2), is dependent on the interpretation of its binder *every boy* in a similar way to a trace. Chapter 5 addresses what is particular to chains, namely the semantic content of the trace position.

(2) Every boy$_i$ was riding his$_i$ bicycle.

The notation used to represent the fact that *which book* and *[x, book]* in (1) belong together semantically was chosen *ad hoc*, and I could have used instead any of the notations in (3) or infinitely many other notations. In (3a) and (3b), the a variable index is written in different positions. In (3c), the non-antecedents are marked instead of the antecedent. And in (3d), the dependency is graphically indicated in a manner similar to Higginbotham (1983). Probably, (3d) would have been the most appropriate

notation for the previous two chapters, since it doesn't suggest anything about the mechanism.

(3) a. [Which book]$_x$ did John read [x, book]

b. Which$_x$ book$_x$ did John read book$_x$

c. [Which book] .did .John .read [book]

d. which book did John read book

Obviously the notation isn't interesting; the mechanism is. Of the three models mentioned at the beginning, two use variables and assignment functions. Namely, λ-calculus and an extended first order logic with restrictors have variable in common. The third view, combinatorial logic, does without variables, but instead employs more complex rules of combination. I call the former view the *Variables View*, the latter the *Combinatorial View*. While I presuppose knowledge of the basics of the variables view, I briefly introduce the combinatorial view below just before the beginning of section 4.1. Before doing that, I sketch the kind of argument for the variables view I detail in section 4.1.

The main difference between the two views is that on the variables view the dependents in one dependency relation are different from those in any other dependency relation, namely the variables involved are different.[1] On the combinatorial view, however, the dependents are more or less semantically vacuous, and there's no

[1]Technically, it's only required to choose different variables for dependencies that overlap. In the discussion surrounding (11), I present reasons to believe that non-overlapping dependencies might also involve different indices.

semantic difference between the dependent of one dependency and those of another one. (This is, as mentioned, introduced below in detail.) Using the identity condition of focus semantics from the previous chapter, it's possible to distinguish the two views empirically: Consider a situation sketched in (4) where the domain of identity, ∼P, contains a dependent element but not the phrase it depends on. Furthermore, the antecedent in (4) contains a dependent element in a position corresponding to that in the domain of identity ∼P, but the phrase it depends on is different.

(4) binder … antecedent(… dependent …) … binder … domain of identity (∼P)(… dependent …)

On the variables view, the contributions of the dependents to the meaning of the antecedent and the meaning of ∼P could potentially be different in (4). Namely, the variables chosen for the dependencies could differ. On the combinatorial view, as explained below, the contribution the dependents make to the meaning of the antecedent and ∼P in (4) are identical. It turns out that, on the variables view, there's some reason that the variables chosen for the two dependencies actually must be different. Then, there's a clear difference in prediction made for a structure like (4). On the variables view, the dependent in ∼P must be focussed, since otherwise the identity condition is violated. On the combinatorial view, no focus is required in the same situation.

To test for the difference in predictions between the two views, a focus structure and interpretation like that in (4) must be argued for in an actual example. The dependent element of (4) is a pronoun that receives a sloppy reading, which is easy

to test for.[2] The other part of the situation in (4) is that a domain is subject to the identity condition, that includes the sloppy pronoun, but not the binder. Since the question which domains are subject to the identity condition is getting important, the following terminology is convenient: I call the domains that the identity condition must apply to *Focus Domains* following Truckenbrodt (1995) and I'll keep Rooth's notation to mark the domains that are subject to the identity condition with a $\sim$-mark in the focus semantic representation of a sentence.

I present two arguments distinguishing the variables view from the combinatorial view based on the difference in prediction for the situation in (4). The first argument relies on additional restriction on the placement of focus that Schwarzschild (1998) argues for. Informally, the requirement is that, if a phrase is focussed, it must be different from the antecedent—Schwarzschild calls this the *Avoid F(ocus)* Principle. Assuming this, the argument for the variables view comes from the fact that the dependent pronoun can optionally be focussed as in (5). On the variables view, because of the difference in variable name, Avoid F is satisfied if the focus domain in (4) is considered. On the combinatorial view, since there is no difference between the two pronouns, Avoid F cannot be satisfied for any choice of focus domain. Therefore, the focus placement in (5) is predicted to violate Avoid F on the combinatorial view.

(5) Every boy_i is riding his_i bike and every man_j is riding $[\text{HIS}_j]_F$ bike.

[2]It seems impossible to create the configuration in (4) with the dependent element being a trace, because the possibility of intermediate landing sites usually makes is possible that there is another binder close enough to the dependent to be part of $\sim$P.

While the first argument involved an optional focus domain, the second argument looks at a situation where a certain focus domain is forced. In that case, the variables view predicts that a sloppy pronoun must be focussed, while the combinatorial view predicts that it must be destressed. Below, I argue based on arguments of Schwarzschild (1998) that essentially every branching F-marked constituent must be a focus domain. Additional assumptions of Schwarzschild (1998) concerning the relationship between the placement of F-marks and the placement of pitch accent are introduced below and important for the argument. Together, they yield a very precise picture of the focus structure of (6d) if (6d) is part of the discourse in (6), and only *left* bears pitch accent in (6d). The focus structure argued for is indicated in (6d). The prediction of the variables view is that (6d) is ill-formed without focus on *her*, while the combinatorial view predicts (6d) to be well-formed. Since there is a contrast between (6d) and both (7a) and (7b) when part of the same discourse, I conclude that the prediction of the variables view is confirmed in (6d).

(6) a. A: Who cut the carrots?

b. B: John didn't. He$_i$ broke his$_i$ right hand.

c. A: Did Mary cut the carrots?

d. B: No. *Mary$_j$ cut $\underbrace{\underbrace{[\text{her}_j\ [\text{LEFT}]_F\ \text{hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

(7) a. B: No. Mary$_j$ cut $\underbrace{\underbrace{[\text{his}_i\ [\text{LEFT}]_F\ \text{hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

b. B: No. $\underbrace{\text{Mary}_j \text{ cut } \underbrace{[[\text{HER}_j]_F\ [\text{LEFT}]_F\ \text{hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

These two arguments for the variables view are presented in detail in sections 4.1.1 and 4.1.2. Together sections 4.1.2 and 4.1.1 argue not only that the variables view is more appropriate than the combinatorial view of binding, but also that the names of variables matter for the identity of meaning considerations relevant for focus and destressing. This assumption, that the indices of variables matter for semantic identity considerations, is underlying the index identity account of Kennedy's puzzle. This account is mentioned and argued against in the introduction of the previous chapter 3. The apparent conflict between the argument against the index identity view presented in chapter 3 and the arguments for the index identity requirement of section 4.1 is addressed in section 4.2. Retracing a line of argumentation of Heim (1997a), I show that the argument against the index identity view of Kennedy's puzzle argues only against one of the two popular incarnations of the variables view. I follow Heim (1997a) in calling the two implementations the *Formulas view* and the *Predicates view*. The difference between the two is indicated by the notation in (8). The formulas view, indicated in (8a), uses as a model an extension of standard first order logic to allow restricted quantification. The two arguments of a quantifier, restrictor and scope, are formulas with an unbound variable, which the operator binds. The predicates view uses λ-calculus as the model: the two arguments of a quantifier are one-place predicates, and the quantifier itself is a function from tuples of predicates to truth values. For index identity, the difference between the two views

is that on the latter the λ-predicate is a constituent where the index of the moved phrase is bound, but that doesn't include the moved phrase itself. As I show below, this difference favors the predicates view.

(8) a. [Which$_x$ book(x)] did John read [x, book]

b. [Which book] λx did John read [x, book]

4.1 Variables or Combinators

This section contrasts the view of binding based on the notion of a variable with that of combinatorial logic by looking at the focus domains that include only a sloppy pronoun, but not the binder of it. (See the discussion in footnote 2 about using traces instead of pronouns.) It starts by introducing the two views, and in subsections 4.1.1 and 4.1.2 argues that the indices of the variables view matter for the focus structure of a sentence. Namely, 4.1.1 shows that focus can force indices to be distinct, while 4.1.2 shows that absence of focus can under special circumstances force indices to be identical.

Both views of dependency I discuss, the variables view and the combinatorial view, are taken from mathematical logic. Since the variables view is older and more popular, I explicate it first.

Within mathematical logic, the status of variables has been viewed differently. In the first versions of first order logic, Frege (1884) and Whitehead and Russell (1910), a variable has no status other than marking a dependency for the statement of inference rules quantificational statements. An unbound variable, on this view, had

no well-defined meaning. With the advent of model theory, Tarski (1936), variables did get a meaning, namely the refer to a value that's provided by an assignment, a kind of storage and retrieval mechanism. This later concept of a variable is what has become to be the major model for dependent reference in linguistics. The use of variables and assignments in semantics is well-known and very clearly presented in recent textbooks (Larson and Segal 1995, Heim and Kratzer 1998). The essential idea is that the meaning of a phrase XP is relative to an assignment functions that must at least assign a value to the variables that occur free in XP. Secondly, the meaning of a complex phrase XP = [X Y] is a combination of the meanings of its parts X and Y relative to the same assignment function, except when one of the parts is a binder. In case one of the parts of XP is a new binder, the assignment relative to which the meaning of the sister is considered is modified, as shown in (9b) for the empty operator λ. (See section 4.2 below for definitions for quantifiers on the formulas view.)

(9) a. $[\![\text{X Y}]\!]^g = \mathcal{C}([\![\text{X}]\!]^g, [\![\text{Y}]\!]^g)$ where $\mathcal{C}$ is the semantic composition function (see section 1.1)

b. $[\![\lambda\zeta\ \text{Y}]\!]$ = the function $\lambda x.\ [\![Y]\!]^{g[\zeta \mapsto x]}$

The variables view as presented so far leaves it open to which extent the indices of unbound variables matter for the comparison of meanings needed for focus semantics. The following three possibilities come to mind to state the identity requirement for an antecedent XP and a focus domain YP: One, it could be that the

indices of variables don't matter at all for the identity condition. This can be stated as the requirement that there is an assignment g such that the meaning of XP under g is identical to that of YP under g. The two other views have in common the assumption that indices do matter, which can be expressed by an identity requirement that for every assignment g the meaning of XP under g is identical to that of YP under g. The difference between possibility two and three is whether reuse of indices is possible In mathematical logic, the formula in (10a) is well-formed and has the same meaning as (10b), because the two dependencies don't overlap (as long as x and y are unbound within the surrounding material indicated by dots in (10)). Possibility two is to assume that similarly in semantic representations that choice of index for a dependency is free except for the case of overlap.

(10) a. $(\forall x\colon \ldots x \ldots) \ldots (\forall x\colon \ldots x \ldots)$

b. $(\forall x\colon \ldots x \ldots) \ldots (\forall y\colon \ldots y \ldots)$

If reuse of an index for different dependencies was possible, the identity condition could in most cases with variables be satisfied by reusing an index. The third possibility is that the indices of variables do matter and that it's reuse of an index is not possible. The third possibility results in the strongest restriction and it's the version of the variables view advocated by Sag (1976) and Heim (1997a). Henceforth, when I mention the *variables view*, I refer to this third possibility as the 'official' version of the variables view. Heim (1997a) states the requirement that reuse of an index isn't possible as in (11) (see also Sag 1976, Chomsky 1986:75 and below)

(11) **No Meaningless Coindexing:** If an LF contains an occurrence of a variable v that is bound by a node α, then all occurrences of v in this LF must be bound by the same node α. (Heim 1997a:(24))

The results below argue in favor of not just the variables view, but more precisely, the third possibility of explicating it. One argument against the first possibility considered above, that indices don't matter at all, is the observation of McCawley (1976:328) (also Bach, p.c.to Williams 1977), that a deictic pronoun in an elided VP must refer to the same individual as the corresponding deictic pronoun in the antecedent VP does. This is shown for VP-deletion in (12a) and for destressing in (12b). If deictic pronouns are interpreted as unbound variables, for which the discourse provides an appropriate assignment, McCawley's observation can be stated as a requirement that the index of the variable the pronoun *him* is interpreted as must be the same in the focus domain and the antecedent. This requirement follows from the second and the third possibility mentioned above, but not from the first.

(12) a. Betsy saw him$_i$ and Sandy did ⟨see him$_i$⟩.

b. Betsy saw him$_i$ and Sandy *saw him$_i$*.

Note, however, that the argument based on deictic pronouns depends very much on the assumptions made for deictic pronouns. For example, if deictic pronouns are phonetically reduced forms of proper names, the facts in (12) would also be expected on the first view. As already mentioned the arguments I give below for the variables view argue, in fact, for possibility three from above.

Distinguishing the possibilities two and three empirically, I leave for below. It seems, though, that possibility three is also conceptually simpler: If we assume that the computational system of syntax doesn't use variables, variables are introduced at the point where the LF-structure of a sentence is translated into a semantic representation. As mentioned the reuse of an index must be prohibited on both possibilities, if the dependency the index was first used for is overlapping with the one it's being reused for. For example, such a restriction is needed for (13a), where two chains are overlapping. If in the translation of the syntactic representation containing chains into a semantic representation containing variables the indices of variables could be freely chosen, a semantic representation like (13b) must be blocked.

(13) a. ?What man$_i$ do you know what man$_j$ to talk to t_j about t_i?

b. *What man$_x$ do you know what man$_x$ to talk to x about x?

The easiest way to block (13b) is to postulate that different chains are always translated with a different variable index. This is possibility three. If, as possibility two assumes, it's sometimes possible to reuse an index the procedure translating syntactic chains into operator-variable dependencies would need to verify whether the no-overlap condition is satisfied. In particular, since this condition must be checked globally, on the entire structure that is translated, this seems undesirable.[3]

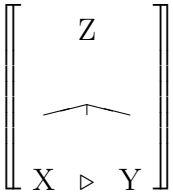
Combinatorial Logic is the only alternative to variables in mathematical logic,

[3]For pronouns, Condition C seems to be such a global condition on the translation of syntactic representation into semantic representation (David Pesetsky, p.c.). But, even if the existence of such global conditions is granted, this doesn't yet justify an unrestricted proliferation of such devices.

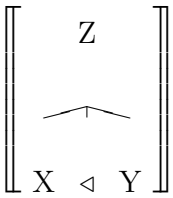
as far as I know (Schönfinkel 1924, Curry 1930, Curry and Feys 1958, Hindley *et al.* 1972). Though far less popular than the variables view, a treatment of dependencies modeled on categorial logic has been proposed by a number of people (Quine 1960, Szabolcsi 1987, Hepple 1990, 1992, Dowty 1992, Jacobson 1992, 1993, 1994, 1998a, 1998b). Since the different adaptations vary in their terminology and range and no standard has emerged, the exposition I give here uses the notation of Curry and Feys (1958).

A constituent XP that contains a dependent element, but not its antecedent is, on the categorial view, always interpreted as a function, that given an appropriate argument yields the interpretation that XP would have if the argument was inserted in the position of the dependent element. The open argument position of the dependent is kept open until the antecedent is encountered. To keep this position open, the semantic composition mechanism of a combinatorial semantics are more flexible. In the following, I annotate the composition rule that applies to determine the interpretation of a complex phrase from its parts in the node dominating the complex phrase. Most advocates of the combinatorial view use some convention like this. As mentioned in the introduction, on the variables the semantic composition rule applying for each phrase is probably predictable from the semantic types of its parts (Klein and Sag 1985, Heim and Kratzer 1998), and therefore I haven't indicated the composition rules above. For the combinatorial view, I am not aware of any discussion of the predictability of the composition rule applying—because of the bigger inventory of composition rules, the result of the variables view doesn't carry over to the combinatorial view.

The definitions of the combinators, I assume, are those marking functional application, function composition, and 'duplication'. For the first two, in addition, the direction must be indicated which I do with the signs ▷ and ◁ following for example Steedman (1996). Functional application, which is usually not indicated by any sign, therefore is indicated by just the direction mark as in (14).

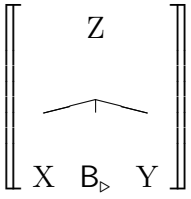
(14) 

is defined as ⟦X⟧(⟦Y⟧)

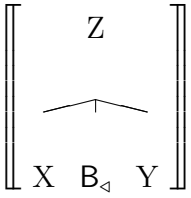


is defined as ⟦Y⟧(⟦X⟧)

Function composition is indicated by the letter B and the direction mark, as defined in (15).

(15) 

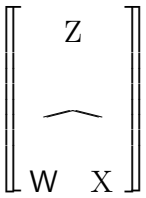
is defined as λx⟦X⟧(⟦Y⟧(x))



is defined as λx⟦Y⟧(⟦X⟧(x))

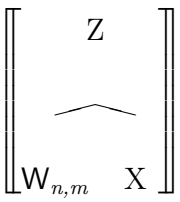
Important for binding is the *Duplicator*. This simple version of the duplicator in (16) when applied to a binary predicate X yields a unary predicate which is derived from

X by applying the same argument twice. In effect the duplicator enforces cobinding of two argument positions. I indicate the points where the duplicator applies with the letter W (Jacobson uses Z instead).

(16) 

is defined as $\lambda x.[\![\mathrm{X}]\!](x)(x)$

In the case of overlapping dependencies some version of the following generalized Duplicator is needed. It's however not needed for any of the cases below.

(17) 

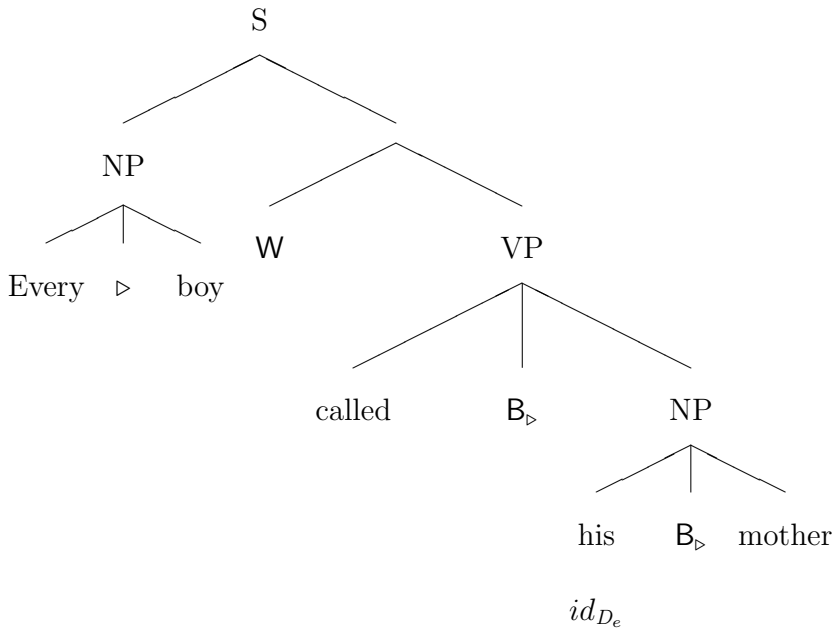
with $n < m$ is defined as

$\lambda \mathrm{x} \lambda y_1 \ldots \lambda y_{n\text{-}1} \lambda y_{n+1} \ldots \lambda y_{m\text{-}1}.[\![\mathrm{X}]\!](x)(y_{m\text{-}1}) \ldots (y_{n+1})(x)(y_{n\text{-}1}) \ldots (y_1)$

Dependent positions can be treated as either semantically empty or be interpreted as the identity function. I follow Jacobson (1998b) in assuming the latter. Consider then the example (18) of a bound pronoun. The semantic representation in (20) with the combinatory rules indicated yields the bound interpretation.

(18) Every boy$_i$ called his$_i$ mother.

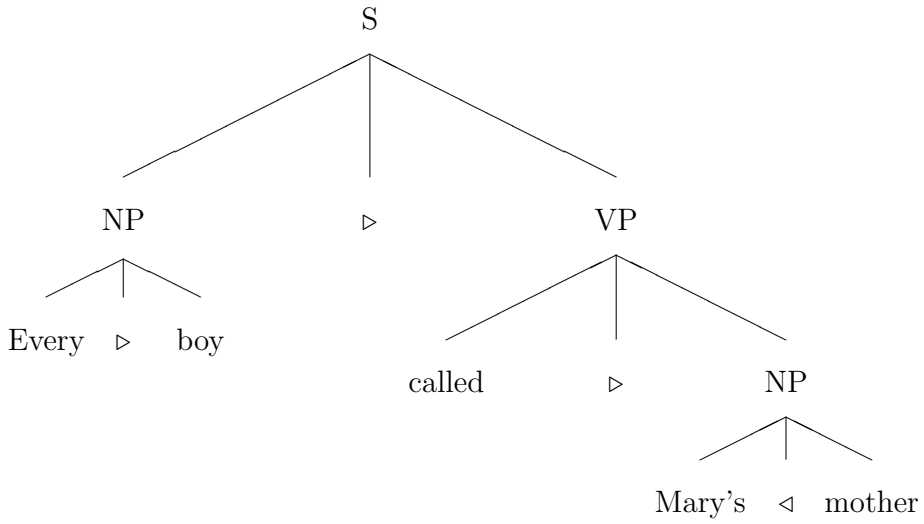
(19)



Compare (18) with an example where there's no binding like (20). In the semantic representation of (20) in (21), the combinatory rules are different from those in (19).

(20) Every boy called Mary's mother.

(21)



As the comparison between (19) and (21) indicates, the relationship between a dependent and its antecedent is notated in the composition principles on the combinatorial view. Therefore, it must be assumed that in the translation of syntactic chains into a semantic representation, the choice of combinator applying in each node is at least partially determined, such that the dependency of a chain is correctly represented. At present it seems to me that this requires annotating the LF-structure with the appropriate composition rule for each node, as I have done above.[4]

In contrast to the variables view, the combinatorial view seems to allow only one possibility with respect to how the identity condition of focus semantics and applies to a phrase that contains a dependent element, but not its antecedent. Since there are no indices, a one dependent means the same another. Hence, there is a difference between the combinatorial view and the official version of the variables view. The prediction of the combinatorial view is that a sloppy pronoun, like his_j in (22), is identical to its antecedent, no matter what domain of identity is under consideration.

[4]Pauline Jacobson (p.c.) notes that, if all restrictions on possible dependencies are part of interpretation, the syntax-semantics mapping needs no restrictions on the composition principles. This would be conceptually a simpler view of the syntax-semantic interface, and there are indeed clear examples of dependencies ruled out for semantic reasons, for example a pronoun *her* presupposes that its antecedent is of female gender. However, there seem to me to be equally clear cases of semantically conceivable dependencies that are impossible because the corresponding syntactic representation cannot be derived. A particularly well understood case is the difference between crossing and nesting dependencies illustrated in (i) (from Pesetsky 1982:268 with minor modifications). As Reinhart (1981), Rudin (1988), Koizumi (1994), and Richards (1997) show, the explanation of (i) is actually an interaction of the Shortest Attract requirement of syntax with a morphological property of English, namely how many Specifiers of CP are possible. Richards (1997) demonstrates that languages with a different morphological property, show the opposite judgement pattern for (i).

(i) a. ?What man$_i$ do you know what man$_j$ to talk to t_j about t_i?
b. *What man$_i$ do you know what man$_j$ to talk to t_i about t_j?

The arguments against the combinatorial view developed in sections 4.1.1 and 4.1.2, apply to the view Jacobson suggests.

(22) Every boy$_i$ called his$_i$ father and every teacher$_j$ *called his$_j$ father*

For example the meaning of antecedent and ∼P indicated in (23), are exactly identical.

(23) Every boy $\triangleleft$ W $\underbrace{[\text{called } \mathsf{B}_{\triangleright} [id_{D_e} \mathsf{B}_{\triangleleft} \text{ father}]]}_{\text{antecedent}}$ and

every $[\text{teacher}]_F \triangleleft$ W $\underbrace{[\text{called } \mathsf{B}_{\triangleright} [id_{D_e} \mathsf{B}_{\triangleleft} \text{ father}]]}_{\sim\text{P}}$

On the variables view, the antecedent and ∼P, as in (23) are not identical. Rather, the extended focus domains in (24) must be considered to license destressing in (22).

(24) $\underbrace{\text{Every boy } \lambda x \ x \text{ called } x\text{'s father}}_{\text{antecedent}}$ and

$\underbrace{\text{every } [\text{teacher}]_F \ \lambda y \ y \text{ called } y\text{'s father}}_{\sim\text{P}}$

Both views predict correctly that destressing is licensed in (22), though the licensing focus domains are different. To be actually able to distinguish the variables view from the combinatorial view empirically, it's necessary to have a better understanding of the distribution of focus domains. Both section 4.1.1 and section 4.1.2 argue first for a certain generalization about the distribution of focus domains and then consider the implication for the question of whether variables or combinators are the correct view. The generalizations about the placement of focus domains are based on the those of Schwarzschild (1998), but different.

4.1.1 Forcing Different Indices

Based on the old observation that focussed material must be 'new' in some sense, Schwarzschild (1994, 1998) argues for a ban against superfluous F-marking, which he calls the Avoid F condition. Obviously, this ban can be absolute since otherwise there would be no F-marking at all. Therefore the ban against F-marking must interact with the factors requiring focus, namely the identity requirements imposed by focus domains ∼P. The nature of this interaction is not obvious. For the moment, I assume the principle Avoid F as defined in (25), according to which the Avoid F condition applies to a structure after the focus domains have been determined. I discuss Schwarzschild's version of defining the Avoid F condition in section 4.1.2 and show below that (25) has empirical advantages.

(25) **Avoid F:** F-mark as little as possible without violating the identity requirements imposed by ∼Ps.

Direct evidence Schwarzschild (1998) gives for the Avoid F condition are question-answer pairs like (26). The answer in (26b) is felicitous where exactly the new material is focussed in the answer. The focus structure in (26c) is infelicitous because there is no focus domain structure such that *Mary* would not have an antecedent. This is expected if the entire answer constitutes a focus domain ∼P. Then focus on the new information, *John*, is required since no antecedent with *John* in the object position is available in the discourse in (26). The F-mark on *Mary* in (26c), however, is avoidable since an antecedent is available that satisfies the identity

condition without focus on *Mary*, as (26b) attests.

(26) a. Who did Mary praise?

b. Mary praised $[\text{JOHN}]_F$.

c. $^*[\text{Mary}]_F$ praised $[\text{John}]_F$

The argument for the variables in this section is based on the observation in (27) that the example of a sloppy reading (22), the sloppy pronoun can optionally be focussed. At first, an optional focus seems to be inconsistent with the idea of an Avoid F condition, and it's indeed inconsistent with Schwarzschild's (1998) statement of Avoid F as I show below. The way Avoid F is stated in (25), however, allows optionality of F-marking in principle, if different choices of focus domains force different amounts of F-marking. This is what I assume to be the case in (27). As was shown in (43a) above, the absence of focus on the sloppy pronoun can be explained easily. But, for the choice of focus domains considered there, since they tolerate the absence of focus on the sloppy pronoun, Avoid F blocks focus on the sloppy pronoun.

(27) Every boy_i called his_i father and every TEAcher_j called HIS_j father.

To satisfy Avoid F, there must be a placement of focus domains such that the focus on *his* is required in (27). At this point, the variables and the combinatorial view diverge: on the variables view there is such a placement of focus domains, namely that in (28). For $\sim\text{P}_1$ in (28) to be identical to the antecedent, the variable y must be focussed to be distinct from the antecedent.

(28) $\underbrace{\text{Every boy } \lambda x\ x\ \underbrace{\text{called } x\text{'s father}}_{\text{antecedent}_1}}_{\text{antecedent}_2}$ and $\underbrace{\text{every [teacher]}_F\ \lambda y\ y\ \underbrace{\text{called [}y\text{'s]}_F\text{ father}}_{\sim \text{P}_1}}_{\sim \text{P}_2}$

On the combinatorial view, on the other hand, there's no distribution of focus domains such that the focus marking on *his* is required. In particular, the choice of focus domain indicated in (29) is identical to the antecedent. Therefore, Avoid F cannot be satisfied for (27) on the combinatorial view.

(29) Every boy $\triangleleft$ W $\underbrace{[\text{called } \mathsf{B}_{\triangleright}\ [id_{D_e}\ \mathsf{B}_{\triangleleft}\ \text{father}]]}_{\text{antecedent}}$ and
every [teacher]$_F$ $\triangleleft$ W $\underbrace{[\text{called } \mathsf{B}_{\triangleright}\ [id_{D_e}\ \mathsf{B}_{\triangleleft}\ \text{father}]]}_{\sim \text{P}}$

Another example making the same point is (30), where in addition we see that the strict reading in (30a) is indeed blocked by the focus on the pronoun. The lack of the strict reading in (29a) is probably predicted by both the variables view ad the combinatorial view, but not important at this point.

(30) a. *John$_i$ called his$_i$ mother and Bill$_j$ called [HIS$_i$]$_F$ mother.

b. John$_i$ called his$_i$ mother and Bill$_j$ called [HIS$_j$]$_F$ mother.

Irene Heim (p.c.) points out that examples like (31) where the ranges of the quantifiers binding the two pronouns overlap don't allow focussing of the sloppy pronoun in the second conjunct. The difference between (31) and (27) is unexpected. The observation might indicate that contrastiveness of a sloppy pronoun requires more than a different index. An different way of thinking about (31) might be to ask

the question whether the semantic relationship of the two quantifiers affects whether a focus domain that doesn't include the quantifier can be considered. At this point, I leave the issue brought up by (31) open.

(31) *I expected every student to call his father, but only every YOUNG student called HIS father.

The first argument for the variables view, was based on the observation that pronouns with different binders can contrast. This was shown to be unexpected on the combinatorial view, while on the variables view the difference in indices provides the necessary contrast.

4.1.2 Forcing Index Identity

The second argument for variables is an attempt to force a focus domain that includes a sloppy pronoun, but not its antecedent. As was argued above, the variables approach predicts that in this case the sloppy pronoun must be focussed, while the combinatorial view predicts that the sloppy pronoun need not be—in fact, because of the Avoid F principle, must not be—focussed.

In the examples of sloppy readings considered so far, it was always possible to extend the domain of focus such that it includes the antecedent of the pronoun. To construct an example where this isn't the case, I again rely on ideas of Schwarzschild (1998). The first relevant observation of Schwarzschild, is that in cases like (32) the answer to the question must obligatorily be a focus domain, since otherwise no focus

would be required in the answer. I propose for the moment to capture this observation by the condition (32). If every sentence in a discourse must be a focus domain, the identity requirement applies and forces new material to be focussed.

(32) Who praised who?

*Mary praised John.

*[MARY]$_F$ praised John.

[Mary]$_F$ praised [John]$_F$

(33) Every sentence in a discourse must be a focus domain.

The condition (33) is derived from other conditions below, but the empirical generalization behind (33) seems correct, and the arguments in the following rely just on (33).

The second important observation of Schwarzschild (1998) is that accent placement within a focussed VP is sensitive to the same constraints as elsewhere. Consider Schwarzschild's (1998) contrast between (34) and (35). In both examples the VP must be focussed since the question is asking for the VP-information. Nevertheless, the pitch accent must be placed on *John* in (34b) and on *praised* in (35b).

(34) a. What did Mary do?

b. She [praised JOHN]$_F$?

(35) a. What did John's mother do?

b. She [PRAISED John]$_F$?

Another class cases of Schwarzschild (1998) showing that the placement of the pitch accent inside an F-marked constituent is affected by the discourse is illustrated by the dialogue in (36). Only the pitch accent on *Donca* is required in (36c). But, the the object of *wreck* must also be F-marked, because otherwise the entire sentence (36c) isn't a licit focus domain: If the object wasn't F-marked, an antecedent of the form *Bill wrecked the convertible X* would be necessary to license the entire sentence as a focus domain. Since such a sentence is not part of the context but by (33) the whole sentence (36c) must a focus domain, either both the subject and the verb, or the object must be F-marked. Because only the object contains a pitch accent, I conclude that the object is F-marked in (36c).

(36) a. John drove the convertible that Barry liked.

b. Aha. And Bill wrecked a boat?

c. No, $\underbrace{\text{Bill wrecked [the convertible that DONCA liked]}_{F}}_{\sim\text{P}}$

One conclusion, Schwarzschild (1998) draws from facts like (34), (35), and (36) concerns the relationship between pitch accent and F-marking. Namely, he proposes that it's necessary and sufficient for the phonetic realization of F-marking that an F-marked phrase contains a pitch accent, with the one exception stated in (37) for examples like (34b), which is however irrelevant for the following.

(37) **Phonological Realization of F-marking:** Every F-marked phrase must contain a pitch accent. (except for an F-marked verb whose complement is

also F-marked and contains a pitch accent)

Condition (37) leaves it open on which word within a complex F-marked phrase the pitch accent falls. But, the placement of pitch in a complex F-marked phrase is determined by the preceeding discourse, in a similar way that determines the placement of pitch within matrix sentences. For example, it's impossible in the context of (36c) to place the pitch accent on the verb *liked*.

(38) *No, Bill wrecked [the convertible that Donca LIKed]$_F$

The focus domain mark on the entire sentence doesn't make any prediction concerning pitch placement within the F-marked phrase, because the effect of the F-marking on the object is to make the information in its scope irrelevant to the focus domain it's part of. In the definition of the presuppositional skeleton (81) in section 3.3 the F-marked constituents of a focus domain were replaced by variables for this reason. For the same reason, any focus domain mark that includes the entire F-marked object in (36c) will not distinguish between (36c) and (38). Therefore, there must be a focus domain within the F-marked object to capture Schwarzschild's observation that the placement of pitch accent within an F-marked constituent is determined by the same discourse considerations that determine pitch placement otherwise. There are a number of possibilities to spell this insight out more precisely I present Schwarzschild's account of (36) first, but am ultimately going to draw slightly different conclusions which are closer to Truckenbrodt's (1995).

Schwarzschild (1998) proposes that all non-F-marked constituents are focus

domains. Furthermore, Schwarzschild states the Avoid-F principle as a global condition, that requires minimization of F-marking by looking at the entire sentence up to the requirement that non F-marked constituents must satisfy the identity condition. Together, the two assumption explain the paradigms in (34), (35) and (36), as Schwarzschild shows in detail. Consider, for example (36) repeated in (39). As already argued, the object in (39c) must be F-marked. However, it remains open whether the subconstituents of the F-marked object are also F-marked. Because of Avoid-F, F-marking is to be avoided here too. And since the antecedent *the convertible that Barry liked* is part of the discourse, F-marking is only required on the noun *Donca*, which must receive the pitch accent. In (39c), I indicated the two F-marked constituents and all the focus domains that Schwarzschild's proposal predicts. Notice though that one attractive aspect of Schwarzschild's proposal is that focus domains need not be indicated, because the presence of a focus domain is indicated by the absence of an F-mark.

(39) a. John drove the convertible that Barry liked.

b. Aha. And Bill wrecked a boat?

c. $\underbrace{\text{Bill}}_{\text{YP}}\ \underbrace{\text{wrecked}}_{\text{YP}}\ [\underbrace{\text{the}}_{\text{YP}}\ \underbrace{\text{convertible}}_{\text{YP}}\ \underbrace{[\text{DONca}]_F\ \underbrace{\text{liked}}_{\text{YP}}]_F}_{\text{YP}}}_{\text{YP}}}_{\text{YP}}}_{\text{YP}}$

Going back to sloppy readings, Schwarzschild's (1998) proposal is only compatible with the combinatorial view of binding. Consider as an example of a sloppy

reading (40) (repeated from (22)). In the second conjunct of (40), only the noun *teacher* must be focussed. Therefore, one of the many focus domains Schwarzschild's proposal predicts is the one indicated in (40). But, as argued above, ∼P in (40) doesn't have an antecedent on the variables view of binding, while the first conjunct provides an antecedent on the combinatorial view of binding. I take this consequence of Schwarzschild's (1998) proposal to be undesirable because of the evidence presented in the previous section against the combinatorial view of binding.

(40) Every boy$_i$ called his$_i$ father and every [TEAcher$_j$]$_F$ $\underbrace{\textit{called his}_j\ \textit{father}}_{\sim\text{P}}$

A second argument against Schwarzschild's (1998) statement of Avoid F and the distribution of focus domains is the optionality of focus that was observed in the previous section. (41) repeats the example from (43a) and (27) where focus on the sloppy pronoun *his* is optional. If Avoid F attempts to minimize the number of F-marks for the entire sentence, F-marking of the sloppy in (41) is predicted to be impossible, because the alternative Focus structure without this F-mark is possible.

(41) Every boy$_i$ called his$_i$ father and every TEAcher$_j$ called his$_j$/HIS$_j$ father.

For these two reasons, I adopt a different proposal concerning the distribution of ∼Ps than Schwarzschild. Recall that the evidence in (34) to (36) shows that non-focussed material within a complex phrase requires a discourse antecedent like that of non-focussed material outside of an F-marked constituent. For destressed material outside of any F-marked constituent this requirement was captured by the

generalization (33), which, however, was insensitive to the focus structure internal to a complex F-marked phrase. It is therefore natural to consider the generalization (42) analogous to (33), which forces all complex F-marked phrases to be focus domains.[5]

(42) Every complex F-marked phrase is a focus domain.

However, (42) makes the wrong prediction for (35) (repeated in (43)). Consider the ∼P indicated in (43b). It requires an antecedent of the form *V John*, which arguably the discourse (43a) doesn't provide an antecedent for.

(43) a. What did John's mother do?

b. She $\underbrace{\underbrace{[[\text{PRAISED}]_F \text{ John}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

It seems that in (43b), a focus domain is required inside the F-marked phrase, but need not include more than the object *John*. Hence, I assume that the requirement for a ∼P is related to the presence non-F-marked phrase, just like in Schwarzschild's (1998) proposal. But, in contrast to Schwarzschild's proposal, I assume that it's sufficient for destressed material to occur in the scope (or domain) of a ∼P without any F-marks intervening. To capture the fact that an intervening F-mark interrupts the licensing between a ∼P and a destressed phrase, I define the notion of *immediate scope* in (45). For the licensing of destressed (i.e. non-F-marked) material, I propose

[5]The restriction to complex F-marked phrases, is needed because imposing (42) on F-marked terminals would lead to circularity: If an F-marked terminal was a focus domain, this focus domain would require domain F-marking of the terminal within this focus. This F-marking would create another even smaller focus domain, which would bring about further requirements *ad infinitum*.

the condition in (44).

(44) Every non-F-marked phrase must be in the immediate scope of a ∼P.

(45) X is in the *immediate scope* of ∼P if there's no F-mark dominating X, but not dominating ∼P (and no other ∼-mark dominating X, but not dominating ∼P)[6]

The condition (44) accounts for the facts in (34) to (36), while allowing both a variables and a combinatorial account of simple examples of sloppy readings like (22). Consider first (35) (repeated in (43) and (46)). Because *praised* is F-marked in (46b), it doesn't need to be in the immediate scope of a ∼P. Therefore, the ∼P inside of the complex F-marked constituent needs to only include *John*, as indicated in (46b).

(46) a. What did John's mother do?

b. $\underbrace{\text{She } [[\text{PRAISED}]_F \overbrace{\text{John}}^{\sim\text{P}}]_F}_{\sim\text{P}}$

Next, consider (36) (repeated in (47)). One possibility of licensing all destressed constituents is the one indicated in (47). There are a number of other possible distributions of focus domains that condition (44) permits, but in all of them there is at least one focus domain within the complex F-marked constituent.

[6]The requirement that the be no intervening ∼-mark is unnecessary at this point. I include it though because then *immediate scope* expresses the intuition that the ∼P that X is in the immediate scope of is the primary one where the discourse requirement of a destressed ∼P is verified. The requirement does play a role below.

(47) a. John drove the convertible that Barry liked.

b. Aha. And Bill wrecked a boat?

c. No, Bill wrecked $\overbrace{\overbrace{[\text{the convertible that DONCA liked}]_F}^{\sim P}}^{\sim P}$

Thirdly, reconsider (22) (repeated in (48)) under the licensing condition (44). For the second conjunct of (48), (44) allows the focus domain structure indicated; namely only one focus domain that contains the sloppy pronoun and its antecedent. As shown above, the focus structure in (48) is predicted to satisfy the identity requirement of ∼P on both the variables and the combinatorial view of binding.

(48) Every boy$_i$ called his$_i$ father and $\underbrace{\text{every } [\text{TEAcher}_j]_F \textit{ called his}_j \textit{ father}}_{\sim P}$

I return now to the question of whether the combinatorial or the variables view of binding is more accurate. In this section so far, it's shown that Schwarzschild (1998) conclusions about the distribution of focus domains are only compatible with the combinatorial view, but a slightly different view of his facts allows us to maintain either the variables or the combinatorial view. The other important result of the discussion above, is that a destressed phrase that occurs in an F-marked constituent must be in the scope of a ∼P that is smaller than the F-marked constituent. In a sense, F-marked constituents are an upper boundary for the extension of focus domains. I show now that this result together with the variables view makes a new prediction about the availability of sloppy readings that is borne out.

Recall that the variables approach requires that a destressed sloppy pronoun is licensed in a focus domain that also includes the binder. In the examples considered so far, it was always possible to choose a focus domain big enough to license a destressed sloppy pronoun. The result of the discussion of Schwarzschild's (1998) data, that F-marking limits the extension of focus domains, can block licensing of a sloppy pronoun. The prediction is that a sloppy pronoun that's part of a F-marked constituent which doesn't include the binder requires F-marking.

To test the prediction, I use discourses similar to Schwarzschild's (1998) example (36), but with a pronominal dependency. Clear examples aren't easy to create. However, all my consultants agreed on the example (6), repeated in (49), from the introduction. With pitch accent only on *left*, (49d) isn't possible in the discourse (49).

(49) a. A: Who cut the carrots?

b. B: John didn't. He_i broke his_i right hand.

c. A: Did Mary cut the carrots?

d. B: No. $^*\text{Mary}_j$ cut her_j LEFT hand.

I show first that (49d) must have a focus structure like (50), in the discourse above. The reasoning is analogous to that in (36) above: Because *Mary cut* is destressed, it must be part of a focus domain with (49c) as its antecedent. But, then the object *her left hand* must be F-marked. Condition (44) forces another focus domain internal to the F-marked object to exist, because the destressed words *her* and *hand* must be licensed by such a focus domain. Hence, there must be focus domain that contains

her, but not its antecedent *Mary*. If we assume that there is a preference to choose a big focus domain, this forces the focus domain shown in (50).[7]

(50) B: No. *Mary$_j$ cut $\underbrace{\underbrace{[\text{her}_j\ [\text{LEFT}]_F\ \text{hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

Then, (49d) is an example where the sloppy pronoun is in a focus domain that the antecedent isn't part of. As discussed above, the variables view predicts such an example to be impossible, while the combinatorial view predicts it be acceptable. If the judgement on (49d) is the one indicated, it therefore argues for the variables view.

As mentioned above, the judgement on (49d) is different when *her* is stressed. This is also predicted by the variables view, because the focus makes the index of the sloppy pronoun irrelevant for the identity condition on focus domains, as discussed in the previous subsection.

(51) B: No. Mary$_j$ cut $\underbrace{\underbrace{[[\text{HER}_j]_F\ [\text{LEFT}]_F\ \text{hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

Another important control are the strict readings in (52). (52a) and (52b) are acceptable with the same focus structure that was impossible for the sloppy reading in (50). For the licensing of (52b), it either needs to assumed that the antecedent *John cut his hand* has an alternative representation where *his* isn't bound by the antecedent

[7] In (49d), if *her* and *LEFT hand* form separate focus domains, with *Mary* the antecedent of the focus domain of *her*, the example would be predicted to be acceptable even on the variables view. At the end of this section I argue that this possibility is blocked by a condition of Truckenbrodt (1995) that requires the maximalization of ∼Ps. The other examples I discuss in the following don't allow such a focus structure.

(Keenan 1971, Sag 1976:125, Reinhart 1981) or, as Rooth (1992b) suggests, that the antecedent has an entailment of the right form to license (52b), e.g. *Somebody cut his hand*, such that indirect identity in the sense of section 3.2 is satisfied.

(52) a. B: No. $\underbrace{\text{Mary}_j \text{ cut } \underbrace{[\text{his}_i\ [\text{LEFT}]_F \text{ hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

b. D: No. $\underbrace{\text{Mary}_j \text{ cut } \underbrace{[\text{John}_i\text{'s } [\text{LEFT}]_F \text{ hand}]_F}_{\sim\text{P}}}_{\sim\text{P}}$

The following two examples, illustrate the same point that (6) made. The example in (53) is initially quite hard to imagine as a discourse. But, once this difficulty is overcome most of my consultants agreed to the indicated judgment.

(53) a. A: John didn't wash the dishes. John damaged the car his father was leasing.

b. B: Aha. Did Mary wash the dishes?

c. A: No. *Mary washed the car her father was SELLING.

Again, it's instructive to compare (53c) with different pitch placements. The pitch placements in (54) improve the example.

(54) a. A: No. Mary washed the car HER father was SELLING.

b. A: No. MARY WASHed the car her father was SELLING.

The contrast between (53c) and the alternatives in (54) is predicted by the variables

view of dependencies. Look at the focus structures of the three examples, as given in (55). (55a) and (55b) are analogous to the previous example. (54b) as shown in (55c) can be licensed with one focus domain that includes both the antecedent and the sloppy pronoun. This ∼P can be licensed under identity to (53a).

(55) a. *Mary washed $\underbrace{\underbrace{[\text{the car her father was [selling]}_F]_F}_{\sim P}}_{\sim P}$

b. Mary washed $\underbrace{\underbrace{[\text{the car [her]}_F \text{ father was [selling]}_F]_F}_{\sim P}}_{\sim P}$

c. $\underbrace{[\text{Mary}]_F\ [\text{washed}]_F \text{ the car her father was [selling]}_F}_{\sim P}$

The example in (56), shows a preference in the predicted direction; namely that *his* needs to be stressed. But, the judgment is even less clear than that in the previous two examples. I suspect that to make (56) a more coherent discourse some people assume that (56b) indicates that every American suspects something that his teacher is something. If (56b) carries such an implicature with it, it could license the focus structure in (57) for (56c). Hence, those people are expected to find (56c) acceptable.

(56) a. A: Every Canadian believes that his teacher is a genius.

b. B: Is that so? Well, every American suspects something.

c. A: You're right. (*)Every American suspects that his teacher is an ALIEN

(57) $\underbrace{\text{Every American suspects that his teacher is an [ALIEN]}_F}_{\sim P}$

This concludes the argument for the variables view, that is the main point of this section. The remainder of this section contains two digressions. The first digression is about how the idea of Truckenbrodt (1995) that the domain of focus need to maximal could be incorporated into the version of Schwarzschild's (1998) system developed above. In particular, I argue that Truckenbrodt's condition can predict some cases where Kennedy's puzzle seemed to arise with A-movement and how it rules out the confound mentioned in footnote 7 above. The second digression contains some remarks towards a potential third argument for the variables view. Namely, it shows that it renders the parallel dependencies requirement (45) for page (45) partially redundant. However, it shows also that, at this point, two subcases of the parallel dependencies requirement remain.

Developing an idea of Rooth (1992a:114), Truckenbrodt (1995) argues that focus domains are also relevant for the phonology of focus and phonological phrasing. Furthermore, Truckenbrodt assumes that normally the focus domain surrounding a focussed phrase is extended to include as much destressed material as possible—the domain of a focus must be maximalized (Truckenbrodt 1995:126–30). Some of the examples above are relevant to the question of how the domain maximalization idea can be incorporated into the set of assumptions argued for above.

The first relevant point is that the domain maximalization condition cannot compare all possible ways of placing focus domains. The reason is the same that led me to abandon an Avoid F condition that compares all possible ways of placing Focus domains and F-marks. Namely, examples like (58a) (repeated from (5)) where focus is optional on a sloppy pronoun. As argued above, there must be two focus domains

in (58a), one surrounding the sloppy pronoun *his*, but not including the binder of it, and one containing the entire clause. (58b), on the other hand, contains only one focus domain—the one indicated. If the focus domain maximalization condition was to force (58b) to only have the one focus domain of (58a), (58a) would be predicted to violate Avoid F. Therefore, the two focus domains in (58a) must be permitted.

(58) a. Every boy$_i$ is riding his$_i$ bike and $\underbrace{\text{every MAN}_j \text{ is riding } \underbrace{[\text{HIS}_j]_F \text{ bike}}_{\sim\text{P}}}_{\sim\text{P}}$

b. Every boy$_i$ is riding his$_i$ bike and $\underbrace{\text{every MAN}_j \text{ is riding his}_j \text{ bike}}_{\sim\text{P}}$

Nevertheless, Truckenbrodt's intuition that focus domains can be too small or rather too vacuous seems to right in other cases. I propose therefore that the domain maximalization requirement only applies to focus domains that are trivial in the sense of (59). At this point, the definition of immediate scope given in (45) becomes important again: Recall that something is in the immediate scope of a focus domain if no F-mark nor ∼-mark dominates it that is inside of the focus domain.

(59) A focus domain ∼P is trivial if either there is no F-mark in the immediate scope of ∼P or there is no destressed material in the immediate scope of ∼P.

Notice that at least in (35) (repeated in (60)) a trivial focus domain was argued to be possible. However, even increasing the scope of $\sim P_{\mathcal{2}}$ in (60b) wouldn't lead to a greater immediate scope of ∼P, because the sister of ∼P is focussed.

(60) a. What did John's mother do?

b. She $\underbrace{[[\text{PRAISED}]_F \underbrace{\text{John}}_{\sim P_2}]_F}_{\sim P_1}$

Therefore, I assume that a trivial ∼P is blocked if an alternative focus structure is possible where the ∼P has more in its immediate scope. In the two cases I talk about now, the domain maximalization blocks a trivial ∼P to be the sister of another ∼P or be immediately dominated by another ∼P. The first case is illustrated by (61) (cf. footnote 7). In (61), $\sim P_2$ is trivial, because it only contains destressed material. Hence, I assume that (61) is blocked because of the possibility to replace $\sim P_2$ and $\sim P_3$ with one focus domain.

(61) B: No. *Mary$_j$ cut $\underbrace{[\underbrace{\text{her}_j}_{\sim P_2} \underbrace{[\text{LEFT}]_F \text{ hand}}_{\sim P_3}]_F}_{\sim P_1}$

The second case, are examples like (62) (repeated from (33b) on page 113). The ill-formedness of examples like (62) was left unexplained in the earlier discussion. Consider now the focus structure for (62) given: $\sim P_1$ and $\sim P_2$ are both trivial.

(62) $\underbrace{*[\text{Every man who said George would } t \underbrace{\text{buy some salmon}}_{\text{antecedent}_1}]_F \underbrace{\text{did } \langle \text{buy some salmon} \rangle}_{\sim P_1}}_{\sim P_2}$

Therefore the restriction on trivial focus domains proposed above requires instead the focus structure in (63) where $\sim P_1$ contains its antecedent. It is conceivable that this configuration is either ungrammatical, or at least hard to parse.

(63) $\underbrace{{}^{*}[\text{Every man who said George would } \underbrace{t \text{ buy some salmon}}_{\text{antecedent}_1}]_F \text{ did } \langle\text{buy some salmon}\rangle}_{\sim \text{P}_1}$

This analysis of (62) lacks a lot of detail at the moment. Nevertheless, I believe that it does look promising in the light of contrasts like those in (64) and similar ones in Heim (1997a).

(64) a. $^{*?}$Every man who wants George to leave should ⟨leave⟩.

b. $^{?}$Every man who wants George to leave did last time.

c. $^{?}$Every man who did wants George to leave.

The remaining paragraphs of this section point towards another potential argument for the variables view. Namely, I show that the variables view predicts some cases of the parallel dependencies generalization in (65) (repeated with minor modifications from (45) on 122) as Rooth (1992b) points out in passing, while the combinatorial view makes no prediction in this respect. The argument is very weak, though, since the variables view doesn't capture all cases that (65) account for, and therefore the condition (65) is still needed. I mention it largely because I feel that the variables view at least gives us a handle on the parallel depedencies requirement, and I hope that the Rooth's account can be extended to all cases of the parallel dependencies condition. Another reason to mention it, is that to show that the instances of focus index sloppiness where the parallel dependencies requirement was seen to apply belong to those cases that follow from the variables view.

(65) *Parallel Dependencies:* If a dependent isn't identical in reference to the corresponding dependent in the antecedent, it must stand in the same structural relationship to its binder as the corresponding dependent in the antecedent.

Consider the contrast in (66) (repeated from (44) on page (44)), which provides direct evidence for (65). The sloppy interpretation in (66a), which satisfies (65), is possible, while the sloppy reading indicated in (66b), which doesn't satisfy (65) is blocked.

(66) a. First, John told Mary$_i$ I was bad-mouthing her$_i$,
and then Sue told Jane$_j$ I was ⟨bad-mouthing her$_j$⟩

b. *First, John told Mary$_i$ I was bad-mouthing her$_i$,
and then Sue$_j$ told Jane I was ⟨bad-mouthing her$_j$⟩

The semantic representation of the second conjunct of (66b) on the variables view—to be explicit, I assume λ-calculus in (67)—is given in (67). On the variables view, the minimal focus domain that can be invoked for the licensing of deletion is one that includes the binder of the variable x.

(67) Sue λx told Jane I was $\underbrace{\underbrace{\text{badmouthing } x}_{\text{elided VP}}}_{\text{minimal } \sim\text{P}}$

What is a possible antecedent for the minimal ∼P indicated in (67)? The fact this it must be identical in meaning to the this ∼P modulo the focussed parts of ∼P restricts the possible antecedents to predicates. Furthermore, I claim the structure

of the antecedent predicate needs to effectively correspond to the structure of (67). The semantic contribution to ∼P of the parts (67) that aren't focussed must be exactly matched by a potential antecedent predicate, while for the focussed parts of (67) the antecedent must contain material that makes an equivalent contribution to its meaning. Since there are few cases where examples with different structures have exactly the same meaning, the semantic identity requirement effectively limits potential antecedents of (67) to predicates with the same internal structure. This can be assumed in the account of the parallel dependencies condition without loss of generality because the discussion of (46) on page (46) shows that if there are cases where predicates with different structure are semantically identical, the parallel dependencies condition is expected to be obviated. But, if the antecedent predicate has the same structure, this means specifically that the variable in the same structural position. In other words, the variables view predicts that any potential antecendent of ∼P in (67), is a predicate denoting phrase that is structurally isomorphic to (67) and where the variable predicated over appears in the same structural position as x does in (67). This predicts that no antecedent is available in (66) for the ∼P indicated in (67).

Before considering other potential choices of ∼P in (67), notice that while the prediction of the variables view pointed out in the previous paragraph doesn't block all cases accounted for by the parallel dependencies requirement. The difference is that the prediction just stated only requires that the antecedent predicate contain a variable in the same position as ∼P, but it doesn't require that there be a direct dependency between the two positions. The parallel dependencies requirement, how-

ever, requires a direct dependency in this antecedent. As Fox (1998c) points out, examples that show that the stronger requirement of the parallel dependencies condition is necessary are those known in the literature as Dahl's puzzle like (68a). The absence of the interpretation paraphrased in (68b) is the crucial fact, which shows that the representation sketched in (68c) impossible.

(68) a. Max said that he saw his mother and Oscar did ⟨say that he saw his mother⟩, too.

b. Max said that Max saw Max's mother and Oscar said that Max saw Oscar's mother, too.

c. Max said that he likes his mother

Now, return to the discussion of (66b) and consider the choice of focus domain as in (69), where the argument of the λ-operator that binds the variable x is also part of the focus domain. The considerations in the following carry over to any focus domain which contains the argument of the relevant λ-operator, also in examples where the focus domain contains other additional material than this argument. Again the question is: What is a possible antecedent for the $\sim$P indicated in (69)?

(69) Sue λx told Jane I was $\underbrace{\text{badmouthing } x}_{\text{elided VP}}$
$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\sim\text{P}}$

The same considerations as above show that all the possible antecedents of $\sim$P that need to be considered correspond to (69) in structure. Furthermore, it can be argued

that the structural positions occupied to x, which refers to *Sue*, and *Sue* must also have be identical in reference in the antecedent, in the cases to consider for the derivation of the parallel dependencies requirement. Since in the cases where the parallel dependencies condition applies the reference of the correspondent of x must be different from that of x, only this situation needs to be considered. But, this difference in reference will block identity, unless it's circumvented by focus in $\sim$P. Since the pronoun corresponding to x in (66b) cannot be focussed, the only way focus can affect the reference of x is to focus *Sue* in (69). Since in all elements of the focus set of (69) the reference of the position of *Sue* and that of the position of x are identical, this is required for the antecedent of $\sim$P as well. In other words, the two positions of the dependency of (69) must have the same reference in any potential antecedent of (69).

The prediction of the variables view just deduced again comes close to rendering the parallel dependencies requirement redundant, but doesn't fully succeed. For the example (66b), the prediction explains that no antecedent is available for the focus domain chosen in (69). There are, however, again examples that show that the stronger requirement of the parallel dependencies condition is needed. Again, Dahl's puzzle represents one class of such examples. An additional class of cases that Rooth (1992b) discusses, are examples like (70) where the two positions in the antecedent have the same reference, but no dependency exists between the two positions.

(70) *5 is (obviously) less than or equal to 5, and (of course) 7 is ⟨less than or equal to 7⟩, too.

Summing up this last point, the variables view covers a substantial amount of cases that are the empirical basis of the parallel dependencies requirement. At this point though the prediction of the variables view doesn't cover the cases (69) and (70), and therefore the parallel dependencies requirement is still needed.

4.2 Predicates or Formulas

The previous section argued that the variables view of dependencies is correct, and that the indices of unbound variables matter for semantic identity of phrases. As mentioned in the introduction, the variables view itself can be spelled out along the lines of two different mathematical models. One view, the *formulas view*, adopts the assumption of first order logic that every quantifier can bind a variable. The other view, the *predicates view*, follows λ-calculus (Church 1932, 1933) in assuming that there is only one operator, λ, that can bind a variable.

Both positions are quite popular in linguistics: for example, Larson and Segal (1995) assume and present in detail the formulas view, while Heim and Kratzer (1998) explicate the predicates view. Heim (1997a) contrasts the two views, and argues that they differ in their predictions in the case of ACD constructions. In this section, I summarize Heim's argumentation, but then argue based on the new data of the previous chapter 3 for the opposite conclusion of Heim's paper; namely, for the predicates view. I then give another argument for the predicates view, based on the distribution of i-within-i reference.

For the example (71), the difference between the two views is represented by

the sketches of semantic representations in (72). In (72a), which exemplifies the formulas view, the quantifier $which_x$ takes two formulas with the unbound variable x as its arguments. On the predicates view, exemplified by (72b), the two arguments of the quantifier are predicates, the lexical predicate *book* and the derived predicate *λx did John read [x, book]*.

(71) Which book did John read?

(72) a. Which$_x$ [x book] [did John read [x, book]]

b. [Which book] λx did John read [x, book]

On both views, binding requires a new semantic composition principle. On the formulas view, the rule has to apply to structures consisting of a quantifier and its two arguments. If we assume, that *which* is essentially an existential quantifier (see the following chapter), and example of such an composition rule is given in (73), and it illustrates the general schema.

(73) $\llbracket$ CP (which$_x$ R N) $\rrbracket^g = 1$ iff an a exists such that $[\![R]\!]^{g[x \mapsto a]} = 1$ and $[\![N]\!]^{g[x \mapsto a]} =$
1

On the predicates view, quantifiers themselves don't require a special composition rule; quantifiers can be understood as functions that take two predicates as an argument and yield a truth value. However, the λ-marking requires the special interpretation rule in (74) that binds a variable and creates a predicate.

(74) $\left[\!\!\left[\begin{array}{c} \text{YP} \\ \diagup\!\diagdown \\ \lambda x \quad \text{Y} \end{array}\right]\!\!\right]^g$ is interpreted as the function $a \mapsto [\![Y]\!]^{g[x \mapsto a]}$

The first argument for the predicates view has to do again with Kennedy's puzzle, as repeated in (75) from (2) on page 95.

(75) a. *Polly visited every town in a country Eric did ⟨visit⟩.

b. Polly visited every town Eric did ⟨visit⟩.

Consider first the semantic representation in (76), which the formulas approach predicts for the examples. As shown in (75), the variable index of the trace in the elided VP is different from that of the corresponding VP in the antecedent only in (76a), which is the representation of the ungrammatical (75a).

(76) a. *every$_x$ [x town in [a$_y$ [y country] [Eric $\underbrace{\text{visited } [y, \text{ country}]}_{\text{elided VP}}$]]] [Polly visited [$x$, town]]

b. every$_x$ [x town Eric $\underbrace{\text{visited } [x, \text{ town}]}_{\text{elided VP}}$] [Polly visited [$x$]town]

Since the indices of the two traces in (76a) are different, the parallel dependencies requirement must be satisfied by the two. This is however not the case, as the discussion of (122) on page 173 showed.

On the predicates approach, the semantic representations of (75) are those given in (77). In both (77a) and (77b), the variables of the trace positions differ, and in both cases there binders are in parallel positions.

(77) a. *[every town in a country λy Eric $\underbrace{\text{visited } [y, \text{country}]}_{\text{elided VP}}$] λx Polly visited [x, town]

b. [every town λy Eric $\underbrace{\text{visited } [y, \text{town}]}_{\text{elided VP}}$] λx Polly visited [x, town]

Therefore, (75a) is predicted to violate the parallel dependencies requirement only on the formulas approach. Heim (1997a) argues based on this observation for the formulas approach. Since it predicts (75a) to be ill-formed, she concludes that it's right. But, in the light of the facts observed in the previous chapter, it turns out that Heim's observation actually can be used for an argument against the formulas approach.

In the previous chapter I showed that (75a) is ruled out by the semantic content of the trace position. Therefore the fact that the formulas approach also rules it out, says little in favor of the formulas approach. In fact, it was shown with (5) on page (5), repeated in (78), all that's wrong with Kennedy's example is the semantic content of the trace.

(78) John visited a town that's near the town Mary did ⟨visit⟩.

Namely, (78) has the same structure as Kennedy's example above. The only difference between the two examples is the lexical content of the trace position. Since (78) is acceptable, it argues against the formulas approach which would predict it to violate the parallel dependencies condition. As shown above, (78) is predicted to be grammatical on the predicates approach.

Heim (1997b) mentions two other cases where ACD is possible, but index identity is not expected on the formulas view. Namely, comparatives as in (79a) and partitives in (79b). The same point as for (78) can be made for the examples in (79).

(79) a. John can run faster than Mary can ⟨run fast⟩.

b. Bill visited the three oldest cities out of the ones that Mary had advised him to ⟨visit t⟩.

The second argument for the formulas approach is based on the distribution of i-within-i reference. I use the term *i-within-i reference* in the following way: A pronoun exhibits *i-within-i reference* with a determiner D if the pronoun covaries in reference with the quantification of the determiner D and occurs inside the DP that determiner D projects. Chomsky (1981:212,229) observes the argument-adjunct distinction with respect to i-within-i reference illustrated in (80). If the pronominal anaphor *itself* occurs in the NP-part of the determiner *a* it cannot exhibit i-within-i reference with this determiner as shown by (80a). If the pronoun occurs in a relative clause adjoined to the NP-part it can refer i-within-i. The contrast in (80) shows that an adjunct occurring inside the NP-part of a DP doesn't allow i-within-i reference. In example (81a) from Vergnaud (1974:31) the pronoun *him* occurs in a relative clause, but one that is adjoined to an argument inside the NP-part of the relevant determiner. i-within-i reference is impossible in (81a), while it's possible in (81b) where the second DP itself occurs inside a relative clause adjoined to the NP-part of the first DP.

(80) a. *Kai drew [a picture of itself$_i$/it$_i$]$_i$

b. Kai drew [a picture showing itself$_i$]$_i$

(81) a. *the son of the woman who killed him was a Nazi (Vergnaud 1974:(62i))

b. the guy buried near the woman who killed him was a Nazi

The generalization illustrated is that *i*-within-*i* with a pronoun and a determiner D is possible if and only if the pronoun occurs outside the NP-part of the the DP projected by D. This is an argument for the predicates because the predicates approach predicts precisely this generalization, while the formulas approach doesn't. First, witness the failure of the formulas approach which is already noted in Higginbotham (1983:416–18) and Jacobson (1994). Recall that, on the formulas approach, both arguments of a quantifier must be open formulas containing a variable. For this reason, the NP-complement on the formulas approach must contain a subject position that contains a variable the quantificational determiner can bind. But, if this subject position can be bound by the determiner, it's predicted that the determiner should also be able to bind variables elsewhere in the NP-part of its complement. Hence, (82) is predicted to be a well-formed semantic representation on the formulas approach, but the DP it corresponds to in (80a) is ill-formed.

(82) *a_x [x picture of x]

The predicates approach predicts the distribution of *i*-within-*i* correctly: Recall that the two arguments of a quantifier are predicates and that the quantifier itself doesn't

bind a variable. Since the NP-part of a DP is a lexical predicate, it's not necessary to postulate a subject position in the NP-part. In fact, on the predicates approach, it's natural to postulate that there's no subject position in the NP-part of a DP that covaries with the determiners quantification. Then, the representation of illicit i-within-i reference is that in (83) (for (80a)), which is ruled out because x isn't bound.

(83) *a [picture of x] λx ...

Since relative clauses are derived predicates, they are predicted to allow i-within-i reference on the predicates approach. Recall that derived predicates are created by the λ-operator. Since the λ-operator can bind variables in its scope, representations like (84) for (80b) are well-formed.

(84) a [[picture] [λx x showing x]]

I conclude that the distribution of i-within-i reference is only predicted by the predicates approach, and therefore argues for it. Notice, by the way, that the combinatorial view of binding also doesn't predict the distribution of i-within-i reference, as Jacobson (1994) shows. On the combinatorial view, there is no difference between derived predicates and lexical predicates. In addition to the two arguments for the predicates view presented in this section, I know of two additional arguments for predicates: Sauerland (1998) presents an argument based on the existence of polyadic quantification, and Nissenbaum (1998) presents an argument based on the distribution of

parasitic gaps. Nissenbaum's (1998) argument is the most ambitious; he claims the existence of λ-operators as independent syntactic heads. While the other three arguments only provide evidence that the complement of a moved phrase as well as relative clauses are interpreted as predicates, they're compatible with Nissenbaum's (1998) stronger claim.

It should be mentioned that the predicates view also predicts that an argument of a lexical predicate cannot bind a pronoun in its scope. In (85a), Mary doesn't bind the pronoun *her*, because only the argument of a derived predicate can bind any pronouns. Hence, the subject must have moved as in (85b) for it to bind the pronoun. But, since it seems that many and maybe all DPs must move a short distance for case reasons, this prediction is maybe not as bothersome as it looks at first.

(85) a. *Mary likes her$_x$ bicycle.

b. Mary λx likes her$_x$ bicycle.

4.3 Summary

This sections investigates the contribution a dependent element to the meaning of a constituent that doesn't contain the binder of it. The tool that is employed to study this question is the semantics of focus and destressing. Hence, the examples considered mainly head the abstract structure in (86). The question that focus semantics can answer for a configuration like (86) is whether the semantic contribution of the dependents to the antecedent and the focus domain are the same or not.

(86) binder ... antecedent[... dependent ...] ... binder ... focus domain (∼P)[... dependent ...]

The results show that the answer to the question depends on how much material intervenes between the the binder and ∼P. One generalization that fits the results is the following: If the focus domain ∼P is smaller the the sister of the binder, the semantic contributions of the dependents differ between the two domains; if the focus domain ∼P is the sister of the binder, the semantic contributions of the two dependents are the same.

This generalization is predicted if λ-calculus is chosen as the mathematical model for the semantics of dependencies. The two other models considered, combinatorial logic and extended predicate calculus, were shown to predict substantially different generalizations which could are inconsistent with the data presented above. Namely, the combinatorial logic model predicts that the contributions of the dependents to the domains should always be identical, while the extended predicate calculus model predicts that the contributions should always be identical.

Chapter 5

Interpreting Moved Quantifiers

The previous chapter argued, that the interpretation of chains involves a mechanism that has the essential properties of variable binding. Still, for many of the structures considered in the chapters 2 and 3, it's not intuitively obvious how the interpretation procedure applies to a chain to yield the correct meaning. For example, consider (1) (repeated from (34) on page 50). The LF-representation of (1a) is given in (1b), with the operator and the trace of the relevant chains marked. It is clear that the variable x cannot refer to a single individual in (1b), because there need not be an single individual paper such Mary told every student to revise it, for (1a) to be a sensible question. But then, the question is what the variable x does refer to in the interpretation of (1b).

(1) a. [Which paper of his$_k$ that Mary$_j$ was given]$_i$ did she$_j$ tell every student$_k$ to revise t_i?

b. $\underbrace{[\text{Which } [\lambda z \text{ Mary}_j \text{ was given } [z]]]}_{\text{operator}} \lambda x$ did she$_j$ tell every student

$\lambda w\ [w]$ to revise $\underbrace{[x, \text{ paper of his}_w]}_{\text{trace}}$

Another interesting observation about (1b) is that the semantic division of questioned information and known information that the surface syntax of English suggests is not transparent in the LF-syntax. For example in (2b), the answer matches the question except for the *wh*-phrase. But, in an LF-representation like (1b) the *wh*-phrase and the rest of the question don't form separate constituents, as they seem to do on the surface in (2a).

(2) a. Q: Which friend of her's did Mary invite.

b. A: Mary invited Bill.

With non-interrogative DPs, an example of the kind of semantic representation entertained is given in (3), repeated from (13) from page 39. The semantic representation of (3a) argued for is (3b). The main feature of (3b) that seems counterintuitive are the three occurrences of *book of Irene's*. It occurs in the trace position inside the relative clause, in the trace position of quantifier raising and in the operator position of quantifier. In fact the matching analysis of relative clauses predicts that it also occurs in the operator position of the relative clause. To interpret the NP-part or any other segment of the restrictor in more than one position seems redundant. But, chapters 2 and 3 showed that the NP-part is often interpreted in the the trace position and section 3.3.3 provided an argument that the NP-part was also represented in the

operator position is relevant at LF.

(3) a. In the end, I asked him to teach the book of Irene's that David wanted me to ⟨ask him to teach⟩

b. $\underbrace{\left[\text{the book of Irene's } \lambda y \text{ that David}_i \text{ wanted me to teach } [y, \text{ book of Irene's}]\right]}_{\text{operator}}$

λx I asked him to teach $\underbrace{[x, \text{ book of Irene's}]}_{\text{trace}}$

The most fundamental problem for interpretation seems to be the one posed by (1), and in a more condensed way by (4), namely that the variable x in (4b) cannot be understood as referring to an individual. One solution for this problem was proposed by Engdahl (1980). She proposes that the variable in (4b) ranges over choice functions. Since Engdahl's (1980) solution for (4a) relies on representations like (4b) I directly adopt it for the case of interrogative quantifiers. Therefore, Engdahl's proposal is presented in some detail in section 5.1.

(4) a. Which friend of her$_i$'s did every student$_i$ invite?

b. Which λx did every student$_i$ invite [x, friend of herself$_i$]

I show then that Engdahl's proposal doesn't straightforwardly carry over to all non-interrogative quantifiers. Rather than concluding that therefore the semantics of interrogative and non-interrogative quantifiers is fundamentally, I show that it's possible to modify Engdahl's proposal such that all quantifiers can be explained as involving quantification over choice functions. At the end of section 5.1 I present an account for

the problem mentioned above that parts of the moved quantifiers must be interpreted in other positions of the chain than the trace position.

Section 5.1 develops Engdahl's proposal. I first present Engdahl's proposal and then go on to show that Engdahl's choice function can be extended to cover all the constructions considered in the previous chapters. The three main difficulties for this extension are the following: First, the fact that interrogative DPs seem to have, as we will see, always existential quantificational force, whereas non-interrogative DPs can be headed by determiners with a different quantificational force. The second difficulty is how to incorporate the contribution of material in the operator position. And, finally I address the interpretation of intermediate traces.

Section 5.2 points out one important prediction of the choice function approach developed in 5.1, namely that it predicts many weak crossover effects. The prediction arises from the the type difference between pronoun and the variables involved in the interpretation of chains, when choice functions are used.

5.1 A Choice Function Approach to All Quantifiers

The goal of this section is to develop a general interpretation procedure for all DP-chains making use of the insights of the previous chapters. Since many of the DPs considered in the previous chapters are *wh*-phrases, one task of the semantics is to account for *wh*-words in questions in a similar way as for other quantificational determiners. Hence, I start the section by summarizing Karttunen's (1977) semantics of questions which treats *wh*-words as existential quantifiers.

It turns out that it's easiest to talk about the meaning of a question when it occurs as the complement of *agree on*—I owe this insight to Lahiri 1991:16–25 and Rullmann and Beck 1997. In other environments, the meaning of questions is obscured either by the difficulty of understanding the semantic contribution of speech acts (in the case of matrix questions) or by the factivity of the question-embedding verb (in the case of other question-embedding verbs). In this section, I only consider the contribution to meaning of a question that appears as the complement of *agree on*, and refer to the specialized literature for the reduction of other cases to this one (Groenendijk and Stokhof 1984, Berman 1991, Lahiri 1991, Dayal 1996, and Hagstrom 1998).

Consider now the example in (5). What is the contribution of the embedded question *which student Lisa invited* to the meaning of (5)?

(5) Bill agrees with John on which student Lisa invited.

Assume that Bill and John both know the concept *student* fully and correctly. Then the truth of (5) implies, that for any student x, (6a) and (6b) must have the same truth value. And conversely, if for any student x, the sentences in (6) both have the same truth value, (5) would be considered true.

(6) a. Bill believes that Lisa invited x

b. John believes that Lisa invited x

If this intuition is any guide, the semantics of *agree on* involves quantification. I adopt

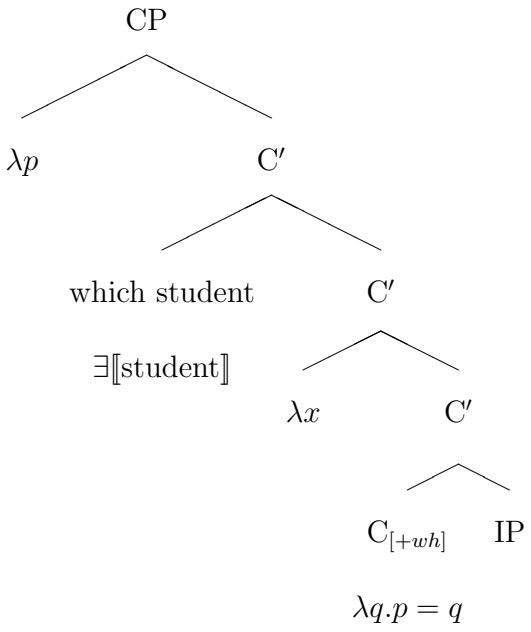
the proposal of Lahiri (1991) that *agree on* involves quantification over propositions. Then, the question must specify the range of propositions *agree on* quantifies over. For (5), the semantics of *agree on* could be given as in (7). This meaning of *agree on* leads to a certain view of the meaning of questions, which is due to Hamblin (1958, 1973). Namely, questions are essentially descriptions of a set of propositions—the set of propositions *agree on* quantifies over.

(7) If Bill believes p, then John believes p and vice versa for all propositions p of the form specified by *which student Lisa invited*

The remaining question is what set of propositions a specific questions is the description of. In the above example, the propositions quantified over are of the form *Lisa invited* x where x is a student.[1] Hence, a proposition p is quantified over if there is a student x such that p is the proposition "Lisa invited x". In this paraphrase, the contribution of the *wh*-word *which* seems to introduce existential quantification over students. This is, in fact, one popular view of the meaning of *wh*-words since Karttunen (1977) and is supported also by the morphological similarity amongst *wh*-words and indefinite determiners in many languages (see for example Cheng 1991 and Hagstrom 1998). It is now possible to isolate the contributions that the elements of a question make towards its meaning on Karttunen's (1977) approach, as given in the tree. The three interpretation rules needed are given in (9).

[1]The example also has a presupposition that Bill and John both believe that Lisa invited only one student. This is not relevant for the point here and I'll ignore it. It is though an interesting aspect of the semantics of questions and I refer to Schwarz (1993) and Dayal (1996) for discussion.

(8)



(9) a. $[\![\mathrm{C}_{[+wh]}]\!] = [\lambda q.p = q]$

b. $[\![\text{which}]\!] = \exists$

c. λp is introduced at CP-level to bind p in $\mathrm{C}_{[+wh]}$

Note that while the interpretation rules (9a) and (9b) specify the meaning of lexical entries, (9c) is unusual as a rule of the translation from syntactic logical form into a more semantic form of representation for two reasons. For one, it's specific to questions but isn't a rule specifying a lexical entry. Secondly, (9c) must introduce a binder λp for the unbound proposition variable that (9a) introduced; hence, (9c) is not a strictly local rule. However, as far as I know, there's at present no satisfying way around this undesirable feature of the semantics of questions.

With the semantics of *wh*-determiners in (9) in mind, look at Engdahl's example. In (10), it's given as the complement of *agree on*. What is the set of propositions

agree on could quantify over in (10)?

(10) Bill agrees with John on which friend of her$_i$'s every student$_i$ invited?

In a neutral context, the meaning of (10) can be elaborated in the following way: If John and Bill both know which individuals are students and who is friends with whom, the truth of (10) entails that for any x and y, if x is a student and y is a friend of x, (11a) and (11b) have the same truth value.[2] Conversely, if (11a) and (11b) have the same truth value for any pair of x and y with x a student x and y a friend of x, (10) would be considered true.

(11) a. Bill believes that x invited a y.

b. John believes that x invited a y.

This paraphrase of (10) using (11) seems to suggest that the subject universal *every student* takes scope outside of the proposition p that's quantified over in the interpretation of (10). This, however, cannot generally be the explanation of examples like (10).[3] While there might be cases where a universal quantifier can take scope outside of a question (cf. Higginbotham and May 1980, Groenendijk and Stokhof 1984, Chierchia 1993, Moltmann and Szabolcsi 1994), examples like Engdahl's are possible when the relevant quantifier cannot take scope outside of the question. In

[2]In a marked context, for example when preceded by a discussion of three kinds of typical friend relationships, boy-friend, oldest friend and grad-school buddy, (10) can be true even when the entailment to (11) doesn't hold.

[3]The argument in the following is I believe due to Engdahl (1986). I haven't been able to verify this, however.

(12), the quantifier that binds a variable in the *wh*-phrase is separated from the question's complementizer position by a finite clause boundary. Furthermore, the higher question-internal subject position in (12) is occupied by the quantifier *a professor*. Since (12) requires that there must be a single professor such that John and Bill believe that he claimed that every student invited somebody for it to have any answer, *a professor* takes obligatorily scope over *every student*. This is expected because of the finite clause boundary. But, if *every student* cannot take scope over *a professor*, it can also not take scope over the +*wh*-Comp that c-commands *a professor*. Nevertheless, the binding of the variable in the fronted *wh*-phrase is possible in (12). Hence, this kind of binding doesn't require quantification to a position outside of the question.

(12) They agree on which friend of her$_i$'s a professor claimed that every student$_i$ invited. (a $\gg$ every, *every $\gg$ a)

Since at least the interpretation example (12) involves a mechanism other scoping the quantifier to a position outside of the questions, I assume that (10), as well, is interpretable without scoping the subject quantifier to a position above the interrogative quantifier. Then the propositions described by the question in (10) must be of the form in (13), where y is a place-holder for the interpretation of the trace. Consider the proposition of the form in (13) if y was restricted to individuals in its interpretation. Then, (13) would entail that every student invited the same person, namely y. But, that entailment is wrong since for the truth of (10) it's not necessary that every student invited the same person. Hence, y must be able to refer to different

individuals covarying with the quantification of the subject. In a way, p in (13) must be equivalent to a conjunction of propositions of the form "x invites y" entertained earlier.

(13) p = 'every student invited y'

Since y cannot be the individual variable that intuition would favor, many less intuitive possibilities are now open. But, the results of chapter 2, I believe, narrow the options down significantly, leaving the proposal of Engdahl (1980) as perhaps the most natural candidate. Recall at this point the conclusion of section 2.2; namely, that the trace position in Engdahl's example must contain the NP-part of the *wh*-phrase. Hence, the propositions described by the questions in (10) must actually be of the form in (14), where y is the variable bound by the *wh*-word.

(14) p = 'every student$_i$ invited [y, friend of her$_i$'s]'

The meaning of the NP-part in the trace position covaries with the quantifier that binds *her*. For one given student that's quantified over by the subject—let's call her Mary—the NP-part *friend of her's* denotes the property *friend of Mary's* that the NP-part denotes. As shown above, the question meaning involves propositions of the form "x invited z" with z being a friend of x. If y in (14) selects one of the individuals that have the property the NP-part denotes, namely being a friend of Mary's, the result is a proposition of the form "Mary invited z", with z a friend of Mary's.

Engdahl's (1980) choice function proposal captures the intuition just expressed,

that the variable bound by the *wh*-word selects an individual that satisfies the property the NP-part expresses. A choice function is a function that assigns to properties individuals which have this property. Formally, this is defined in (15). Sometimes, I'll use the abbreviation *CF* for either the set of choice functions or the term *choice function*.

(15) f of type $\langle\langle e,t\rangle,e\rangle$ is a *Choice Function* if $x(f(x)) = 1$ for all $x \in \text{domain}(f)$

For Engdahl's example (repeated in (16a)) this results in the interpretation represented by (16b) and paraphrased in (16c).

(16) a. which friend of her$_i$'s every student$_i$ invited?

b. $\lambda p \exists \lambda \text{f}\,(f \in D_{\langle\langle e,t\rangle,t\rangle}$ and f is a CF and $\forall y \in \{\text{students}\}\text{: } y$ invited $f(\text{friends}(y))$

c. There is an f such that p means: for every student x, x invited the one that f chooses for the property *friends of x's*

The choice function f could, for example, always select the oldest of the people having a certain property. For this f, the proposition p described by (16b) is "Every student invited her oldest friend". But, the selections by f made could also not correspond to any natural definite description, for example f could choose the oldest friend for one student, the youngest friend for another, and a friend that's neither the oldest nor the youngest for a third student.[4]

[4]One problem of Engdahl's approach is exemplified by the case where two students have exactly the same friends. The properties of being a friend of these two students are extensionally identical, and therefore the choice function should select the same individual for both the students. The

Engdahl's account would carry over straightforwardly to the interpretation of other questions where the *wh*-phrase doesn't contain a bound variable, if the NP-part wasn't interpreted in the head of the chain. Consider the representation (17b) for the example (17a) (repeated from (5)). In (17b), the NP-part in the trace position doesn't vary with any quantifier in the sentence, hence the choice function f is only applied to the property *student* for which it selects a student. Hence, the proposition (17b) describes are exactly those of the form "Lisa invited x", with x being a student. But, given the argument in section 3.3.3 that the NP-part can appear in both the operator and the trace position, the representation of (17a) is probably (17c) rather than (17b). If this is correct, Engdahl account must be slightly modified in the way shown below.

(17) a. Bill agrees with John on which student Lisa invited?

b. which λf Lisa invited [f, student]

c. [which student] λf Lisa invited [f, student]

In the area of non-interrogative quantifiers, Engdahl's (1980) proposal has been widely adopted for wide scope indefinites (Reinhart 1994, 1997, Kratzer 1995, 1998, Ruys 1993, Winter 1997 and Matthewson 1998). In this case, the syntactic processes involved are different, namely the existential quantifier is not related to the

problem arises more sharply in examples like (i) where the property denoted by the NP-part, *ancestor of her's*, is necessarily the same for each person quantified over. I believe the problem indicates that the formal notion of property isn't fine-grained enough to reflect how the denotation of the NP-part is conceptualized in such cases (cf. Kratzer 1998).

(i) Which ancestor of her$_i$'s did every common daughter$_i$ of John and Mary like best?

indefinite that restricts it by a chain as attested by island insensitivity in (18) and further arguments in Ruys's (1993). Hence, the question of whether the NP-part of the indefinite also occurs in the operator position doesn't arise.

(18) a. Mary will leave if we invite a philosopher.

b. $\exists$ [λf Mary will leave if we invite [f, philosopher]]

Both interrogative quantifiers and indefinites involve existential quantification, as motivated above for questions. With universal quantifiers, quantification over choice functions also works straightforwardly, at least if the occurrence of the NP-part in the operator position, is ignored as von Stechow (1996) first pointed out.[5] Consider the example (19a) assuming (19b) as its semantic representation. For (19b) to be true, any way of selecting one of the cliffs must be one such that a girl is climbing the selected cliff. This is sufficient to the intuitive meaning of (19a), namely that for every cliff there's a girl who is climbing it.

(19) a. A (different) girl is climbing every cliff.

b. every λf a girl is climbing [f, cliff]

In addition to the problem of material occurring in the operator position, a generalization of the choice functions to the analysis of all quantifiers faces other significant problems. Consider (20a) with the cardinal quantifier *two* taking wide scope over the subject, assuming (20b) as its semantic representation. I show now

[5]Kai von Fintel (p.c.) drew my attention to this part of von Stechow's paper.

that (20a) is predicted to be true in a situation where there is only one cliff that a boy is climbing.

(20) a. A (different) boy is climbing two cliffs.

b. two λf a boy is climbing [f, cliff]

Assume that College Rock is the only cliff that a boy is climbing. Then, f and g in (21) are two different choice functions that make the predicate 'λf a boy is climbing [f, cliff]' true. Hence, (20b) is predicted to be true in a situation where College Rock is the only cliff climbed by a boy. But, this prediction is of course undesirable.

(21) $f(\llbracket$'cliffs'$\rrbracket)$ = College Rock, $f(\llbracket$mountains$\rrbracket)$ = Everest

$g(\llbracket$'cliffs'$\rrbracket)$ = College Rock, $g(\llbracket$mountains$\rrbracket)$ = Zugspitze

A similar problem arises also with the proportional quantifier *most* in example (22): The predicate over functions in (22b) will be true of infinitely many functions f even if only one girl is climbing a cliff. Hence, it's not easily possible to determine the proportion that would be required for (22a) to be true. More generally, all quantifiers for the interpretation of which the cardinality of the domain is important are problematic for the choice function proposal as developed so far because the cardinality of the set of choice functions that satisfy a predicate is typically either zero or infinite.[6] For existential and also universal quantifiers this property of choice

[6]For the truth of the kind of predicate arising in linguistic examples only the values of the choice function for a finite set of properties is relevant. Under these circumstances, for example, the number of choice functions satisfying the predicate will always be zero or infinite.

functions isn't problematic, but for most other quantifiers it is.

(22) a. A girl is climbing most cliffs.

b. most λf a girl is climbing [f, cliff]

There are probably many more ways out of the problem posed by (20). In the following, I discuss two of them. The first one is to restrict quantification to choice function that vary over choice functions that are minimal in their domains. Imposing this requirement on the choice functions considered ensures that two different ones of the choice functions consider differ in their value for a property actually considered in the evaluation of the sentence under consideration. To refute this approach, I show then a second problem that arises when quantification over choice functions is extended to all non-interrogative quantifiers. The second solution I present is to assume that two choice functions are only considered different if they are pointwise different—they differ in value for every property in their domain. This solution can account for both (21) and the second problem discussed below.

The first way out of the problem of (20) is to restrict the domain of the choice functions looked at to the properties that are 'really relevant'. 'Really relevant' are only the values of the choice function for those properties that it's actually applied to in the evaluation of the sentence in question. For example, in (22b), f is only evaluated for the one property: *cliff*. I assume that in (22) only choice functions defined for this one property are quantified over. The number of such choice functions is the same as the number of cliffs, because the number of ways to select one element from one set

is exactly the number of elements the set has. Hence, by looking only at such choice functions, it's possible to define cardinal and proportional quantifiers in a way that yields the right interpretation for (21) and (22).

Recall, though, that in example (10) above, the choice function must be defined for more than one property: Because the argument of the choice function contains a bound variable, the choice function must at least be defined for all the different properties that arise given the values the bound variable ranges over. Since it is desirable to postulate only one interpretation mechanism for all kinds of DPs, this case must be allowed by the restriction imposed on the set of choice functions a quantifier quantifies over. This is captured by the definition in (23).

(23) $\min(C) = \{ f \in C \mid \forall g \in C\colon \mathrm{domain}(g) \not\subset \mathrm{domain}(f) \}$

The general way to draw the restriction, hence, seems to be to restrict quantification to choice functions that have a minimal domain such that the argument of the quantifier are defined. The meaning of quantifiers can be given by the general schema in (25), which is exemplified in (24) for *two*, *most* and *which*.

(24) a. $[\![\text{two}]\!](S)$ is true if and only if two different elements of $\min(\mathrm{domain}(f))$ make $S(f)$ true.

b. $[\![\text{most}]\!](S)$ is true if and only if more than half of the elements of $\min(\mathrm{domain}(f))$ make $S(f)$ true.

c. $[\![\text{which}]\!](S)$ is true if and only if one $f \in \min(\mathrm{domain}(f))$ makes $S(f)$ true

(25) $[\![\text{Q}]\!](S)$=1 if and only if Q-many of min(domain(S)) are in $\{ f \mid S(f) = 1 \}$

Going back to the example (20), repeated in (26), the correct interpretations are now predicted. The minimal choice functions that the λf predicate in (26b) is defined for, are those that have only the predicate *cliff* in their domain. As argued above, the existence of two such choice functions is equivalent to the existence of two cliffs which a boy is climbing. Notice that whenever the content of the trace position doesn't contain a bound variable, the choice functions quantified over are predicted to be ones with a singleton domain.

(26) a. A (different) boy is climbing two cliffs.

b. two λf a boy is climbing [f, cliff]

Next, consider again the interpretation of Engdahl's example (27) (repeated from (10)). The minimal choice functions that the λf-predicate in (27b) is defined for are those that are defined for exactly all of the predicates *friend(y)* for the values of y quantified over.

(27) a. ... which friend of her$_i$'s every student$_i$ invited?

b. $\lambda p \exists \lambda \text{f}\, (f \in D_{\langle\langle e,t\rangle,t\rangle}$ and f is a CF and $\forall y \in \{\text{students}\}$: y invited $f(\text{friends}(y))$

c. There is an f such that p means: for every student x, x invited the one that f chooses from the set of friends of x's

Because the semantics given for (27) is general to all DPs, the question arises

whether the possibility that a bound variable in the lexical content of the trace can generally lead to quantification over choice functions with a domain of cardinality greater than one. The question is what kind of interpretation would be predicted for such a case. In the case of an indefinite quantifier, the resulting reading would be equivalent to a narrow scope reading. But, consider a possible wide scope construal of (28a) with the counting quantifier *two* taking scope over *every*. The semantic representation of (28a) I'm entertaining is shown in (28b).

(28) a. Every student brought two relatives of his$_i$.

b. two λx every student$_i$ brought [x, relatives of his$_i$]

The reading (28b) is predicted to have according to the previous section is I claim not available for (28a): Consider a situation with two students, one of which, brought two relatives of his, Lynn and Eve, but the other, Bill, brought only one, Sue. Intuitively, (28a) is false in such a situation. But, there are then two choice functions that make the λx-predicate in (28b) true—namely, f, which selects Lynn for the property *relatives of John* and Sue for the property *relative of Bill*, and g, which selects Eve for the property *relatives of John* and Sue for the property *relative of Bill*. Hence, (28b) is predicted to be true, which is incorrect.

The problem arises generally in the situation in (29), when Q_b is a strong quantifier, and Q_a is a quantifier sensitive to the cardinality of its domain. It's generally difficult to obtain wide scope of one strong quantifier over another, but seems to be marginally possible in examples like (30) (cf. Beghelli 1993, 1995, Sato-

Zhu 1996). Therefore, the incorrect reading is predicted in (28).

(29) $Q_a\ \lambda x\ \ldots\ [Q_b\ NP]_i \ldots [x, \ldots\ pro_i \ldots]$

(30) Every student read exactly two books.

The problem brought to light by (28) and (30) is not a problem specific to the approach I'm developing here. Rather it seems to arise by necessity from the assumption that all DPs have a uniform semantics. These uniform semantics need to provide an account for examples like Engdahl's (10), where a bound variable occurs inside a *wh*-quantifier. But, then the question whether this kind of binding is also possible with examples involving other quantifiers is unavoidable, and probably some representation equivalent to (29) must be allowed. Instead of giving up the idea of a uniform DP-semantics as Engdahl (1980) and others do, I want to maintain that representations like (29) are possible. But then, the interpretation of (28) cannot be the one given above. I propose that counting quantifiers individuate choice functions differently than assumed above. If we assume that the quantifier *two* requires that there are two choice functions that are different in their value on every argument they have in common, defined as the pointwise different relation in (31), (28b) is correctly predicted to be false in the situation laid out above.

(31) f is *pointwise different* from g if and only if

$\forall x \in \text{domain}(f) \cap \text{domain}(g)\colon f(x) \neq g(x)$

More generally, it's true that, for any set S of finite sets, the maximum number of

choice functions with domain S that are pointwise different from each other for each possible pair is equal to the cardinality of the smallest set $s \in S$. Therefore, if counting quantifiers require the choice functions that satisfy their domain to be pointwise different, even representations like (29) are interpreted correctly. This approach can also accommodate proportional quantifiers if it's assumed that here the maximum number of pointwise different choice functions that make the scope of the quantifier true is compared to the maximum number of pointwise different choice functions that make it false. Therefore, individuating choice functions with the pointwise different relation makes it possible to maintain a uniform semantics for interrogative and non-interrogative determiners.

Notice that the pointwise different requirement renders the restriction to minimal choice functions superfluous. Recall, that the observation that lead to the introduction of the minimality requirement was the following. An example like (32a) (repeated from (20)) with the representation in (32b) is predicted to true if the two choice functions f and g defined by (33) are considered. However, the requirement that the choice functions considered in the interpretation of a quantifier must be pointwise different doesn't permit the consideration of f and g as defined in (33). Therefore, the pointwise different requirement can replace the minimality condition.

(32) a. A (different) boy is climbing two cliffs.

b. two λf a boy is climbing $[f$, cliff$]$

(33) $f(\llbracket$'cliffs'$\rrbracket)$ = College Rock, $f(\llbracket$mountains$\rrbracket)$ = Everest

$g(\llbracket$'cliffs'$\rrbracket)$ = College Rock, $g(\llbracket$mountains$\rrbracket)$ = Zugspitze

The remainder of this section addresses the question of lexical content in positions other than the lowest trace position. At this time, what I can say about the issue is mostly mapping out the issues that arise and to show that the account is not incoherent because of these issues. To begin with, recall the problem: The schema for the definition of quantifiers in (25) above assumes that only the operator is interpreted in the operator position of a chain. In chapter 2, I argued that this assumption is incorrect: Both NP-parts and relative clauses can occur in the operator position of a chain. The potential occurence of NP-parts in the operator position was argued for in section 3.3.3. Recall that also relative clauses, according to chapters 2 and 3, can occur in the operator position of a chain as well, and in contrast to the NP-part don't even need to be repeated in the trace position. One of the arguments given in section 2.2 for this conclusion is Freidin's (1986) observation that Condition C can be obviated by over *wh*-movement in examples like (34) (repeated from (20) on page 44). The relative clause in (34) cannot occur in the trace position since its subject doesn't trigger a Condition C violation in this position.

(34) [Which argument that John$_i$ had criticized]$_j$ did he$_i$ accept t_j in the end?

It turns out that a relative clause in the operator position is actually easier to interpret there, then the NP-part is. The reason is that, as argued in section 2.4 relatives clauses usually contain an internal head. Recall that the matching analysis of relative clauses proposed in section 2.4 claims that the relative clause internal trace position is occupied by the NP-part of the relative clause head. Hence, the

LF-representation of (34) proposed is the one in (35). If the relative clause internal trace position also contains an NP-part, it seems natural to interpret relative clauses as predicates of choice functions in the same way as it was proposed for the sister of a moved quantifier above. The relative clause in (35) then denotes a predicate of choice functions that is true if the choice function assigns to the property *argument* an argument that John had criticized. Hence, if the choice functions that *which* quantifies over in (35) is restricted to those that satisfy the predicate the relative provides, the correct interpretation results in (35).

(35) [which argument λy John$_i$ had criticized [y, argument]] λx did he$_i$ accept [x, argument] in the end.

Generally, if a relative clauses has a non-empty internal head, it is interpreted as a predicate of choice functions. And, the relative clause is true if the choice function selects an individual that intuitively would make the relative clause true. Hence, relative clauses that have an internal head can be combined with the operator they share a position with as restrictors of the operator. The new definition schema for quantifiers is given in next.

(36) $[\![\text{Q}]\!](S)$=1 if and only if Q-many pointwise different choice functions of $\min(\text{domain}(S)) \cap \{f \mid R(f) = 1\}$ are in $\{f \mid S(f) = 1\}$

The NP-part is not as easily interpretable in the operator position as a relative clause is because the NP-part denotes a predicate of individuals, not of choice

functions. This mismatch in semantic types seems to arise naturally from the considerations above. Namely, the discussion of Engdahl's example (10) above showed that derived predicates with lexical material in an internal position must be, at least in some cases, be interpreted with a variable of type other than that of individuals. Hence, they denote a predicate of something other then individuals. The NP-part, on the other hand, must be a predicate of individuals since it can occur as an NP, for example, in a copular construction like (37).

(37) John is a student.

At this point, there doesn't seem to be an elegant way to let the NP-part in the operator position contribute to the interpretation of the chain. One way of making it interpretable is to assume that a predicate of individuals can be converted into a (trivial) predicate of choice functions by the type shifting rule in (38). Then, the NP-part in the operator position can be interpreted in the same way the relative clause was interpreted.

(38) $p^{et} \longrightarrow \lambda f^{\langle et\rangle e}(f(p) = f(p))$

The type-mismatch problem arises not only with the NP-part, which on the approach taken here seems to be redundant in the operator position anyway, but also for those relative clauses that don't contain an internal head. Recall that for examples like (39a) (repeated from (34) on page 50), the NP-part of the moved phrase has to reconstruct because it contains a bound variable, but the relative clause has to be represented in

the operator position because otherwise (39a) would violate Condition C. As argued in section 2.4, the fact that matching in a matching relative clause applies at LF predicts for (39a) the LF-representation in (39b) where the NP-part isn't represented in the relative clause internal trace position. But then, the null hypothesis is that the relative clause in (39b) is interpreted as a predicate of individuals. If this is right, the type-shifting rule in (38) must apply for (39b) to be interpretable.

(39) a. [Which paper of his$_k$ that Mary$_j$ was given]$_i$ did she$_j$ tell every student$_k$ to revise t_i?

b. $\Big[$Which $[\lambda z$ Mary$_j$ was given $[z]]\Big]$ λx did she$_j$ tell every student$_i$ to revise $[x$, paper of his$_i]$?

A type-shifting rule like (38) is also required in the other direction, because just like the NP-part isn't interpretable as a predicate of individuals in the operator position, a relative clause isn't interpretable in the trace position as a predicate of choice functions. One argument for this are examples like (40a) where the relative clause contains a bound pronoun. The LF-representation predicted for (40a) is given in (40b). The relative clause in (40b) is predicted to be interpreted as a predicate of choice functions that is true if the value for the property *paper* is a paper that relevant student *he*$_j$ read. In (40b), however, the interpretation of the relative clause must be a property of individual, so that it can serve as the argument of the choice function variable x. Hence, I assume that some kind of type-shifting can apply, like the rule given in (41). For the relative clause in (40b) the rule in (41) results in a

predicate that's true of a paper if he_j read it, which is appropriate.

(40) a. [Which paper that he_j read]$_i$ did every student_j like t_i?

b. [which paper] λx every student_j liked [x, paper λy he_j read [y, paper]]

(41) If $\forall f \in \text{domain}(P)\ \forall p \in \text{domain}(f)$: $p(x)$ and $\text{domain}(P) \neq \emptyset$,

$P^{\langle\langle et\rangle e\rangle t} \longrightarrow \lambda x\ (P(f_x) = 1)$, where f_x any $f \in \text{domain}(P)$ with $\exists p$: $f(p) = x$

With the type-shifting rule (41), the predicate that is the argument of *which paper* in (40b) requires that the choice function x be defined for the properties *paper that he likes* for each of the students. However, the NP-part, in the operator position is interpreted as a predicate of choice functions that requires that they are defined for the property *paper*. It's easy to see the interpretation assigned to (41b) is nevertheless correct. But, the fact that sometimes the content of traces that in some sense belong to the same operator have different lexical content, does give rise to a problem in the following.

Finally, material interpreted in intermediate traces requires the type-shifting just mentioned, but also raises the other problem just hinted at. Consider the example in (42a), repeated from (31) on page 48, and its LF-representation in (42b). The intermediate trace of the chain in (42b) contains the relative clause and the NP-part while the lowest trace and the operator position of the chain contain only the NP-part. Notice that the higher part of the chain in (42b), the operator and the intermediate trace, resembles the chain in (40), and the same interpretation procedure can apply. But, how does the predicate created by λy that contains the lowest trace, contribute

to the interpretation?

(42) a. [Which paper that he$_k$ gave to Mary$_j$]$_i$ did every student$_k$ think t'_i that she$_j$ would like t_i?

b. $\underbrace{\text{[Which paper]}}_{\text{operator}}$ λx every student$_i$ think $\underbrace{[x, \text{ paper}, \lambda z \text{ he}_i \text{ gave } [z, \text{ paper}] \text{ to Mary}_j]}_{\text{intermediate trace}}$

λy she$_j$ would like $\underbrace{[y, \text{ paper}]}_{\text{lowest trace}}$

For the moment, I pursue an approach in line with the structure given in (42b). Then, the natural proposal seems to be to apply the λy-predicate to the choice function x. But, this predicts the wrong interpretation for (42b): The denotation of the trace in the λy-predicate is the value the choice function y assigns to the property *paper*. But, the denotation of the trace in the intermediate position is the value x assigns to the property *paper he$_i$ gave to Mary*. Now, consider a situation where there are papers that every student gave to Mary, but no student thinks Mary would like the paper he gave her. Rather, each student thinks Mary would like the paper P, which nobody gave her. Intuitively, (41) should have no correct answer in such a situation. But, for a choice function that assigns to the property *paper* the value P, the λx-predicate in (42b) is true.

The problem with (42b), on the above account of how it's interpreted, is that there is no relationship as what the choice functions quantified over assign to the lowest trace that contains just the NP-part and to the intermediate trace that contains the NP-part and the relative clause. To remedy this problem, the λy-predicate in (42b) must apply not to the choice function x, but to the result of applying x to the

content of the intermediate trace. Since, this is an individual, another type shifting rule is needed. Namely, one like (43), which assigns to an individual x a choice function that chooses x for all properties that of x.

(43) $\zeta^e \longrightarrow x_\zeta$ where x_ζ is the choice function with
domain$(x_\zeta) = \{p^{et} \mid p(\zeta) = 1\}$ and $x(p) = \zeta$ for all p

5.2 Predictions of the Approach

The main prediction of the approach to the interpretation of quantifiers developed in the previous section is, of course, that it assigns the right interpretation to all the structures that were hypothesized in chapters 2 and 3. The approach makes two further predictions which are worth mentioning. The first prediction concern weak crossover effects: I show that the bijection principle of Koopman and Sportiche (1982), which is one well-known generalization about weak crossover effects, follows from the proposal of the previous section. The predictions stems from the fact that chains with lexical content in the trace position were seen to involve a variable of a higher type than that of individuals. This predicts effectively that A-bar chains, which generally do have lexical content in the trace position, and A-chains, which don't, differ with respect to the type of the variable involved. It seems natural to relate the possibility to bind a pronoun to this different in type. The other prediction I point out below concerns the question what a chain with lexical content in the trace position can be headed by.

The account of weak crossover effects is predicted in the following way. A

consequence of the system developed in the previous section is that all dependencies where the trace position has lexical content involve binding of a variable of a type other than the type e of individuals. It's shown above that the higher type is required in case the lexical content of the restrictor contains a pronoun that is bound only the trace position, as in examples like (44) (repeated from (4)) for which Engdahl's (1980) choice function proposal was adopted. For all other chains with lexical material in the trace position the motivation to use a higher type was also seen to provide an account for the the appearance of lexical material in the trace position, and therefore renders all other potential accounts using a lower type superfluous.

(44) Which friend of her$_i$'s did every student$_i$ invite?

I believe that the proposed higher type is corroborated by the following analysis of weak crossover effects. Weak crossover effects are cases where the a moved DP cannot bind a pronoun from its derived position. Consider, for example, the contrasts in (45) and (46). Only the a)-examples, where the trace position of the *wh*-word c-commands the pronoun *his* allow binding.

(45) a. Who$_i$ fed his$_i$ dog? (Wasow 1972:135)

b.$^{*?}$Who$_i$ was his$_i$ dog fed by?

(46) a. Which boy$_i$ received a postcard from his$_i$ sister?

b.$^{??}$Which boy$_i$ did his$_i$ sister send a postcard to?

I show now that the weak crossover effects in (45) and (46) is in fact predicted by

the approach outlined in the previous section. Recall from section 2.3 that A-bar movement chains require that the NP-part is present in the trace position. Therefore, the examples in (45) and (46) have LF-representations where the trace position has lexical content. For example, (47) shows the LF-representation for (46b). In (47), the variable f must range over choice functions for the chain to be interpretable. But then, it seems plausible that the operator binding this choice function variable cannot also bind a pronoun, since pronouns are plausibly of the type e of individuals.

(47) $^{??}$[Which boy] λf did his$_f$ sister send a postcard to [f, boy]

The remaining question for the account of (45) and (46) is why binding of the pronoun by the *wh*-quantifier seems possible in (45a) and (46a). Recall, though, from the discussion of (85) on page (85) that a special mechanism must be postulated to explain why a DP that doesn't seem to have moved and therefore is not by virtue of this movement the argument of a λ-predicate can nevertheless act as a binder, a problem that Heim and Kratzer (1998) also point out for the predicates approach. Whatever is the answer to this question, will allow the subject trace in (45a) and (46a) to act as a binder. Specifically the mechanism allowing the binding could be a short, string-vacuous A-movement step preceding the A-bar movement.

A-movement is predicted to obviate weak crossover because it doesn't require that the NP-part be represented in the trace position. That A-movement does obviate weak crossover is, of course, well known, and illustrated here by (48).

(48) Which girl$_i$ seemed to her$_i$ brother to be a good player.

Therefore, this account of weak cross over predicts that a pronoun can only be bound if it's c-command by a trace in an A-position. This is, in effect, equivalent to the bijection principle of Koopman and Sportiche (1982), since the lowest trace of any NP occupies an A-position. It should be mentioned, though, that the account also inherits all potential problems that of the bijection principle (see Safir 1984).

The second prediction of the approach developed in section 5.1 is a restriction to essentially chains headed by a quantificational DP. This restriction seems to maybe reflect the particular data considered in section there; namely, data involving DP-chains headed by a quantificational determiner. While this is possible, I present one argument in section 6.1 that the restriction is in some sense real.

5.3 Summary

In this chapter, I provided interpretation rules that assign the right interpretation to all the examples of the previous chapters. The main tenet of the system laid out in this section was that all determiner phrases involve the same interpretation principles. It was shown that this assumption lead to the account of Engdahl (1980), which involves a variable ranging over choice function to express the semantic dependency in a chain.

Engdahl's proposal, which was originally only intended for interrogative quantifiers, which all have existential force, is shown to raise problems when it's carried over to cardinal non-interrogative quantifiers. Both problems can be solved, however, if it's assumed that the lexical entries of quantifiers are such that two choice functions quantified are only considered different it they are pointwise different.

The choice function proposal developed in 5.1 was seen to predict the effect of the weak crossover condition.

Chapter 6

Conclusion/Outlook

This conclusion doesn't provide a summary of what was accomplished in the preceding chapters. An overview of the thesis is given in section 1.2. Rather, it contains two tentative remarks concerning the completeness of the account presented in the previous chapters for the interpretation of chains.

My aim is here to make the claim plausible that the account of the syntax-semantics interface developed in this thesis covers all cases of chains that arise. The two sections look at the two cases that, at first, seem to show that this completeness claim is wrong assuming. I present, in each case, an analysis that is compatible with the completeness claim and then argue with new facts that this new analysis is superior to the first analysis that isn't compatible with completeness. This result constitutes the strongest support for the exhaustiveness claim that seems possible. It is, obviously, always possible that the system developed proves incomplete in other respects, either ones I overlooked or ones that are discovered in the future.

The restriction of the account discussed in section 6.1 is was pointed out at

the end of section 5.2: The syntactic and semantic rules presented consider only the case of a chain headed by a quantificational DP and at least the mechanism that interprets chains with lexical content in the trace position, cannot straightforwardly account for any other case. Section 6.1 summarizes one argument and presents a second argument that only traces of type e (and maybe other non-functional types like t) arise at the level of logical form. This result is actually even stronger than the restriction just mentioned, since the interpretation of chains where the trace is of a higher type than e but has no lexical content seems possible. Therefore, the result implies that that the restriction of the account addressed is unproblematic.

Section 6.2 concerns the restriction of the account to cases of a DP-chain where the quantificational determiner is interpreted in the head position of the chain. I show first that this restriction does actually make the account of the syntax-semantics interface properties of chains easier. I then develop a new account for scope reconstruction phenomena that assumes that in scope reconstruction cases movement is actually not seen by interpretation at all, but takes place in the PF-branch of grammar. This account is seen to predict a generalization about the availability of scope reconstruction that is otherwise unexplained.

6.1 The Type of Traces

This section argues that the type of the trace position of a chain must be the type of individuals e.[1] If this claim is correct, it entails that only the two types of variables

[1]With respect to type of truth values t, there is to my knowledge no evidence that t is different from e, and the distinction drawn between the two types seems to be merely for expository purposes.

inside a trace that are proposed in chapter 5 arise. Namely, the type of individuals if the trace doesn't have any lexical content and the type of choice functions if the trace has lexical content.

A restriction on the types of traces is first proposed by Heycock (1995), Beck (1996) and Fox (1998b). While the proposals and the evidence differ, all three present evidence only for the existence of traces of type e. Heycock (1995), in effect, proposes a restriction to type e. The evidence Fox (1998b) uses to argue for the restriction, for example, involves an interaction between scope reconstruction of A-moved quantifiers and Condition C. He points out that in examples like (1) scope reconstruction is blocked (see also 1997, and Sportiche 1996). This correlation is unexpected, if the type of the variable corresponding to the A-trace could be the type of generalized quantifiers $\langle e, et \rangle$ because this type achieves the effect of narrow scope while syntactically representing the A-moved quantifier in the higher position (von Stechow 1993, Cresti 1995, Rullmann 1995, Chierchia 1995). In contrast, the interaction between Condition C and scope reconstruction is predicted if the trace position can only correspond to a variable of type e. Then, the moved quantifier must be syntactically represented in the trace position for narrow scope. Therefore, the interaction between Condition C and scope reconstruction argues for a restriction on the type of traces. The data in section 6.2, where scope reconstruction is discussed, provide another argument to assume that to achieve scope reconstruction by means of a higher type trace must be blocked.

The same seems to be true for degrees.

(1) A student of David$_i$'s seems to him$_i$ to be at the party. ($\exists \gg$ seem, seem $\gg \exists$) (Fox 1998a:(46a))

In this section, I present further empirical evidence for a restriction of the type of traces to the type e from quantifier float in Japanese. In this construction, there's a trace position associated with the moved nominal phrase in the complement position of the numeral quantifier. I claim that the type of this complement position can be either e or et resulting in two distinct interpretations, and show that the interpretations associated with the higher type et require that the moved nominal phrase be interpreted entirely in the complement position of the quantifier.

The argument relies on a new account of the partitive/cardinal ambiguity found with cardinal floating quantifiers in Japanese (Kitagawa and Kuroda 1992, Ishii 1997). Ishii (1997) observes that if the direct object that a floating quantifier is associated with occupies a VP-adjoined position, only the partitive interpretation is available: example (2) is infelicitous in a situation where only three books are salient.

(2) John-wa [urenokotta hon-o]$_i$ Mary-ni [t_1 san-satu] ageta
John$_{\text{TOP}}$ left unsold books$_{\text{ACC}}$ Mary$_{\text{DAT}}$ three-CL gave

'John gave Mary three (of the) unsold books.' (partitive, *cardinal)

Ishii (1997) also notes that examples like (3), where the nominal phrase associated with the floated quantifier occupies an IP-adjoined position, allow both a partitive and a cardinal interpretation. (Actually, it seems impossible to assess the presence of the partitive interpretation in examples like (3) if a cardinal interpretation is available,

since the former entails the latter.)

(3) [Urenokotta hon-o]$_i$ John-wa Mary-ni [t_1 san-satu] ageta
left unsold books$_{\text{ACC}}$ John$_{\text{TOP}}$ Mary$_{\text{DAT}}$ three-CL gave

'John gave Mary three (of the) unsold books.' (partitive, cardinal)

I claim that the cardinal interpretation of (3) requires reconstruction of the nominal phrase *urenokotta hon-o* to a the complement position of the quantifier *san-satu*. In support of this claim, I show that the availability of the cardinal reading correlates with the availability of reconstruction in two cases. The first is (4). In contrast to (3), (4) doesn't allow a cardinal interpretation. Since reconstruction is blocked by Condition C in (4), reconstruction is required for a cardinal interpretation.

(4) [Mary-ga$_j$ sukina hon-o]$_i$ John-wa kanozyo-ni$_j$ [t_1 san-satu] ageta
Mary$_{\text{NOM}}$ likes books$_{\text{ACC}}$ John$_{\text{TOP}}$ her$_{\text{DAT}}$ three-Cl gave

'John gave Mary three of the books she liked.' (partitive, *cardinal)

The second correlation between the availability of reconstruction is involves a parallelism of the data in (5) to the contrast between (2) and (3). Saito (1992) shows that scrambling to a VP-adjoined position cannot reconstruct for anaphor binding, while IP-adjoined scrambling can. The contrast in (5) shows that Saito's observation also holds for scrambling that strands a floated quantifier in the base position.

(5) a. John-ga [Hanako-to Mary-ni]$_i$ otagai$_i$-no hon-o ni-satu ageta
John$_{\text{NOM}}$ Hanako-and Mary-to each other$_{\text{GEN}}$ book$_{\text{ACC}}$ two-Cl gave

'John gave Hanako and Mary two books of each other's.'

b. *John-ga [otagai$_i$-no hon-o]$_j$ [Hanako-to Mary-ni]$_i$ [t_2 ni-satu] ageta
John$_{\text{NOM}}$ each other$_{\text{GEN}}$ book$_{\text{ACC}}$ Hanako-and Mary$_{\text{DAT}}$ two-Cl gave

c. [otagai$_i$-no hon-o]$_j$ John-ga [Hanako-to Mary-ni]$_i$ [t_2 ni-satu] ageta
each other$_{\text{GEN}}$ book$_{\text{ACC}}$ John$_{\text{NOM}}$ Hanako-and Mary-to two-Cl gave

Examples (2) and (3) showed that only scrambling to an IP-adjoined position allows a cardinal interpretation. Hence, the availability of reconstruction (in (5)) again correlates with the availability of the cardinal interpretation. I conclude that the cardinal interpretation requires reconstruction.

At this point, the generalization is the following: If the sister of the floated quantifier is a trace at LF, only the partitive reading is available. How does this generalization relate to the type of the trace. I claim that the cardinal interpretation of a numeral requires a complement of type $\langle e, t\rangle$, the type of first order properties. The partitive interpretation of a numeral, on the other hand, takes a complement of type e, the type of individuals. This is suggested by the English examples in (6):

(6) three $\underbrace{\text{books}}_{\langle e,t\rangle}$ vs. three of $\underbrace{\text{the books}}_{e}$

If we assume that, in Japanese as well, the difference between the cardinal and the partitive interpretation is represented by the type of the complement of the (floated) quantifier, the generalization I arrived at above follows from the restriction of traces to be of type e straightforwardly: If, at LF, the trace that's the sister of the floating quantifier is visible at LF, it must be interpreted as of type e and, therefore, only the partitive interpretation is available. Hence, the Japanese facts support the generalization that traces must be of type e.

An alternative explanation of the above generalization would be the following. Assume that floating quantifier constructions can either be generated by movement or base-generated with a pronominal element *pro* occupying the complement position of the quantifier (Kitagawa and Kuroda 1992). Furthermore, assume that when the complement of the quantifier is *pro*, only the partitive interpretation is available, and the if Q-float is generated by movement, this movement must reconstruct. Then the facts above follow, without appeal to the condition on the type of traces. But, the following facts argue that Q-float must always be generated by movement (see also Miyagawa 1989). The argument is based on the ungrammaticality of (7), where the floating quantifier is contained in a fronted VP, while the associated NP is stranded.

(7) *[Mary-ni [t_1 san-satu] age-sae]$_j$ [Urenokotta hon-o]$_i$ John-wa t_2 sita
Mary$_{\text{DAT}}$ three-CL give-even left unsold books$_{\text{ACC}}$ John$_{\text{TOP}}$ did

Notice that material stranded by VP-fronting can bind a variable in the fronted VP, as shown in (8). Therefore, the ungrammaticality of (7) argues that the relation of the associate of the floated quantifier and quantifier is not just one of binding, but one derived by movement. Under this assumption, the ungrammaticality of (7) follows from the proper binding condition (or recent proposals to derive proper binding condition effects from shortest attract; Takano 1993, Kitahara 1994, Yatsushiro 1997).

(8) [Mary-ni zibun$_i$-no hon-o san-satu age-sae]$_j$ daremo$_i$-ga t_2 sita
Mary$_{\text{DAT}}$ self$_{\text{GEN}}$ book$_{\text{ACC}}$ three-CL give-even everybody$_{\text{NOM}}$ did

'Give three books of his to Mary, everybody did.'

I conclude that trace may only be of type e. Hence, there are only two types

possible for the variable in a chain: e if the trace has no lexical content and $(et)e$, the type of choice functions if the trace has lexical content.

6.2 Scope (or Total) Reconstruction

The restriction of the interpretation procedure for chains considered in this section is that the moved quantificational determiner must be interpreted in the operator position of the chain. I have made this assumption throughout and will argue below that it allows the account to be simpler than otherwise possible. The main case where the assumption seems to be wrong are cases of scope reconstruction like (9) under the interpretation where *two* takes scope below *likely* (see (11) below).

(9) [Two people from New York]$_i$ are likely to t_i win the lottery next weekend.

That the quantification determiner of a moved DP must be interpreted in the higher of position follows from an assumption argued for in section 4.2. There I argued that the sister of a moved constituent is always interpreted as a λ-predicate. This implies that at least parts of the moved constituent are interpreted in the derived position, since otherwise this λ-predicate wouldn't have any argument. But, if any part of a moved chain is interpreted in the operator position the determiner head must be interpreted there. Hence, it follows that at least the D-head of a moved DP is always interpreted in the derived position.[2]

[2]This argument is weakened, however, that the facts below argue that, for example for VP-fronting, an interpretive mechanism must apply where all material is interpreted in the trace position.

This chapter argues for a new analysis of scope reconstruction phenomena in A-chains. Specifically, I mean by scope reconstruction cases where all material of the moved phrase seems to contribute to interpretation only in the trace position. Saito (1992), for example, uses the term *total* reconstruction for such cases. The type of scope reconstruction seems to be mainly available with A-movement. I propose that A-movement can take place in the PF-branch of the derivation and therefore be not noticed at the LF-interface. The argument for this proposal is based on the unavailability of reconstruction when the A-moved element doesn't c-commands its trace in the overt form as first observed by Barss (1986). As I show, only the PF-movement analysis straightforwardly predicts this restriction on reconstruction.

In the remainder of the introduction, I summarize test for whether scope reconstructions is possible in A-chains. May (1977) first notice scope reconstructions in A-chains looking at quantifier scope. He observed that in raising constructions the raised subject is scopally ambiguous with respect to a scope bearing element that intervenes between the trace of the raised subject and its overt position. This is illustrated by the examples (10a) and (11a), where the two different readings are paraphrased in b. and c. In (10a), the wide scope reading (paraphrase (10b)) is salient because our world knowledge about skiing competitions at the Olympic games tells us that the possibility of there being two gold medal winners in one competition is vanishingly small. In (11a), on the other hand, the narrow scope reading paraphrased in (11c) is the only one compatible with our world knowledge that, in a lottery, it's never the case that a particular individual has more than a very small chance of winning.

(10) a. [Two Germans]$_i$ are likely to t_i win the Gold Medal in this skiing race.

b. Two Germans have a good chance of winning. (two $\gg$ likely)

c.#There is a good chance that two Germans will win. (likely $\gg$ two)

(11) a. [Two people from New York]$_i$ are likely to t_i win the lottery next weekend.

b.#Two New Yorkers have a good chance of winning. (two $\gg$ likely)

c. There is a good chance that two New Yorkers will win. (likely $\gg$ two)

These following three tests for scope reconstruction which rely on grammaticality judgments are used below. The first of test uses negative polarity licensing in addition to Scope as a test for the scope reconstruction (Linebarger 1980, 1987). As is well known, a negative polarity item (henceforth NPI) must be c-commanded by negation or a downward entailing operator. What Linebarger shows is that the scope reconstruction of an A-chain can feed NPI-licensing. This is illustrated in (12). Neither in (12a) nor in (12b) does the negation c-command the NPI *anything* in the overt form. Nevertheless, the NPI in (12a) can be licensed and the NPI-licensing seems to force a scopal construal where negation takes scope over the subject. Given that there is an A-trace of the subject below negation, it seems reasonable to assume that the scope reconstruction of the subject A-chain feeds NPI-licensing in (12a).

(12) a. [A doctor who knows anything about acupuncture]$_i$ isn't t_i available.

b. *[A doctor who knows anything about acupuncture]$_i$ is t_i available.

In raising constructions as well, the narrow scope interpretation can feed NPI-licensing. This is illustrated in (13a), which contrast with example (13b), where there's no negation c-commanding the A-trace, as well as with (13c), where negation is present, but the A-trace is not c-commanded by it.

(13) a. [A doctor with any reputation]$_i$ is likely not to be t_i available.

b. *[A doctor with any reputation]$_i$ is likely to be t_i available.

c. *[A doctor with any reputation]$_i$ is t_i anxious for John not to be available.

A second test for the availability of a narrow scope interpretation using grammaticality was discovered by Burzio (1986).[3] It uses binomial *each* as test. The contrast between (14a) and (14b) shows that normally binomial *each* must be c-commanded by a distributive noun phrase in the overt form.

(14) a. The athletes demanded one translator each.

b. *One translator each welcomed the athletes.

As Burzio notes, there's one exception to this generalization which is illustrated in (15): Binomial *each* attached to the direct object can be licensed by a distributive *to*-phrase and, as Safir and Stowell (1987) point out certain other prepositional phrases.[4]

(15) The Olympic Committee assigned one translator each to the athletes.

[3]Richard Kayne (p.c.) first drew my attention to Burzio's work.

[4]As David Pesetsky (p.c.) pointed out to me, licensing of direct object binomial *each* by the following PP might itself involve a scope reconstruction of an A-chain, assuming the direct object moved from a position below the goal-PP to its surface position. See Pesetsky (1994:221) and footnote 11 on page 58 for corroborating data.

For our purposes, Burzio's most important observation is that the scope reconstruction in an A-chain can feed *each*-licensing in the pre-PP position before a prepositional phrase. This is shown in (16) for A-movement in passives, and in (17) for two-step A-movement, one step being movement to the subject position of a passive and the second step being movement to the subject position of a raising construction.

(16) a. [One translator each]$_i$ was assigned t_i to the athletes.

b. *[One translator each] gave a speech to the athletes.

(17) a. [One translator each]$_i$ is likely to t_i be assigned t_i to the athletes.

b. *[One translator each]$_i$ is likely to t_i give a speech to the athletes.

6.2.1 A PF-movement Account of Scope Reconstruction

The existence of a scope reconstruction in A-chains being established, consider the proposals that have been made to derive the narrow scope interpretation. The three proposals I know of are LF-lowering (May 1977, 1985, Chomsky 1995), the Copy Theory of movement (Wasow 1972:139, Burzio 1986, Chomsky 1993, Hornstein 1995) and Semantic Reconstruction (von Stechow 1993, Cresti 1995, Rullmann 1995, Chierchia 1995). I don't have room to summarize these proposals here in detail—in a nutshell, LF-lowering assumes that covert movement doesn't have to be to a c-commanding position in the tree and thereby can undo the effect of overt raising. The Copy Theory of movement assumes instead that a full copy of the moved phrase is left in the trace position and the interpretive component of grammar can look at this lower copy rather than the higher one. Semantic Reconstruction, finally, assumes that the

semantics of an A-chain dependency can optionally be of a higher semantic type, which leads to a scope reconstruction. All three proposals have in common that they assume that overt A-movement is followed by an invisible undoing operation as the traditional term 'scope reconstruction' for the scope reconstruction suggests. I, as already mentioned, believe that the term 'reconstruction' is misleading and propose that no undoing of movement is necessary. Rather, I propose that A-movement in the cases of a narrow interpretation, is not seen by the interpretive component of grammar because it takes place in the PF-branch of grammar.

The question that needs to be answered by any account of narrow interpretation phenomena in A-chains is the following: What is the derivation of the PF-LF-pair in (18)?[5]

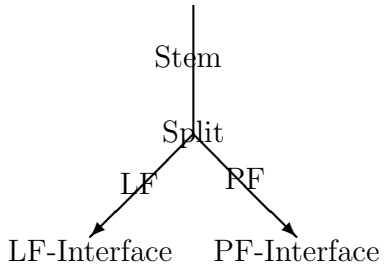
(18) PF: Two people are likely to win the lottery
LF: are likely to two people win the lottery

My answer to this question relies on the T-model of grammar (also sometimes called the Y-model or inverted Y-model) of Chomsky and Lasnik (1977), which I assume here in the form given in Chomsky (1995). The T-model embodies three partially interrelated assumptions: One, it assumes that complex representations are built up and modified by simple operations, generalized transformations, which are inherently ordered. Two, it assumes that operations can apply either having an effect on both LF and PF, or their effect can be limited to only LF, or only to PF. Three, there is link between the ordering of operations and where they have an effect: namely the

[5]The PF-representations here and in the following are given in the form before real phonology has applied.

operations that are visible only to one of LF or PF follow operations that are visible to both. All three assumptions together have the consequence that LF-PF pairs are derived by a partially ordered set of transformations that has the graphical shape shown in (19).[6]

(19) T-model (Chomsky and Lasnik 1977)



For the moment, assumption the of the T-model, that an operation can have an effect at only one of the interfaces if it applies in one of the branches, is what we need. This allows us to analyze LF-PF mismatches as operations that apply in one of the branches. It seems natural to propose that (18) is derived by PF-movement of *two people* from the embedded subject position. This is the derivation I propose generally derives scope reconstructions in A-chains. In other words, I propose that A-movement in general can optionally take place in the PF-branch of the derivation, instead of taking place in the stem. For an illustration, consider (20).

[6]The T-model incorporates an additional assumption, namely that operations which take more than one simple representations as input (in Chomsky 1995 the only operation of this type is Merge) must have an effect at both LF and PF. This assumption derives that there is exactly one Split point in the derivation of one LF-PF pair and that the branch segments of any derivation are totally ordered. This assumption, however, is not important for anything I'll say in the following.

(20) [Two people]$_x$ are likely to t_x win the lottery.

The proposal is that (20) has two possible derivations. In one derivation raising of *two people* takes place in the stem and therefore the result of raising is visible to both LF and PF. This derivation therefore gives rise to wide scope of *two people* over *likely*. Crucially, I assume that there's no way raising in the stem can be covertly undone. So, this derivation yields only the wide scope interpretation. The second derivation is one where raising of *two people* is takes place at PF, and its application is visible only to PF, but not to LF. This derivation leads to a narrow scope interpretation of *two people* below *likely*, because raising is not seen by the LF-interface.[7]

One immediate ramification the PF-movement proposal makes concerns the level at which it is verified that obligatory overt movements have indeed taken place. The PF-movement approach is incompatible with the view taken for example by Chomsky (1995) that this verification only takes place at LF. Rather, PF must be the level where the verification takes place for overt movement. At least, the morphological requirement that triggers raising in (20)—the EPP-feature if current work on the topic is to be believed—must be checked at PF. This consequence however, as far as I can see, doesn't cause any new problems; on the contrary, it now follows that the EPP must universally be satisfied overtly (cf. Chomsky 1995).

[7]It is technically conceivable, that in a derivation where raising is delayed until PF, quantifier raising applies in the LF-branch to bring about the wide scope interpretation. I don't have any evidence bearing on this possibility.

6.2.2 The Scope Freezing Generalization

In this section, I show that the PF-movement approach predicts a generalization Barss (1986) first hinted at regarding the availability of scope reconstructions, and then argue that the generalization is indeed true. This generalization is the following Scope Freezing Generalization (SFG).

(21) *SFG:* A moved quantifier QP cannot be interpreted in an A-trace position, if the trace isn't c-commanded by the overt position of QP.

The SFG blocks a scope reconstruction in cases where the trace left by A-movement is inside a a constituent that subsequently undergoes movement itself. One such case is example (22) from Barss (1986), who based solely on (22) suggests an analysis that would account for the SFG. I address Barss's account of the SFG at the end of this section.

(22) [How likely to t_{QP} address every rally]$_{wh}$ is [some politician]$_{QP}$ t_{wh}? (some≫likely, *likely≫some)

Barss (1986) presents only the example (22) in support of the SFG. In the following, I present some more. The kind of constructions that are relevant to testing the SFG are ones where subsequent A-bar movement destroys the c-command relationship between an A moved phrase and its trace, as it happened in (22). Example (23a) shows that such a stranded A-moved phrase is capable of taking scope below a c-commanding quantifier. (23b) shows that the stranded phrase can also take scope below a c-

commanding *likely*. Therefore, the lack of narrow scope of the stranded phrase in (22) and the examples in the following must be due to the lack of c-command.

(23) a. [Every journalist]$_{\forall}$ asked [how likely to $t_{\exists}$ address every rally]$_{wh}$ [some politician]$_{\exists}$ is t_{wh}. ($\forall \gg \exists$, $\exists \gg \forall$)

b. John is **likely**$_1$ to find out [how likely$_2$ to $t_{\exists}$ address every rally]$_{wh}$ [some politician]$_{\exists}$ is t_{wh}. (likely$_1$ $\gg \exists$, $\exists \gg$ likely$_1$)

The judgment in Barss's (1986) example can be sharpened by using *each*-licensing as introduced above as a test. As we see in (24), Barss's judgment now shows up as a grammaticality contrast: (24a) shows again that the scope reconstruction can feed *each* licensing. In (24b), where the SFG correctly blocks the narrow interpretation, *each* cannot be licensed. (24c), on the other hand, without *each* is grammatical, but it only has the reading with scope of *one* over *likely*.

(24) a. [One translator each]$_{QP}$ is likely to be assigned t_{QP} to the athletes.

b. *[How likely to be assigned t_{QP} to the athletes]$_{wh}$ is [one translator each]$_{QP}$ t_{wh}?

c. [How likely to be assigned t_{QP} to the athletes]$_{wh}$ is [one translator]$_{QP}$ t_{wh}?

In (22) and (24) it was *wh*-movement that destroyed the c-command relationship between the A-moved QP and its trace. The contrasts in (25) and (26) show that other types of A-bar movement, namely topicalization in (25) and *though*-raising in

(26), have the same effect.[8]

(25) a. *... and [likely to be assigned t_{QP} to the athletes]$_{top}$ [one translator each]$_{QP}$ is t_{top}.

b. ... and [likely to be assigned t_{QP} to the athletes]$_{top}$ [one translator]$_{QP}$ is t_{top}.

(26) a. *[Likely to be assigned t_{QP} to the athletes]$_{tr}$ though [one translator each]$_{QP}$ is t_{tr}, there were still complaints.

b. [Likely to be assigned t_{QP} to the athletes]$_{tr}$ though [one translator each]$_{QP}$ is t_{tr}, there were still complaints.

While in questions NPIs are independently licensed, with topicalization and though raising we can also use NPI-licensing as a test for the availability of a scope reconstruction. As the data in (27) and (28) show, the result from NPI-licensing confirms the *each*-licensing data.

(27) a. *... and [certain to be not t_{QP} available]$_{top}$, [a doctor with any reputation]$_{QP}$ was t_{top}.

b. ... and [certain to be not t_{QP} available]$_{top}$, [a doctor from cardiology]$_{QP}$ was t_{top}.

[8]Examples of VP-fronting like (25) are best if they are preceded by the same sentence with a non-fronted VP as in (i). The dots preceding all examples of VP-topicalization serve as a reminder to look at them in such a context.

(i) Martin said that one translator (each) is likely to be assigned to the athletes and, likely to be assigned to the athletes, one translator (*each) is.

(28) a. *[Certain to be not t_{QP} available]$_{tr}$ though [a doctor with any reputation]$_{QP}$ is t_{tr}, patients were waiting.

b. [Certain to be not t_{QP} available]$_{tr}$ though [a doctor from cardiology]$_{QP}$ is t_{tr}, patients were waiting.

A-movement of subjects has been argued to take place not only in raising constructions, but also with all other subjects from the VP-internal underlying subject position to the EPP-position. Hornstein (1995) and Johnson and Tomioka (1997) argue that inverse scope of the object over the subject in English transitive clauses requires a scope reconstruction of the subject chain from the VP-internal subject position to its overt position. Therefore, the SFG predicts that A-bar movement of the VP will block inverse scope in transitive clauses. In fact, this prediction seems to be a well-known fact (Fox, p.c. referring to Truckenbrodt, p.c.), though I don't know who first made this observation nor whether this has ever been made in print. The contrasts in (29) and (30) show the prediction. While (29a) and (30a) allow inverse scope, this interpretation is not available in (29b) and (30b).

(29) a. ... and [a policeman]$_{QP}$ t_{QP} stood in front of every bank. ($\forall \gg \exists$, $\exists \gg \forall$)

b. ... and [t_{QP} stand in front of every bank]$_{top}$ [a policeman]$_{QP}$ did t_{top}. ($\forall \gg \exists$, $^{*}\exists \gg \forall$)

(30) a. Though [enough of us]$_{QP}$ were t_{QP} defending every gate, the enemy broke through. (enough$\gg\forall$, $\forall\gg$enough)

b. [t_{QP} Defending every gate]$_{tr}$, though [enough of us]$_{QP}$ were t_{top}, the enemy broke through. (enough$\gg\forall$, $^{*}\forall\gg$enough)

In sum, the SFG seems be corroborated by a number of tests. I show now that the SFG is a consequence of the PF-movement analysis of narrow scope phenomena in conjunction with the three assumptions in (31). Each of these additional assumptions are independently motivated. I assume that assumption (31a), that *wh*-movement and other types of A-bar movement take place in the stem, follows from the nature of A-bar movement. The c-command condition on movement in (31b) could follow from a better understanding of movement as discussed by Chomsky (1995). I also discuss this assumption in section below in the context of a quantifier lowering analysis. Of the T-architecture, I will make use of the order it imposes on the operations in a derivation; specifically, that movement in the PF-branch takes place after movement in the stem.

(31) a. Overt A-bar movement must take place in the stem.

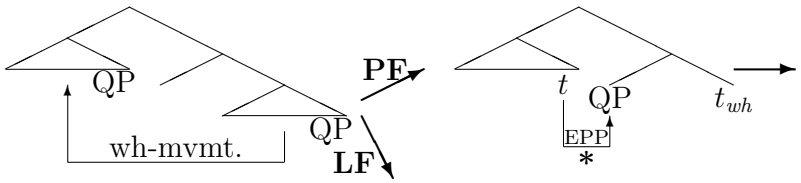
b. c-command: Movement must target a position that c-commands the moving item. (Chomsky 1995)

c. T-architecture: PF-movement must take place later than stem movement. (Chomsky and Lasnik 1977)

Consider now the derivations of a structure like (30) that would lead to narrow and wide scope respectively. First, look at a potential derivation for narrow scope in (32).

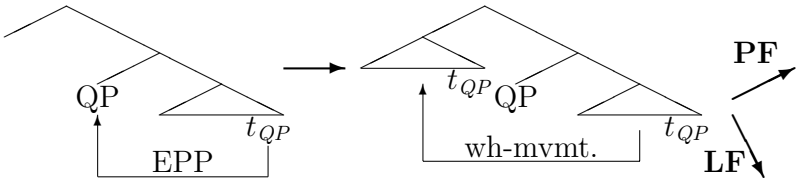
For narrow scope, raising must be delayed until PF. But, *wh*-movement must take place in the stem by assumption (31a). Assuming the T-model, the derivation in (32) is forced. It violates either the EPP or the c-command condition on movement.

(32) Failing Derivation for Narrow Scope



For wide scope, on the other hand, EPP-raising takes place in the stem, and can therefore precede *wh*-movement, as shown in (33). As the derivation in (33) shows, the EPP can be satisfied without incurring a violation of the c-command condition.

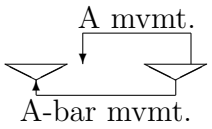
(33) Derivation for Wide Scope



I conclude that the SFG is a consequence of the PF-movement approach to narrow scope phenomena. Before I begin to consider alternatives views of scope reconstruction that actually explain the SFG, note that both the copy theory as well a semantic approaches to scope reconstruction phenomena in A chains seem to offer no perspective in accounting for SFG in an insightful way. On either account the operation bringing about the scope reconstruction is different from movement, and therefore a sensitivity of this operation to c-command would have to stipulated.

The strength of the PF-movement account, in this respect, is that the c-command sensitivity of movement is an independently argued for property of movement which carries over to PF-movement.

Now, consider two alternative explanations that might be given for the SFG. Note here, that the evidence for the SFG came entirely from examples with the structure in (34).

(34) 

The first potential explanation of the SFG was brought to my attention by David Pesetsky and Željko Bošković, and is based on two assumptions of Lasnik and Saito (1992): One, the generalized proper binding condition (GPBC), that traces must not be unbound at any point of the derivation, and two, the assumption that a control analysis is optionally possible in all raising constructions. These two assumptions force a control analysis for all examples that have the structure in (34) as Lasnik and Saito (1992:140-42) point out. control structures generally don't allow scope reconstruction, Lasnik and Saito's (1992) analysis of structures of type (34) predicts the SFG.

However, the assumptions underlying this account are at best controversial. As Takano (1993), Kitahara (1994), and Müller (1996) show, the GPBC is not correct in the form suggested by Lasnik and Saito (1992) and a empirically more accurate condition accounting for all data attributed to the GPBC follows from the general economy condition shortest attract. But, this condition allows a raising analysis for

structure such as (34). Moreover, the assumption that a control analysis is possible for raising structures misses some distinctions between the two: As Wurmbrand (1998) points out, real control can be 'imperfect' as in (35a): The PRO can refer to a plural entity that the subject is a member of. Raising on the other hand doesn't allow 'imperfect' readings, as (35b) shows.

(35) a. The mayor decided to PRO_{they} gather in the lobby.

b. *The mayor was likely to PRO_{they} gather in the lobby.

The second alternative account of the SFG is the analysis Barss (1986) gives for the example (22). He relies on a Q-lowering analysis of scope reconstruction phenomena in A-chains and proposes that the c-command condition on movement cannot only be satisfied by the landing site c-commanding the origin site, but is also satisfied if the origin site c-commands the landing site. This modified, symmetric, c-command condition allows lowering, but only to a position that is c-commanded by the origin site. Barss (1986) claims that his account blocks lowering in a structure like (34) because the landing site inside the fronted constituent here isn't c-commanded by the origin site.

It is first not clear that Barss's (1986) account actually predict the SFG. Consider a derivation, where the position inside the fronted constituent is reached by two steps of movement: The first step raises the raised subject to a position above the fronted constituent, and the second step lowers the subject into the fronted constituent. This derivation doesn't violate Barss's weakened c-command condition.

Secondly, Barss's (1986) account inherits the problems that Q-lowering has. In particular, the absence of overt lowering will need to be explained, which is not trivial in many cases: Consider e.g. Japanese scrambling: Saito (1992) shows that in Japanese a *wh*-phrase can be scrambled to a position outside of its scope domain as in (36a). He therefore argues that Japanese scrambling can be freely undone. Nevertheless, it is still impossible to scramble a phrase to a lower position in Japanese as the ungrammaticality of (36b) attests.

(36) a. dono hon-o$_i$ Masao-ga Hanako-ga t_i tosyokan-kara karidasita ka
which book Masao$_{\text{NOM}}$ Hanako$_{\text{NOM}}$ library-from checked-out Q

siritagatteiru
want-to-know

'Masao wants to know which book Hanako checked out from the library.'

b. *Hanako-ga t_i Masao-ga Taro-ni$_i$ waratta-to omowa-seta
Hanako$_{\text{NOM}}$ Masao$_{\text{NOM}}$ Taro$_{\text{DAT}}$ laughed-that believe-made

Finally, it seems to me quite likely that a strict c-command condition on movement could easily be derived as a consequence of more general principles of syntactic derivations—an issue that has received a lot of attention in recent work (cf. Chomsky 1995). The symmetric c-command condition of Barss (1986), on the other hand, seems to be a mere stipulation at this point, and it would be more natural to assume no such restriction on lowering. Therefore, I conclude that c-command is a general property of all movement. But then, the PF-movement account of scope reconstruction in A chains is the only account of the SFG left.

References

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2.83–135.

Alexiadou, Artemis, and Elena Anagnastopoulou. 1997. Toward a uniform account of scrambling and clitic doubling. In *German: Syntactic Problems – Problematic Syntax*, ed. by Werner Abraham and Elly van Gelderen, 143–161. Tübingen: Niemeyer.

Barss, Andrew. 1986. *Chains and Anaphoric Dependence. On Reconstruction and its Implications*. Cambridge: Massachusetts Institute of Technology dissertation.

Beck, Sigrid. 1996. *Wh-Constructions and Transparent Logical Form*. Universität Tübingen dissertation.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In *Proceedings of NELS 23*, 65–80, Amherst. GLSA.

——. 1995. *The Phrase Structure of Quantification Scope*. Los Angeles: University of California dissertation.

Berman, Arlene, and Michael Szamosi. 1972. Observations on sentential stress. *Language* 48.304–325.

Berman, Steve. 1991. *On the Semantics and Logical Form of Wh-Clauses*. Amherst: University of Massachusetts dissertation.

Brame, Michael K. 1968. A new analysis of the relative clause: Evidence for an interpretive theory. Unpublished Manuscript, MIT.

Bresnan, Joan. 1971. Sentence stress and syntactic transformations. *Language* 47.257–281.

——. 1972. Stress and syntax: A reply. *Language* 48.326–342.

Brody, Michael. 1979. Infinitivals, relative clauses, and deletion. Mimeograph, University of London.

Büring, Daniel. 1995. The great scope inversion conspiracy. In *Proceedings of SALT 5*, ed. by Mandy Simons and Teresa Galloway, 37–53, Cornell. CLC Publications.

——. 1996. *The 59th-Street Bridge Accent*. Universität Tübingen dissertation.

——. 1998. Focus and topic in a complex model of discourse. Manuscript, Cologne University, Germany.

Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Kluwer.

Carlson, Greg N. 1977. Amount relatives. *Language* 58.520–542.

Cheng, Lisa Lai-Shen. 1991. *On the Typology of Wh-Questions*. Cambridge: Massachusetts Institute of Technology dissertation.
Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1.181–234.
——. 1995. *Dynamics of Meaning*. University of Chicago Press.
Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.
——. 1972a. Some empirical issues in the theory of transformational grammar. In *Goals of Linguistic Theory*, ed. by Stanley Peters, Englewood Cliffs, New Jersey: Prentice Hall. (reprinted in Chomsky 1972b).
——. 1972b. *Studies on Semantics in Generative Grammar*. The Hague, The Netherlands: Mouton.
——. 1977. On wh-movement. In *Formal Syntax*, ed. by Peter Culicover, Tom Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.
——. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Berlin: Mouton de Gruyter.
——. 1986. *Barriers*. Cambridge, Massachusetts: MIT Press.
——. 1993. A minimalist program for linguistic theory. In *The View from Building 20, Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Ken Hale and Jay Keyser, 1–52. MIT Press.
——. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
——, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8.425–504.
Chung, Sandra, James McCloskey, and Bill Ladusaw. 1995. Sluicing and logical form. *Natural Language Semantics* 3.239–282.
Church, Alonzo. 1932. A set of postulates for the foundation of logic. *Annals of Mathematics (2)* 33.346–366.
——. 1933. A set of postulates for the foundation of logic (second paper). *Annals of Mathematics (2)* 34.839–864.
Cinque, Giugliemo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24.239–297.
Cooper, Robin. 1983. *Quantification and Syntactic Theory*. Dordrecht, The Netherlands: Reidel.
Cormack, Anabel. 1984. VP anaphora: Variables and scope. In *Varieties of Formal Semantics*, ed. by Fred Landman and F. Veltman, Dordrecht: Reidel.
Cresti, Diana. 1995. Extraction and reconstruction. *Natural Language Semantics* 3.79–122.
——. 1997. Some considerations on *wh*-decomposition and unselective binding. In *Proceedings of the Tübingen Workshop on Reconstruction*, 147–176. Universität Tübingen, Germany.
Curry, Haskell B. 1930. Grundlagen der kombinatorischen Logik. *American Journal of Mathematics* 52.509–536.
——, and Robert Feys. 1958. *Combinatory Logic, Volume I*. Amsterdam: North-Holland.
Dayal, Veneeta. 1996. *Locality in Wh Quantification*. Dordrecht, The Netherlands: Kluwer.

Déprez, Viviane. 1990. *On the Typology of Syntactic Positions and the Nature of Chains: Move α to the Specifier of Functional Projections*. Cambridge: Massachusetts Institute of Technology dissertation.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: MIT Press.

Dowty, David R. 1992. 'Variable-free' syntax, variable-binding syntax, the natural deduction Lambek calculus, and the crossover constraint. In *Proceedings of WCCFL 11*, 161–176, Stanford, California. Stanford University, Center for the Study of Language and Information.

Engdahl, Elisabeth. 1980. *The Syntax and Semantics of Questions in Swedish*. Amherst: University of Massachusetts dissertation.

——. 1986. *Constituent Questions*. Dordrecht, The Netherlands: Reidel.

Evans, Frederic. 1988. Binding into anaphoric verb phrases. In *Proceedings of ESCOL 5*, ed. by Joyce Powers and Kenneth de Jong, 122–129, Columbus. Ohio State University, Working Papers in Linguistics.

Fiengo, Robert, and Robert May. 1994. *Indices and Identity*. Cambridge, Massachusetts: MIT Press.

Fodor, Janet Dean, and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5.355–398.

Fox, Danny. 1995a. Economy and scope. *Natural Language Semantics* 3.283–341.

——. 1995b. Condition C effects in ACD. In *Papers on Minimalist Syntax, MITWPL 27*, ed. by Rob Pensalfini and Hiroyuki Ura, 105–119. Cambridge, Massachusetts: MIT Working Papers in Linguistics.

——. 1996. Where does binding theory *really* apply? Presentation at the LF reading group, MIT, Cambridge, Massachusetts.

——. 1997. Reconstruction, variable binding and the interpretation of chains. In *Proceedings of the Tübingen Workshop on Reconstruction*, 71–118. Universität Tübingen, Germany.

——. 1998a. *Economy and Semantic Interpretation*. Cambridge: Massachusetts Institute of Technology dissertation. (in progress).

——. 1998b. Reconstruction, variable binding and the interpretation of chains. *Linguistic Inquiry* (to appear).

——. 1998c. Locality in variable binding. In *Is the Best Good Enough?*, ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 129–144. Cambridge, Massachusetts: MIT Press and MIT Working Papers in Linguistics.

——, and Jon Nissenbaum. 1998. A-bar traces and the copy theory of movement. Notes, MIT, Cambridge, Massachusetts.

Frege, Gottlob. 1884. *Grundlagen der Arithmetik*. Breslau.

Freidin, Robert. 1986. Fundamental issues in the theory of binding. In *Studies in the Acquisition of Anaphora, Volume I*, ed. by Barbara Lust, 151–188. Dordrecht: Reidel.

Gabbay, Dov M., and J.M.E. Moravscik. 1974. Branching quantifiers, English and Montague Grammar. *Theoretical Linguistics* 1.139–157.

Geach, P.T. 1964. Referring expressions again. *Analysis* 24. (reprinted as Geach 1972.

—— 1972. Theory of reference and syntax. In *Logic Matters*, chapter 3.4, 97–102. Berkeley and Los Angeles, California: University of California Press.

Grinder, J., and Paul Postal. 1971. Missing antecedents. *Linguistic Inquiry* 2.269–312.

Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies in the Semantics of Questions and the Pragmatics of Answers*. The Netherlands: University of Amsterdam dissertation.

Grosu, Alexander, and Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6.125–170.

Hackl, Martin, and Jon Nissenbaum. 1998. Variable modal force in *for*-infinitival relative clauses. Manuscript, Massachusetts Institute of Technology, Cambridge.

Hagstrom, Paul. 1998. *Decomposing Questions*. Cambridge, Massachusetts: Massachusetts Institute of Technology dissertation. (in progress).

Hamblin, C.L. 1958. Questions. *Australasian Journal of Philosophy* 36.159–168.

—— 1973. Questions in Montague English. *Foundations of Language* 10.41–53. Reprinted in Partee (1976).

Hardt, Daniel. 1992. VP ellipsis and semantic identity. In *Proceedings of SALT II*, ed. by Chris Barker and David Dowty, 145–161, Columbus. Ohio State University, Working Papers in Linguistics.

Heim, Irene. 1987. Where does the definiteness restriction apply? Evidence from the definiteness of variables. In *The Representation of (In)definiteness*, ed. by Eric Reuland and Alice ter Meulen, chapter 2, 21–42. Cambridge, Massachusetts: MIT Press.

——. 1997a. Predicates or formulas? Evidence from ellipsis. In *Proceedings of SALT VII*, ed. by Aaron Lawson and Eun Cho, 197–221. Ithaca, New York: CLC Publications.

——. 1997b. Predicates or formulas? Evidence from ellipsis. Presentation at LF reading group, MIT.

——, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Hepple, Mark. 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach*. Scotland: University of Edinburgh dissertation.

——. 1992. Command and domain constraints in categorial theory of binding. In *Proceedings of the Amsterdam Colloquium*, University of Amsterdam, The Netherlands.

Heycock, Caroline. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26.547–570.

Higginbotham, James. 1983. Logical form, binding, and nominals. *Linguistic Inquiry* 14.679–708.

——, and Robert May. 1980. Questions, quantifiers, and crossing. *The Linguistic Review* 1.41–80.

Hindley, J.R., B. Lercher, and J.P. Seldin. 1972. *Introduction to Combinatory Logic*. Cambridge University Press.

Hintikka, Jaakko. 1974. Quantifiers vs. quantification theory. *Linguistic Inquiry* 5.153–177.

Höhle, Tilman. 1979. 'Normalbetonung' und 'normale Wortstellung': Eine pragmatische Explikation. *Leuvense Bijdragen* 68.385–437.
——. 1982. Explikation für "normale Betonung" und "normale Wortstellung". In *Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*, ed. by Werner Abraham, 75–153. Tübingen: Narr.
Hornstein, Norbert. 1995. *The Grammar of Logical From: From GB to Minimalism*. Cambridge: Blackwell.
Ishii, Yasuo. 1997. Scrambling and the weak-strong distinction in Japanese. In *Is the Logic Clear?* Papers in Honor of Howard Lasnik, UConn WPL 8, ed. by Joeng-Seok Kim, Satoshi Oku, and Sandra Stjepanović, 89–112. Storrs: University of Connecticut.
Jackendoff, Ray. 1968. An interpretive theory of pronouns and reflexives. PEGS paper number 27.
——. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.
——. 1977. *$\bar{X}$-bar Syntax: A Stude of Phrase Structure*. Cambridge, Massachusetts: MIT Press.
Jacobson, Pauline. 1992. Antecedent contained deletion in a variable free semantics. In *Proceedings of SALT II*, 193–213, Columbus. Ohio State University, Working Papers in Linguistics.
——. 1993. Bach-Peters sentences in a variable-free semantics. In *Proceedings of the Eigth Amsterdam Colloquium*.
——. 1994. *i*-within-*i* effects in a variable free semantics and a categorial syntax. In *Proceedings of the Ninth Amsterdam Colloquium*. University of Amsterdam, The Netherlands.
——. 1998a. ACE and pied-piping: Evidence for a variable-free semantics. Presentation at SALT 8, MIT.
——. 1998b. Towards a variable-free semantics. *Linguistics and Philosophy* (to appear).
Jayaseelan, K.A. 1990. Incomplete VP deletion and gapping. *Linguistic Analysis* 20.64–81.
Johnson, Kyle. 1996. When verb phrases go missing. *GLOT* 2.3–9.
——, and Satoshi Tomioka. 1997. Lowering and mid-size clauses. In *Proceedings of the Tübingen Workshop on Reconstruction*, 177–198. Universität Tübingen, Germany.
Karttunen, Lauri. 1977. The syntax and semantics of questions. *Linguistics and Philosophy* 1.1–44.
Keenan, Edward. 1971. Names, quantifiers, and a solution to the sloppy-identity problem. *Papers in Linguistics* 4.211–232.
Kennedy, Christopher. 1994. Argument contained ellipsis. Linguistics Research Center Report LRC-94-03, University of California, Santa Cruz.
Kitagawa, Yoshihisa, and S.-Y. Kuroda. 1992. Passive in Japanese. Manuscript, University of Rochester and UCSD.
Kitahara, Hisatsugu. 1994. Restricting *ambiguous* rule-application. In *Formal Approaches to Japanese Linguistics I*, ed. by Masatoshi Koizumi and Hiroyuki Ura, 179–209, Cambridge. MIT Working Papers in Linguistics.

Klein, Ewan, and Ivan A. Sag. 1985. Type-driven translation. *Linguistics and Philosophy* 8.163–201.

Koizumi, Masatoshi. 1994. Layered specifiers. In *Proceedings of NELS 24*, ed. by Mercè Gonzàlez, 255–269, Amherst. GLSA.

Koopman, Hilda, and Dominique Sportiche. 1982. Variables and the bijection principle. *The Linguistic Review* 2.139–160.

Kratzer, Angelika. 1991. The representation of focus. In von Stechow and Wunderlich (1991), chapter 40, 825–834.

——. 1995. Scope or pseudoscope? Are there wide-scope indefinites? Manuscript, University of Massachusetts, Amherst.

——. 1998. Scope or pseudoscope? Are there wide-scope indefinites? In *Events in Grammar*, ed. by Susan Rothstein, Dordrecht: Kluwer.

Kuno, Susumo. 1997. Binding theory in the minimalist program. Manuscript, Harvard University, Cambridge, Massachusetts.

Lahiri, Utpal. 1991. *Embedded Interrogatives and Predicates that Embed them*. Cambridge: Massachusetts Institute of Technology dissertation.

Lakoff, George. 1968. Pronouns and reference. Mimeograph, Indiana University Linguistics Club, University of Indiana, Bloomington (published as Lakoff 1976).

——. 1972. The global nature of the Nuclear Stress Rule. *Language* 48.285–303.

——. 1976. Pronouns and reference. In *Notes from the Linguistic Underground, Syntax and Semantics, Volume 7*, ed. by James D. McCawley, chapter 16, 275–335. New York: Academic Press.

Langacker, Ronald W. 1969. Pronominalization and the chain of command. In *Modern Studies in English*, ed. by David A. Reibel and Sanford A. Schane, chapter 12, 160–186. Englewood Cliffs, New Jersey: Prentice Hall.

Lappin, Shalom. 1984. VP anaphora, quantifier scope, and logical form. *Linguistic Analysis* 13.273–315.

Larson, Richard K., and Robert May. 1990. Antecedent containment or vacuous movement: Reply to Baltin. *Linguistic Inquiry* 21.103–122.

Larson, Richard K., and Gabriel Segal. 1995. *Knowledge of Meaning: An Introduction to Semantic Theory*. Cambridge, Massachusetts: MIT Press.

Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2.1–22.

——. 1995. A note on pseudogapping. In *Papers on Minimalist Syntax, MITWPL 27*, ed. by Rob Pensalfini and Hiroyuki Ura, 143–164. Cambridge: MIT Working Papers in Linguistics.

——, and Mamoru Saito. 1992. *Move α: Conditions on Its Application and Output*, volume 22 of *Current Studies in Linguistics*. Cambridge: MIT Press.

Lebeaux, David. 1988. *Language Acquisition and the Form of Grammar*. Amherst: University of Massachusetts dissertation.

——. 1992. Relative clauses, licensing, and the nature of the derivation. In *Perspectives on Phrase Structure: Heads and Licensing*, ed. by Susan Rothstein and Margaret Speas, volume 25 of *Syntax and Semantics*, 209–239. New York: Academic Press.

——. 1995. Where does the binding theory apply? In *University of Maryland Working Papers in Linguistics 3*, 63–88. College Park: University of Maryland.

——. 1998. Where does the binding theory apply? (version 2). Technical Report 98-044, NEC Research Institute, Princeton, New Jersey.

Lechner, Winfried. 1998. On semantic and syntactic reconstruction. *Studia Linguistica*.

Lees, Robert B. 1960. *The Grammar of English Nominalizations*. The Hague: Mouton.

—— 1961. The constituent structure of noun phrases. *American Speech* 36.159–168.

Linebarger, Marcia. 1980. *The Grammar of Negative Polarity*. Cambridge: Massachusetts Institute of Technology dissertation. Distributed by MITWPL.

——. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10.325–387.

Lobeck, Anne. 1992. *Ellipsis*. Oxford: Oxford University Press.

Matthewson, Lisa. 1998. On the interpretation of wide-scope indefinites. Manuscript, Massachusetts Institute of Technology, Cambridge (revised form to appear in *Natural Language Semantics*).

May, Robert. 1977. *The Grammar of Quantification*. Cambridge: Massachusetts Institute of Technology dissertation.

——. 1985. *Logical Form: Its Structure and Derivation*. Cambridge: MIT Press.

McCawley, John. 1976. Notes on Jackendoff's theory of anaphora. *Linguistic Inquiry* 7.319–341.

Merchant, Jason. 1998a. On the extent of trace deletion in ACD. Manuscript, University of Utrecht, The Netherlands and University of California, Santa Cruz.

——. 1998b. Guess what i found and where. In *Proceedings of WCCFL 18*, Stanford, California. Center for the Study of Language and Information.

Miyagawa, Shigeru. 1989. *Structure and Case Marking in Japanese*. Number 22 in Syntax and Semantics. New York: Academic Press.

Moltmann, Friederike, and Anna Szabolcsi. 1994. Scope interactions with pair-list quantifiers. In *Proceedings of NELS 24*, ed. by Mercé González, 381–395, Amherst. University of Massachusetts, GLSA.

Müller, Gereon. 1993. *On Deriving Movement Type Asymmetries*. Tübingen: Universität Tübingen dissertation.

——. 1996. *Incomplete Category Fronting*. Habilitationsschrift. Universität Tübingen.

——. 1998. *Incomplete Category Fronting*. Dordrecht, The Netherlands: Kluwer.

Munn, Alan. 1994. A minimalist account of reconstruction asymmetries. In *Proceedings of NELS 24*, ed. by Mercè Gonzàlez, 397–410, Amherst. University of Massachusetts, GLSA.

Nissenbaum, Jon. 1998. Movement and derived predicates: Evidence from parasitic gaps. In *The Interpretive Tract, MITWPL 25*, ed. by Uli Sauerland and Orin Percus, 247–295. Cambridge: MIT Working Papers in Linguistics.

Partee, Barbara H. (ed.) 1976. *Montague Grammar*. New York: Academic Press.

Pesetsky, David. 1982. *Paths and Categories*. Cambridge: Massachusetts Institute of Technology dissertation.

——. 1989. *Wh*-in-situ: Movement and unselective binding. In *The Representation of (In)definiteness*, ed. by Eric Reuland and A. ter Meulen, 98–129. Cambridge: MIT Press.

——. 1994. *Zero Syntax: Experiencers and Cascades*. Cambridge: MIT Press.
Pica, Pierre, and William Snyder. 1994. Weak crossover, quantifer scope and minimalism. In *Proceedings of WCCFL 13*, 334–349, Stanford, California. Center for the Study of Language and Information, Stanford University.
Quine, Willard van Orman. 1960. Variables explained away. *Proceedings of the American Philosophical Society* 104. (reprinted as Quine 1995).
——. 1995. Variables explained away. In *Selected Logic Papers,* Enlarged Edition, 227–235. Cambridge, Massachusetts: Harvard University Press.
Reinhart, Tanya. 1976. *The Syntactic Domain of Anaphora*. Cambridge: Massachusetts Institute of Technology dissertation.
——. 1981. A second COMP position. In *The Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, ed. by Adriana Belletti, Luciana Brandi, and Luigi Rizzi, 517–557. Pisa, Italy: Scuola Normale Superiore di Pisa.
——. 1994. Wh-in-situ in the framework of the minimalist program. OTS Working Papers 94/03, Utrecht University, Utrecht, The Netherlands.
——. 1997. Quantifier scope: How the labor is divided between QR and choice functions. *Linguistics and Philosophy* 20.335–397.
Richards, Norvin. 1997. *What moves where when in which language?*. Cambridge: Massachusetts Institute of Technology dissertation.
Riemsdijk, Henk van, and Edwin Williams. 1981. NP-structure. *The Linguistic Review* 1.171–217.
Ristad, Sven Eric. 1990. *Computational Structure of Human Language*. Cambridge: Massachusetts Institute of Technology dissertation.
Rizzi, Luigi. 1990. *Relativized Minimality*. Number 16 in Linguistic Inquiry Monographs. Cambridge: MIT Press.
Romero, Maribel. 1996. The (cor)relation of scope reconstruction and connectivity effects. Manuscript, University of Massachusetts, Amherst.
——. 1997. Problems for a semantic account of scope reconstruction. In *Proceedings of the Tübingen Workshop on Reconstruction*, 119–146. Universität Tübingen.
Rooth, Mats. 1985. *Association with Focus*. Amherst: University of Massachusetts dissertation.
——. 1992a. A theory of focus interpretation. *Natural Language Semantics* 1.75–116.
——. 1992b. Ellipsis redundancy and reduction redundancy. In *Proceedings of the Stuttgart Ellipsis Workshop*, ed. by Steve Berman and Arild Hestvik. Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29, IBM Germany, Heidelberg.
——. 1996. Focus. In *The Handbook of Contemporary Semantic Theory*, ed. by Shalom Lappin, 271–297. Oxford: Blackwell.
Ross, John R. 1967. On the cyclic nature of pronominalization in English. In *To Honor Roman Jakobson*, 1669–1682. The Hague, The Netherlands: Mouton. (reprinted as Ross 1969a).
—— 1969a. On the cyclic nature of pronominalization in English. In *Modern Studies in English*, ed. by David A. Reibel and Sanford A. Schane, chapter 13, 187–200. Englewood Cliffs, New Jersey: Prentice Hall.

Ross, John Robert. 1969b. Guess who? In *CLS 5*, ed. by Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252–278, Chicago. Chicago Linguistics Society.

Rudin, Catherine. 1988. On multiple questions and multiple wh fronting. *Natural Language & Linguistic Theory* 6.445–501.

Rullmann, Hotze. 1995. *Maximality in the Semantics of Wh-Constructions*. Amherst: University of Massachusetts dissertation.

——, and Sigrid Beck. 1997. Reconstruction and the interpretation of *which*-phrases. In *Proceedings of the Tübingen Workshop on Reconstruction*, 215–249. Universität Tübingen, Germany.

Ruys, Eddie. 1993. *The Scope of Indefinites*. Utrecht University dissertation.

Safir, Ken. 1984. Multiple variable binding. *Linguistic Inquiry* 15.603–638.

——. 1998. Reconstruction and bound anaphora: Copy theory without deletion at LF. Manuscript, Rutgers University, New Brunswick, New Jersey.

——, and Tim Stowell. 1987. Binominal each. In *Proceedings of NELS 18*, 426–450. Amherst: GLSA.

Sag, Ivan. 1976. *Deletion and Logical Form*. Cambridge: Massachusetts Institute of Technology dissertation.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1.69–118.

Sato-Zhu, Eriko. 1996. *The Logical Intepretation of English and Japanese Sentences*. Stony Brook: State University of New York dissertation.

Sauerland, Uli. 1998. Plurals, derived predicates and reciprocals. In *The Interpretive Tract, MITWPL 25*, ed. by Uli Sauerland and Orin Percus, 177–204. Cambridge: MIT Working Papers in Linguistics.

Schachter, Paul. 1973. Focus and relativization. *Language* 49.19–46.

Schönfinkel, M. 1924. Über die Bausteine der mathematischen Logik. *Mathematische Annalen* 92.305–316.

Schütze, Carson. 1995. Pp attachment and argumenthood. In *Papers on Language Processing and Acquisition, MITWPL 26*, ed. by Carson Schütze, Jennifer Ganger, and Kevin Broihier, 95–151. Cambridge: MIT Working Papers in Linguistics.

Schwarz, Bernhard. 1993. Gewisse Fälle eingebetteter Fragesätze. Master's thesis, Universität Tübingen, Germany.

Schwarzschild, Roger. 1994. The contrastiveness of associated foci. Manuscript, Hebrew University of Jerusalem, Israel.

——. 1998. Givenness and optimal focus. *Natural Language Semantics* (to appear in revised form).

Selkirk, Elizabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In *The Handbook of Phonological Theory*, ed. by John Goldsmith, chapter 16, 550–569. London: Blackwell.

Sharvit, Yael. 1996a. Functional dependencies and indirect binding. In *Semantics and Linguistic Theory VI*, ed. by Teresa Galloway and Justin Spence, 227–244, Ithaca, New York. Cornell University, CLC Publications.

——. 1996b. *The Syntax and Semantics of Functional Relative Clauses*. New Brunswick, New Jersey: Rutgers University dissertation.
——. 1998. Possessive wh-expressions and reconstruction. In *Proceedings of NELS 28*, Amherst. University of Massachusetts, GLSA. (to appear).
Sportiche, Dominique. 1996. A-reconstruction and constituent structure. Handout, University of California, Los Angeles.
Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, Massachusetts: MIT Press.
Sternefeld, Wolfgang. 1998. The semantics of reconstruction and connectivity. Arbeitspapier 97, SFB 340, Universität Tübingen and Universität Stuttgart, Germany.
Szabolcsi, Anna. 1987. Bound variables in syntax (are there any?). In *Proceedings of the Amsterdam Colloquium*, 331–351. University of Amsterdam, The Netherlands.
Tada, Hiroaki. 1993. *A/A-Bar Partition in Derivation*. Cambridge: Massachusetts Institute of Technology dissertation.
Takano, Yuji. 1993. Minimalism and proper binding. Manuscript, UC Irvine.
——. 1994. Unbound traces and indeterminacy of derivation. In *Current Topics in English and Japanese*, ed. by Masura Nakamura, 229–253. Tokyo: Hituzi Syoboo.
——. 1995. Predicate fronting and internal subjects. *Linguistic Inquiry* 26.327–340.
Tancredi, Christopher. 1992. *Deletion, Deaccenting and Presupposition*. Cambridge: Massachusetts Institute of Technology dissertation.
Tarski, A. 1936. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica* 1.261–405.
Truckenbrodt, Hubert. 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Cambridge: Massachusetts Institute of Technology dissertation.
Vergnaud, Jean Roger. 1974. *French Relative Clauses*. Cambridge: Massachusetts Institute of Technology dissertation.
von Stechow, Arnim. 1993. Die Aufgaben der Syntax (the objectives of syntax). In *Syntax: Ein internationales Handuch der zeitgenössischen Forschung* (Syntax: An International Handbook of Contemporary Research), ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 1–88. Berlin: de Gruyter.
——. 1996. Some remarks on choice functions and LF-movement. In *Proceedings of the Konstanz Workshop "Reference and Anaphorical Relations"*, ed. by Klaus von Heusinger and Urs Egli. Constance: Universität Konstanz.
——, and Dieter Wunderlich (eds.) 1991. *Semantik: Ein internationales Handuch der zeitgenössischen Forschung* (Semantics: An International Handbook of Contemporary Research). Berlin: de Gruyter.
Wasow, Thomas. 1972. *Anaphoric Relations in English*. Cambridge: Massachusetts Institute of Technology dissertation.
——. 1979. *Anaphora in Generative Grammar*. Ghent: Story Scientia.

Webelhuth, Gert. 1995. X-bar theory and case theory. In *Government and Binding Theory and the Minimalist Program*, ed. by Gert Webelhuth, chapter 1, 15–96. Cambridge: Blackwell.

Whitehead, Alfred N., and Bertrand Russell. 1910. *Principia Mathematica. Volume I*. Cambridge, Great Britain.

Wilkinson, Karina. 1991. *Studies in the Semantics of Generic Noun Phrases*. Amherst: University of Massachusetts dissertation.

Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry* 8.101–139.

Winter, Yoad. 1997. Dependency and distributivity of plural definites. Manuscript, Utrecht University.

Wold, Dag. 1995. Identity in ellipsis: Focal structure and phonetic deletion. Manuscript, MIT.

——. 1996. Long distance selective binding: The case of focus. In *Proceedings of SALT 6*, ed. by Teresa Galloway and Justin Spence, 311–328, Ithaca, New York. Cornell University, CLC Publications.

——. 1998. Cambridge: Massachusetts Institute of Technology dissertation. (in progress).

Wurmbrand, Susi. 1998. Downsizing infinitives. In *The Interpretive Tract, MITWPL 25*, ed. by Uli Sauerland and Orin Percus, 141–175. MIT Working Papers in Linguistics.

Yatsushiro, Kazuko. 1997. VP-scrambling in Japanese. In *Is the Logic Clear?* Papers in Honor of Howard Lasnik. UConn WPL 8, ed. by Joeng-Seok Kim, Satoshi Oku, and Sandra Stjepanović, 325–338. Storrs: University of Connecticut, Working Papers in Linguistics.

Zubizaretta, Maria Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge, Massachusetts: MIT Press.