Anaphora Resolved

Floris Roelofsen

Anaphora Resolved

ILLC Dissertation Series DS-2008-09

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@science.uva.nl
homepage: http://www.illc.uva.nl/

Anaphora Resolved

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op donderdag 9 oktober 2008, te 12.00 uur

door

Floris Roelofsen

geboren te Amsterdam

Promotiecommissie

Promotor:
prof.dr. J.A.G. Groenendijk

Overige leden:
prof.dr. D. Fox
prof.dr. A. Szabolcsi
prof.dr. E.J. Reuland
prof.dr. H.J. Bennis
prof.dr.ir. R.J.H. Scha
prof.dr. M.J.B. Stokhof
dr. M.D. Aloni

Faculteit der Geesteswetenschappen

Printed and bound by PrintPartners Ipskamp, Enschede.

ISBN: 978–90–5776–183–6

Notre tête est ronde pour permettre à la pensée de changer de direction.

Francis Picabia

Dedicated to the memory of Tanya Reinhart.

Contents

Preface

Preview

One of the central questions in linguistics is how, or to what extent, the meaning of complex linguistic objects (e.g., sentences) can be derived from the meaning of their more basic constituents (e.g., words) and the way in which these basic constituents are put together.[1]

A major stumbling block for this enterprise is that the meaning of certain words is highly context-dependent. For example, it is impossible to define the meaning of personal pronouns like *he*, *she*, and *it* without referring to the context in which these expressions are used. Such expressions are called *anaphoric expressions*, or simply *anaphora*, and will be the main topic of this dissertation.

Anaphoric expressions have received a great deal of attention, both from linguists and from philosophers. To illustrate why they are so fascinating, let me consider a straightforward definition of the meaning of *he* and point out several facts which it fails to capture. The following definition is probably more or less what would spring to mind first, and can indeed be found in many dictionaries:[2]

(0.1) *he*: male person or animal previously referred to.

This works for many cases. For instance, in (0.2a) *he* can be taken to refer to John, in (0.2b) *him* can be taken to refer to Max, and in (0.2c) *his* can be taken to refer to the lion which is also referred to by the subject noun phrase of the sentence.

[1]Note that this question is not only central in linguistics, but also in other sciences. For instance, physicists try to explain how, or to what extent, the characteristics of complex physical objects (e.g., molecules) can be derived from the characteristics of their basic constituents (i.e., atoms) and the way in which these basic constituents are put together.

[2]This particular formulation is taken from the Oxford ESL dictionary by A.S. Hornby and C.A. Ruse, Oxford University Press, 1999, p.278.

(0.2) a. John says that he didn't sleep.
b. Max hopes that Mary likes him.
c. The lion devoured his prey.

But the definition in (0.1) does certainly not capture all cases. For example, consider (0.3a) and its possible reading in (0.3b). On this reading, *he* is not interpreted as referring to someone in particular, but rather, it seems, as a variable ranging over a domain of several individuals, just as variables in logic do.

(0.3) a. Every student hopes that he will pass the exam.
b. Every student x hopes that x will pass the exam.

A similar interpretation is possible if *he* occurs in a question:

(0.4) a. Which student thinks that he passed the exam?
b. Which student x thinks that x passed the exam?

Other cases which are not captured by the definition in (0.1) are ones in which *he* does refer, but not to an individual that has been referred to *previously*. For example, in (0.5a) *he* may be taken to refer to Max, even though Max has not been referred to previously. A similar remark applies to (0.5b). Such cases are sometimes called *backward anaphora*.

(0.5) a. When Mary finally kissed him, Max was very happy.
b. Before he left the house, Fred closed all the windows.

Another important shortcoming of the definition in (0.1) is that it does not capture the fact that the interpretation of pronouns is systematically restricted. For example, in (0.6a) *him* cannot be taken to refer to John, even though John has been referred to previously, and a similar remark applies to (0.6b).

(0.6) a. John hates him.
b. John bought him a present.

There are also restrictions on the interpretation of pronouns as variables. For example, (0.7a) cannot be interpreted as in (0.7b), and (0.8a) cannot be interpreted as in (0.8b).

(0.7) a. Every student bought him a present.
b. Every student x bought x a present.

(0.8) a. Which student does he like best?
b. Which student x does x like best?

Finally, there are certain restrictions on backward anaphora. For example, in (0.9a) *he* cannot be taken to refer to Max, and in (0.9b) *he* cannot be taken to refer to Fred.

(0.9) a. He was very happy when Mary finally kissed Max.
b. He closed all the windows, before Fred left the house.

Pronouns are by far the most widely studied kind of anaphora. But there are other anaphoric mechanisms as well. One that will receive considerable attention in this dissertation is *verb phrase ellipsis* (VP ellipsis for short). This kind of anaphora is exemplified in (0.10).

(0.10) a. Sue went to school after Mary did.
b. Sue went to school after Mary went to school.

(0.10a) is most naturally interpreted as in (0.10b), i.e., the auxiliary in the subordinate clause, *did*, is interpreted as *went to school*. Of course, this interpretation is highly context-dependent, just like the interpretation of pronouns.

In fact, pronominal anaphora and VP ellipsis behave alike in many ways (and indeed, one of the claims that will be defended in this dissertation is that they should receive a unified treatment). For instance, example (0.11) shows that backward anaphora is possible with VP ellipsis, just as with pronouns, while (0.12) shows that this mechanism is restricted in certain ways, again, just as in the case of pronouns.

(0.11) a. After Mary did, Sue went to school as well.
b. If nobody else does, you must ask a question yourself.

(0.12) a. Sue did after Mary went to school.
b. You must if nobody else asks a question.

Pronominal anaphora and VP ellipsis also interact in interesting ways. For example, as illustrated in (0.13), VP ellipsis of a verb phrase which contains a pronoun often gives rise to a particular kind of ambiguity. The second clause in (0.13) can be interpreted as in (0.13a), but also as in (0.13b).

(0.13) John talks about his children all the time, and Fred does too.
a. Fred also talks about John's children all the time.
b. Fred also talks about his own children all the time.

Again, there are interesting restrictions on this kind of ambiguity. For instance, if we consider a verb phrase which contains not one, but two pronouns, we would expect to get at least four possible interpretations. But example (0.14) shows that this expectation is not always born out. In particular, (0.14) cannot be interpreted as in (0.14d).

(0.14) Max said that he called his mother, and Bob did too.
a. Bob also said that Max called Max's mother.
b. Bob also said that Bob called Bob's mother.
c. Bob also said that Bob called Max's mother.
d. Bob also said that Max called Bob's mother.

These, then, are some of the puzzling facts that have to be explained. A more systematic and comprehensive presentation of the data will follow of course. The purpose here is merely to illustrate why anaphora have fascinated so many generations of linguists and philosophers, and in particular why the present dissertation should make for interesting reading.

The dissertation is divided into two parts. The first part discusses several existing theories of anaphora. The theories are evaluated and compared, some problems are pointed out, and possible solutions are suggested, leaving the fundamental ideas of the original theories intact. In the second part, however, these fundamental ideas are reexamined in more detail. Eventually, some of them must be refuted, and a completely different theory is proposed. The most important characteristics of the new proposal are (i) that pronominal anaphora and VP ellipsis are treated in a unified way, and (ii) that the meaning of anaphora is always contextually retrieved. In particular, it is not encoded syntactically, as many current theories assume.

Intended Audience

The issues discussed in this dissertation are central in linguistic theory. They are typically discussed in a very first introductory course in linguistics, so they will be familiar, at least to some extent, to everyone in the field. Therefore, this dissertation should be accessible and of interest not only for anaphora specialists, but also for linguists specialized in other subfields, and for students.

The dissertation is intended to be self-contained, but may be a bit dense for novice students. The ideal background is provided by introductory textbooks that deal with the syntax-semantics interface, such as Heim and Kratzer (1998), and ones that deal more specifically with the logical tools used in semantics, such as Gamut (1991). Other pointers to background reading will be provided along the way.

The formal framework presented in the first chapter may be of special interest to students. This framework is assumed in most contemporary work on semantics and the syntax-semantics interface, but it is hardly ever spelled out in detail. Thus, reading this first chapter will not only help to understand the rest of this dissertation, but also to get a better grasp of the background assumptions made in other contemporary work.

Advanced students and researchers, even those who are not anaphora experts, will probably be sufficiently familiar with the framework presented in chapter 1 to merely glance through it at first and only read parts of it more carefully when needed. Chapter 2, however, will be of particular interest to this audience, as it provides a detailed overview of some of the most prominent existing analyses of pronominal anaphora. These analyses are often closely tied to very general ideas about the relation between linguistic form and meaning, which have played, and continue to play, a major role in linguistic theorizing. Familiarity with these ideas and with the empirical findings that have been adduced as evidence for or against them will be a vital enrichment for anyone in the field.

Anaphora experts may want to proceed directly to chapters 3, 4, and 5, where the really novel ideas are presented. Again, however, it should be emphasized that these chapters are of interest not only for specialists. The arguments presented, though specifically concerned with anaphora, have immediate and significant consequences for the general conception of the relation between linguistic form and meaning.

Acknowledgments

I came to Amsterdam four years ago for two reasons. First, I wanted to learn more about logic, and second, I wanted to learn more about dance. During the first semester here, I took classes every day at the ILLC from 8:00AM till 5:00PM and at the Dance Academy from 5:00PM till 10:00PM. For the many good memories I have from that period, I am very grateful to my sister Eva, with whom I was sharing an apartment at the time, and who literally kept me alive by preparing supersized pasta meals every day when I came home after dancing; my fellow students at the ILLC, especially Yanjing, Edgar, Gustaaf, and Scott; my fellow dance students, especially Orfee; and my teachers, especially Paul Dekker, Dick de Jongh and Michiel van Lambalgen.

During the second semester I worked on my Master thesis with Johan van Benthem. Johan was excellent, and I also benefited very much from interacting with the lively dynamic logic community in- and outside of the ILLC. As part of the project, I visited Luciano Serafini in Trento for one month, and David Pearce in Madrid for two months. I am very grateful to Luciano and David for their hospitality. This was also the period during which I first got to spend some real time together with my girlfriend Ana. We had met the summer before in France, but after that we both had to go our own way, she back to Mexico, and I to Amsterdam. In Madrid, we got together again, and that was fantastic.

That Spring the University of Amsterdam announced that it would fund a PhD position at the Philosophy of Language branch of the ILLC. As the funding would not come from an external funding agency, the applicants were completely free to write their own research proposals. I am very grateful to the University for providing this exceptional opportunity, and especially to Michiel van Lambalgen, who convinced me that doing a PhD at the ILLC would make me very happy, and helped me in putting together a successful project proposal.

When applying, I knew that the project would be supervised by Jeroen Groenendijk, but for some reason, I had never met Jeroen. I had read some of his work of course, which was very impressive, but good researchers are not necessarily good supervisors. What really convinced me that Jeroen had to be an excellent supervisor was the work of his students. I was especially impressed by the dissertations of Maria Aloni, Jelle Gerbrandy, and Balder ten Cate. Not only are these dissertations of exceptional quality, they also exhibit great versatility. Each of them makes a profound contribution to a completely different field. This means, of course, that Maria, Jelle and Balder were very talented students, but also that Jeroen granted them the freedom to develop their own independent thoughts, and successfully guided them in developing those thoughts, even though they were not (in some cases not even remotely) his own. This kind of flexibility is, for me, one of the hallmarks of true wisdom, and certainly one of the most important characteristics of a good thesis supervisor. Jeroen has completely lived up to my expectations in this respect, and I am extremely grateful for that.

During the first year of the project, Barbara Grosz invited me to spend a semester at Harvard. We taught a course on computational linguistics together, and we started working on a paper with Rebecca Nesson. I am very grateful to Barbara for inviting me, to Rebecca for our very pleasant and fruitful collaboration, and to Yonatan, Ping, Koby and the local swing and salsa community for much fun.

The paper that Barbara, Rebecca and I had started working on was about anaphora, but it was written from the perspective of computational linguistics and psycholinguistics. At the time, I didn't know much about what theoretical linguists have to say about the subject. It is remarkable how these communities lead quite separate lifes. I met a psycholinguist the other day who told me that only after thirteen years of doing psycholinguistic research had be encountered the term *c-command*, which is probably the most ubiquitous term in the syntactically oriented theoretical linguistics literature.

So, when I came back from Harvard, I wanted to become acquainted with the theoretical perspective, and this is how I got to know Tanya Reinhart. We met a few times at her place in Amsterdam, which is just around the corner from my office at the Philosophy department, and talked for hours on end. It was a great shock that she suddenly passed away just a few months later. This dissertation has, in great part, been inspired by Tanya and is therefore dedicated to her memory.

Over the last two years, the dissertation has gradually taken its shape. I am very grateful to the audiences of workshops and colloquia in Szklarska Poreba, Paris, Leiden, Nijmegen, Utrecht, Amsterdam, Harvard and MIT for providing me with feedback along the way. Especially, I would like to thank Philippe Schlenker, Emmanuel Chemla, Danny Fox, Irene Heim, Kai von Fintel, Stuart Shieber, Anna Szabolcsi, Bart Geurts, Eric Reuland, Anna Chernilovskaya, Maria Aloni, Salvador Mascarenhas, Paul Dekker, and Edgar Andrade for detailed comments. I am also grateful to my colleagues at the ILLC, in particular my office mates Michael and Tikitu, for providing such a pleasant and inspiring environment to work in.

Finally, I thank Ana, my family and friends for four times five hundred twenty five thousand six hundred minutes. How do you measure the life of a man? In truths that he learned, or in times that he cried? In bridges he burned, or in the way that he died? How about lo–oo–oo–oo–oo–oo–ove?
How about lo–oo–oo–oo–oo–oo–ove?[3]

Amsterdam
July, 2008.

[3] Jonathan Larson, *Rent*, 1996. `www.youtube.com/watch?v=hj7LRuusFqo`

Part I

Bound and Referential Pronouns: Pushing the Limit

Chapter 1

Framework

This part of the dissertation evaluates and refines some of the most prominent theories of pronominal anaphora that have been developed within the framework of Generative Grammar. These theories are particularly concerned with third person singular pronouns such as *he* and *she*, which I will henceforth simply refer to as *pronouns*. The most important common characteristic of the theories to be discussed is that they all assume a fundamental distinction between *bound* and *referential* pronouns. The remainder of this introductory chapter is dedicated to motivating this distinction, and defining the formal syntax and semantics of a fragment of English in which bound and referential pronouns are clearly distinguished. Chapter 2 will discuss several accounts of how binding and coreference are constrained, and chapter 3 will attempt to resolve the issues that are raised and/or left open by these accounts.

1.1 Bound and Referential Pronouns

Pronouns can be interpreted in at least two distinct ways.[1] First, they can be interpreted as *bound variables*. For example, sentence (1.1) below has a reading which says that every man has the property of being an x such that x thinks that x will win. Or in slightly more formal terms, that every man has the property [λx. x thinks that x will win]. On this reading, *he* is interpreted as a variable x which is bound by a λ-operator.

(1.1) Every man thinks he will win.

Second, pronouns can be taken to *refer* to some contextually salient entity. In (1.2) for example, *he* may be taken to refer to John.

(1.2) John is in good shape. I think he will win.

[1] For an early discussion of this distinction, see Partee (1978).

Further motivation for the distinction between bound and referential pronouns comes from the fact that it naturally explains certain ambiguities that arise when pronouns occur in focus constructions and in elliptical constructions.

Ambiguity in focus constructions. Consider the following sentence (capital letters are used here to indicate a pitch accent):

(1.3) MAX called his mother.

Suppose that the pronoun is anaphorically related to *Max*. Then the sentence has two readings. The first says that Max called his mother, and suggests that nobody else called Max's mother. That is, Max has the property [$\lambda x.\ x$ called Max's mother], and nobody else does. The second reading says that Max called his mother and suggests that other people didn't call their mother. That is, Max has the property [$\lambda x.\ x$ called x's mother] and the other people don't. This ambiguity is naturally explained in terms of the distinction between bound and referential pronouns. Interpreting *his* as referring to a contextually salient individual, in this case Max, yields the first reading, while interpreting the pronoun as a bound variable gives us the second reading.[2]

A similar ambiguity arises in constructions which involve focus-sensitive operators such as *only* and *even*. Consider the following example:

(1.4) Only MAX called his mother.

Suppose that the pronoun is anaphorically related to *Max*. Then the sentence has two readings. The first says that only Max has the property [$\lambda x.\ x$ called x's mother] (nobody else called his own mother); the second reading says that only Max has the property [$\lambda x.\ x$ called Max's mother] (nobody else called Max's mother). The distinction between bound and referential pronouns provides a natural explanation of this ambiguity. On the first reading, *his* is interpreted as a bound variable; on the second reading, it is interpreted as referring to Max. Of course, similar examples can be constructed with other focus-sensitive operators.

Ambiguity in elliptical constructions. Consider (1.5), a simple case of VP ellipsis.

(1.5) Max called his mother and Bob did too.

a. ... Bob called his own mother too. [sloppy]

b. ... Bob called Max's mother too. [strict]

[2]Even more readings are obtained, of course, if the pronoun in (1.3) is not taken to refer to Max but to some other contextually salient individual. Such readings are left out of consideration here and in the examples below, because they are not really relevant for the point being made.

Suppose that the pronoun in the source clause (*Max called his mother*) is anaphorically related to *Max*. Then, as first observed by Ross (1967), the target clause (*Bob did too*) has two readings: (1.5a) and (1.5b). The first reading is called *sloppy*; the second is called *strict*.

Keenan (1971) first suggested that this ambiguity can be explained in terms of the distinction between bound and referential pronouns. If the pronoun in the source clause is interpreted as a bound variable, then the source clause as a whole says that Max has the property [$\lambda x.\ x$ called x's mother], and the target clause says that Bill has that property too. This gives us the sloppy reading in (1.5a). If the pronoun is taken to refer to the most salient individual in the utterance context—here, plausibly Max—then the source clause says that Max has the property [$\lambda x.\ x$ called Max's mother], and the target clause, again, says that Bill has that property too. This gives us the strict reading in (1.5b).

This concludes the informal characterization of and motivation for the distinction between bound and referential pronouns. In the next section, I will formally define the syntax and the semantics of a fragment of English in which bound and referential pronouns are clearly distinguished.

1.2 Basic Framework

The fragment to be defined here will include most of the example sentences to be discussed. To keep the framework as simple as possible, the syntax will be allowed to overgenerate considerably. I will not discuss any syntactic constraints that could be deployed to combat this overgeneration. My aim here is merely to set up a precise and convenient terminology, so that the discussion below will be clear and my claims falsifiable.

1.2.1 Syntax

To facilitate the discussion below, I will assume the old Government and Binding architecture (Chomsky, 1981), in which there are four levels of syntactic representation: Deep Structure (DS), Surface Structure (SS), Logical Form (LF), and Phonological Form (PF):

$$\begin{tikzcd}[arrows=dash] & \mathrm{DS} \arrow[d] & \\ & \mathrm{SS} \arrow[dl] \arrow[dr] & \\ \mathrm{LF} & & \mathrm{PF} \end{tikzcd}$$

(PS 1)	S	$\rightarrow$	NP VP	[sentences]
(PS 2)	CP	$\rightarrow$	C S	[complement phrases]
(PS 3)	VP	$\rightarrow$	IV	[verb phrases]
(PS 4)	VP	$\rightarrow$	TV NP	[verb phrases]
(PS 5)	VP	$\rightarrow$	AV CP	[verb phrases]
(PS 6)	NP	$\rightarrow$	POS RN	[noun phrases]
(PS 7)	NP	$\rightarrow$	DET CN	[noun phrases]
(PS 8)	CN	$\rightarrow$	CN S	[common nouns]
(PS 9)	POS	$\rightarrow$	NP 's	[possessives]
(LI 1)	DET	$\rightarrow$	a, the, every, some, no, ...	[determiners]
(LI 2)	CN	$\rightarrow$	man, girl, ...	[common nouns]
(LI 3)	RN	$\rightarrow$	mother, friend, ...	[relational nouns]
(LI 4)	IV	$\rightarrow$	sing, walk, ...	[intransitive verbs]
(LI 5)	TV	$\rightarrow$	call, love ...	[transitive verbs]
(LI 6)	AV	$\rightarrow$	know, say ...	[attitude verbs]
(LI 7)	C	$\rightarrow$	that	[complementizer]
(LI 8)	NP	$\rightarrow$	John, Mary, Max, Lucie, ... who, whom, he$_n$, she$_n$, it$_n$, ... he, she, it, ...	[noun phrases]

Table 1.1: Phrase structure rules and lexical insertion rules for the generation of deep structures.

The DS component of our fragment consists of all trees (or labeled bracketings) that can be generated by the phrase structure rules (PS 1-9) and the lexical insertion rules (LI 1-8) in table 1.1.

Bound pronouns come with a binding index, which is adjoined to the pronoun in subscript (e.g., [she$_2$]). Referential pronouns do not have an index. The general form of possessives is [NP 's]. This generates instances such as [John 's], [every girl 's], and [he$_1$'s]. I will often write [his$_1$] instead of [he$_1$'s], and similarly for other pronominal possessives. Also, I will often simply refer to such pronominal possessives as pronouns, as I already did in the informal discussion above.

I will assume that surface structures are obtained from deep structures by wh-movement, and that logical forms are obtained from surface structures by quantifier raising. If a wh-element moves it receives a binder index n, which is adjoined to it in superscript (e.g., [who]^3). It also leaves behind a trace which has that same index n as its binding index (e.g., the trace of [who]^3 would be t_3).

(1.6) $[_{\text{S}}\ \text{X}\ [_{\text{NP}}\ \text{wh}]\ \text{Y}] \;\Rightarrow\; [_{\text{S}}\ [_{\text{NP}}\ \text{wh}]^n\ [_{\text{S}}\ \text{X}\ \text{t}_n\ \text{Y}]]$ (wh-movement)

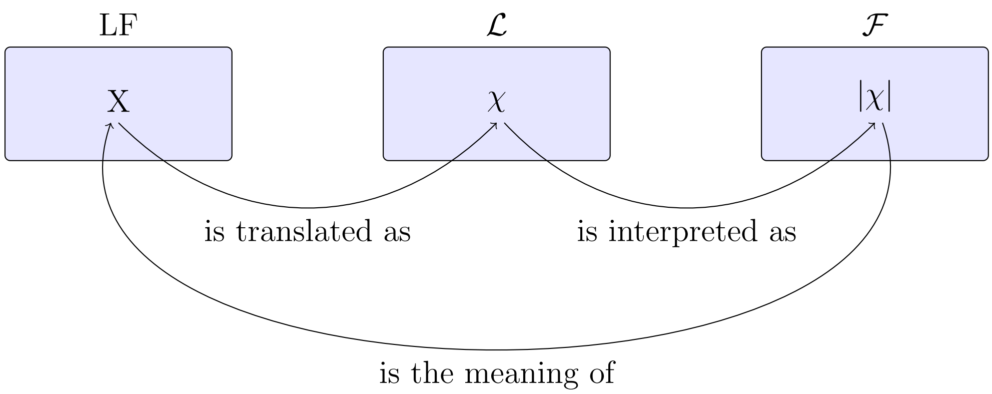
The same goes for quantifier raising: if a noun phrase undergoes QR it receives a binder index n and leaves behind a trace which has that same index n as its binding index.

(1.7) $[_{\text{S}}\ \text{X}\ [_{\text{NP}}\ \text{Z}]\ \text{Y}] \quad \Rightarrow \quad [_{\text{S}}\ [_{\text{NP}}\ \text{Z}]^n\ [_{\text{S}}\ \text{X}\ \text{t}_n\ \text{Y}]]$ (quantifier raising)

Finally, phonological forms are obtained from surface structures by contracting pronominal possessives (e.g. [he_1 's] becomes [his_1]) and deleting all indices, brackets, and traces.

1.2.2 Semantics

The semantic component of our framework associates logical form constituents with their meaning. A standard way of doing so consists in the following three steps. First, a space of possible meanings is defined. Such a space of meanings is called a *frame* $\mathcal{F}$. Second, a formal language $\mathcal{L}$ is defined, and each expression in $\mathcal{L}$ is assigned a meaning in $\mathcal{F}$. Finally, logical form constituents are translated into $\mathcal{L}$-expressions. This is pictured below. Each logical form constituent X is translated into an $\mathcal{L}$-expression χ, which in turn is assigned a meaning $|\chi|$ in $\mathcal{F}$. $|\chi|$ is then called the meaning of X.



I will take $\mathcal{F}$ and $\mathcal{L}$ to be a frame and a language of two-sorted type theory (Ty2) (Gallin, 1975).[3] Below, I will first define Ty2 in general, and then specify the particular Ty2 frame $\mathcal{F}$ and the particular Ty2 language $\mathcal{L}$ that we will use.

Two-sorted Type Theory

We start with the basis: a definition of the *types* in two-sorted type theory. In n-sorted type theory there are $n+1$ basic types and infinitely many complex types. Thus, in the particular case of 2-sorted type theory there are 3 basic types and infinitely many complex types.

[3] In general, $\mathcal{F}$ and $\mathcal{L}$ are taken to be a frame and a language of n-sorted type theory, where n depends on the complexity of the fragment of natural language that is being described.

1.1. Definition. [Types]
The set Ω of Ty2 types is the smallest set of strings such that:

1. $e, s, t \in \Omega$
2. If $\tau, \sigma \in \Omega$, then $(\tau\sigma) \in \Omega$

Outer brackets of complex types will often be omitted. For example, $(s(et))$ will often be abbreviated as $s(et)$.

Given Ω, we can define the class of Ty2 frames and the class of Ty2 languages.

1.2. Definition. [Frames]
A Ty2 frame F is a set of objects $\bigcup_{\tau\in\Omega} D^F_\tau$ such that:

- $D^F_e \neq \emptyset$
- $D^F_s \neq \emptyset$
- $D^F_t = \{0, 1\}$
- $D^F_{\tau\sigma} = \{f \mid f : D_\tau \to D_\sigma\}$ for every complex type $\tau\sigma$

Note that the letter F *ranges over* Ty2 frames here. In particular, it should not be confused with the letter $\mathcal{F}$, which denotes the particular frame whose elements will be associated with the logical form constituents in our fragment of English (this particular frame will be defined below).

For every Ty2 frame F and every Ty2 type τ, D^F_τ is the set of objects of type τ in F. Table 1.2 lists some names that are customarily used for objects of certain types in Ty2 frames.

Objects of type	are called
t	truth values
s	possible worlds
e	individuals
et	properties
$e(et)$	binary relations
se	individual concepts
$s(et)$	property concepts
$s(e(et))$	binary relation concepts
st	propositions

Table 1.2: Names for objects of certain types.

1.3. Definition. [Languages]
A Ty2 language L is a set of expressions $\bigcup_{\tau\in\Omega} E^L_\tau$ such that:

- For every Ty2 type τ, E^L_τ contains a countable set of constants of type τ and a countable set of variables of type τ.

- If φ and ψ are expressions of type t (*formulas*) then $\neg\varphi$ and $(\varphi \wedge \psi)$ are also formulas;
- If φ and ψ are expressions of the same type, then $\varphi = \psi$ is a formula;
- If φ is a formula and x is a variable of any type, then $\forall x.\varphi$ is a formula;
- If φ is a formula and x is a variable of type e, then $\iota x.\varphi$ is an expression of type e.
- If φ is an expression of type σ and x is a variable of type τ, then $\lambda x.\varphi$ is an expression of type $\tau\sigma$;
- If φ is an expression of type $(\tau\sigma)$ and ψ is an expression of type τ, then $\varphi(\psi)$ is an expression of type σ;

Other logical operators ($\exists$, $\vee$, $\rightarrow$, $\leftrightarrow$) are used as abbreviations:

- $\exists x.\varphi$ abbreviates $\neg\forall x.\neg\varphi$
- $\varphi \vee \psi$ abbreviates $\neg(\neg\varphi \wedge \neg\psi)$
- $\varphi \rightarrow \psi$ abbreviates $\neg(\varphi \wedge \neg\psi)$
- $\varphi \leftrightarrow \psi$ abbreviates $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$

Expressions are sometimes subscripted with their type. For example, we may write φ_t to indicate that φ is of type t. Finally, note the difference between the letters L and $\mathcal{L}$. L is used here to range over Ty2 languages, whereas $\mathcal{L}$ denotes the particular Ty2 language whose elements will be associated with the logical form constituents in our fragment of English ($\mathcal{L}$ will be defined below).

Thus, we have defined what Ty2 frames and Ty2 languages are. Now, given a certain Ty2 language L and a certain Ty2 frame F, we must specify how the expressions in L are assigned a meaning in F. This is done by means of *interpretation functions* and an *assignment functions*.

1.4. Definition. [Interpretation functions and assignment functions]
Let L be a Ty2 language and F a Ty2 frame. Then, an *interpretation function* I for L and F is a function that maps every *constant* in L to an object in F, such that for every type τ and every constant $c_\tau \in E^L_\tau$ we have $I(c_\tau) \in D^F_\tau$. That is, I maps every constant of type τ to an object of type τ. Similarly, an *assignment function* g for L and F maps every *variable* of type τ to an object of type τ. If g is an assignment function, we write $g[d/x]$ for the assignment function g' defined by $g'(x) = d$ and $g'(y) = g(y)$ if $y \neq x$.

1.5. Definition. [Interpretation]
Let L be a Ty2 language, F a Ty2 frame, I an interpretation function for L and F, g an assignment function for L and F, and φ an expression in L. Then the

interpretation $|\varphi|^{F,I,g}$ of φ in F given I and g is recursively defined as follows:[4]

$$
\begin{array}{lcl}
|c| & = & I(c) \text{ if } c \text{ is a constant} \\
|x| & = & g(x) \text{ if } x \text{ is a variable} \\
|\neg\varphi| & = & 1 \text{ iff } |\varphi| = 0 \\
|\varphi \wedge \psi| & = & 1 \text{ iff } |\varphi| = 1 \text{ and } |\psi| = 1 \\
|\varphi = \psi| & = & 1 \text{ iff } |\varphi| = |\psi| \\
|\forall x_\tau.\varphi|^{F,I,g} & = & 1 \text{ iff } |\varphi|^{F,I,g[d/x]} = 1 \text{ for all } d \in D_\tau^F \\
|\iota x_e.\varphi|^{F,I,g} & = & \begin{cases} \text{the unique object } d \in D_e^F \text{ such that } |\varphi|^{F,I,g[d/x]} = 1 \\ \text{undefined if such a unique object does not exist} \end{cases} \\
|\lambda x_\tau.\varphi|^{F,I,g} & = & \text{the function } f \text{ with domain } D_\tau^F \text{ such that} \\
 & & \text{for all } d \in D_\tau^F\text{: } f(d) = |\varphi|^{F,I,g[d/x]} \\
|\varphi(\psi)| & = & |\varphi|(|\psi|)
\end{array}
$$

Notice that the interpretation of expressions of the form $\iota x_e.\varphi$ may be undefined. To keep things simple, I have not specified how this may affect the definedness of more complex expressions which contain expressions of this kind as a subexpression. This problem is a particular instance of a more general problem, which is known as the problem of *presupposition projection*. This is an important problem in itself, but I will not go into it here. The reader is referred to (Beaver, 1997; Geurts, 1999) and the references given there.

Next, we define what it means for two Ty2 expressions to be *equivalent*.

1.6. Definition. [Equivalence]
Let L be a Ty2 language and let φ and ψ be expressions in L. Then:

- φ and ψ are equivalent iff $|\varphi|^{F,I,g} = |\psi|^{F,I,g}$ for all F, I, and g.[5]
- φ and ψ are equivalent given a particular frame F' iff $|\varphi|^{F',I,g} = |\psi|^{F',I,g}$ for all I, and g.
- φ and ψ are equivalent given a particular frame F' and a particular interpretation function I' iff $|\varphi|^{F',I',g} = |\psi|^{F',I',g}$ for all g.

We may also define what it means for one Ty2 expression to *entail* another. We are especially interested in entailment between expressions of type st, because these are the expressions that will be associated with *sentential* logical form constituents.

1.7. Definition. [Entailment]
Let L be a Ty2 language and let φ and ψ be expressions of type (st) in L. Then:

[4] Whenever possible, I simply write $|\varphi|$ instead of $|\varphi|^{F,I,g}$.

[5] Provided, of course, that I is an interpretation function for L and F, and g is an assignment function for L and F. Henceforth, this qualification will be left implicit.

- φ entails ψ iff
  for all F, I, and g, and for all $w \in D_s^F$ such that
  $|\varphi|^{F,I,g}(w) = 1$ we also have $|\psi|^{F,I,g}(w) = 1$

- φ entails ψ given a particular frame F' iff
  for all I and g, and for all $w \in D_s^{F'}$ such that
  $|\varphi|^{F',I,g}(w) = 1$ we also have $|\psi|^{F',I,g}(w) = 1$

- φ entails ψ given a particular frame F'
  and a particular interpretation function I' iff
  for all g and for all $w \in D_s^{F'}$ such that
  $|\varphi|^{F',I',g}(w) = 1$ we also have $|\psi|^{F',I',g}(w) = 1$

Finally, it should be remarked that Ty2 expressions can be converted into other Ty2 expressions by α-conversion and β-reduction. α-conversion can be thought of as re-naming of bound variables and β-reduction as applying λ-expressions to their arguments. For example:

$$\begin{array}{lll} (\lambda x.x = y) & \text{can be } \alpha\text{-converted into} & (\lambda z.z = y) \\ (\lambda x.x = y)(z) & \text{can be } \beta\text{-reduced to} & (z = y) \end{array}$$

If ψ can be obtained from φ by (repeatedly) applying α-conversion and/or β-reduction, then we will simply say that φ can be *reduced* to ψ. If φ can be reduced to ψ, then φ and ψ are always equivalent (for a proof of this fact, as well as proper definitions of α-conversion and β-reduction see Andrews, 1986). This means that the picture we started out with in the beginning of this section is in fact a little bit more complicated. Each logical form constituent X is translated into an $\mathcal{L}$-expression χ. This expression may be *reducible* to other $\mathcal{L}$-expressions $\chi', \chi'', \ldots$. In any case, $\chi, \chi', \chi'', \ldots$ will be equivalent, that is, they will all be associated with the same meaning $|\chi|$. Thus, $|\chi|$ will be called *the* meaning of X, and $\chi, \chi', \chi'', \ldots$ will all be called possible translations of X.

This concludes my presentation of Ty2. For more detail, I refer to Gallin (1975) and Andrews (1986).

Fixing $\mathcal{F}$, $\mathcal{L}$, and $\mathcal{I}$

Let me now specify $\mathcal{F}$ and $\mathcal{L}$, the particular Ty2 frame and Ty2 language whose elements will be associated with the LF constituents in our fragment of English. We will take $\mathcal{F}$ to be the most general frame, containing all possible meanings. This means, in particular, that $D_e^{\mathcal{F}}$ will consist of all possible individuals and that $D_s^{\mathcal{F}}$ will consist of all possible worlds. Next let us define $\mathcal{L}$. To do so we must fix its inventory of constants and its inventory of variables. The constants in $\mathcal{L}$ correspond to the content words in our fragment of English.[6] Some of the

[6]There is a traditional distinction between content words and function words. Names, nouns, verbs, adjectives, and most adverbs are considered to be content words, while determiners,

constants in $\mathcal{L}$ are listed in table 1.3, and some of the variables in $\mathcal{L}$ are listed in table 1.4. Notice that $\mathcal{L}$ contains two kinds of variables ranging over individuals: $x_1, x_2, \ldots$ will be used in the translation of traces, whereas $x, x', \ldots$ will be used for all other purposes (see, for example, the translation of *every* in table 1.5 below).

JOHN	se	individual concept
SING	$s(et)$	property concept
MAN	$s(et)$	property concept
MOTHER	$s(e(et))$	binary relation concept
LOVE	$s(e(et))$	binary relation concept
SAY	$s((st)(et))$	

Table 1.3: Some constants in $\mathcal{L}$.

$w, w', \ldots$	s	worlds
$x, x', \ldots$	e	individuals
$x_1, x_2, \ldots$	e	individuals
$P, P', \ldots$	et	properties
$R, R', \ldots$	$e(et)$	binary relations
$p, p', \ldots$	st	propositions

Table 1.4: Some variables in $\mathcal{L}$.

Apart from $\mathcal{F}$ and $\mathcal{L}$, we will also fix the interpretation function $\mathcal{I}$ that maps all the constants in $\mathcal{L}$ onto appropriate meanings in $\mathcal{F}$. $\mathcal{I}$ is taken to be the most general interpretation function that respects the way in which different content words are conventionally related. For example, $\mathcal{I}$ must be such that for every world w in $D_s^{\mathcal{F}}$, $\mathcal{I}(\text{TIGER})(w)$ (the set of tigers in w) is a subset of $\mathcal{I}(\text{ANIMAL})(w)$ (the set of animals in w).

From Logical Form Constituents to Ty2 Expressions

Now we are ready to specify how logical form constituents are translated into $\mathcal{L}$-expressions. This is done in two steps. First, the translation of terminal nodes is defined and second, the translation of non-terminal nodes is defined in terms of the translations of their daughter nodes. The translation function $[\![\ ]\!]^{\mathrm{C}}$ will have a context-parameter C, which reflects the idea that the interpretation of some words, in particular referential pronouns, depends on the context of use.

pronouns, complementizers, auxiliaries, expletives, etc. are considered to be function words. The only content words in our fragment are names, nouns, and verbs.

Node		Translation	Type
$[\![\text{man}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda x.\textsc{man}(w)(x)$	$s(et)$
$[\![\text{sing}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda x.\textsc{sing}(w)(x)$	$s(et)$
$[\![\text{mother}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda x.\lambda y.\textsc{mother}(w)(x)(y)$	$s(e(et))$
$[\![\text{love}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda x.\lambda y.\textsc{love}(w)(x)(y)$	$s(e(et))$
$[\![\text{say}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda p.\lambda x.\textsc{say}(w)(p)(x)$	$s((st)(et))$
$[\![\text{John}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\textsc{john}(w)$	se
$[\![\text{he}]\!]^{\mathrm{C}}$	$=$	$[\![\textsc{ant}^{\mathrm{C}}(\text{he})]\!]^{\mathrm{C}}$	se
$[\![\text{he}_n]\!]^{\mathrm{C}}$	$=$	$\lambda w.x_n$	se
$[\![\text{t}_n]\!]^{\mathrm{C}}$	$=$	$\lambda w.x_n$	se
$[\![\text{that}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda p.p(w)$	$s((st)t)$
$[\![\text{who}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda P.P$	$s((et)(et))$
$[\![\text{every}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda P.\lambda P'.\forall x(P(x) \rightarrow P'(x))$	$s((et)((et)t))$
$[\![\text{the}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda P.\iota x.P(x)$	$s((et)e)$
$[\![\text{'s}]\!]^{\mathrm{C}}$	$=$	$\lambda w.\lambda x.\lambda R.\iota x'.R(x)(x')$	$s(e((e(et))e))$

Table 1.5: Translations of some terminal LF nodes.

The Translation of Terminal Nodes. Table 1.5 specifies the translation of some terminal LF nodes. Other terminal LF nodes are translated analogously.

A referential pronoun [he] occurring in a context C is translated as $[\![\textsc{ant}^{\mathrm{C}}(\text{he})]\!]^{\mathrm{C}}$ where $\textsc{ant}^{\mathrm{C}}(\text{he})$ is the *antecedent* of [he] in C. The antecedent of a referential pronoun must always be a referential expression itself: an expression of type (se) whose translation does not contain any free variables (traces, in particular, do *not* count as referential expressions). If a referential expression A is the antecedent of a pronoun P in a context C, then I will say that P is *resolved* to A in C and write P = A next to the LF under consideration. For example, if (1.8) is considered in a context in which [she] is resolved to [Mary] I will write she = Mary next to it, as in (1.9).

(1.8) [Mary] [says that she likes John]

(1.9) [Mary] [says that she likes John] she = Mary

Notice that apart from the translation of referential pronouns, all the other translations in table 1.5 are context-independent. This is a simplification, which I permit myself here in order to focus exclusively on the interpretation of pronouns. In general, the translation of other nodes may also be context-dependent (I am thinking, for example, of the domain restrictions of determiners).

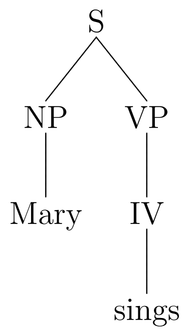
Finally, notice that every terminal LF node X is translated into a type-theoretical expression χ of type $s\tau$ (where τ may be different in each case). This means that X is always associated with a function $|\chi|$ from possible worlds to objects of type τ. Such objects (functions from possible worlds to other objects) are called *intensional objects*. Accordingly, $|\chi|$ is called the *intension* of X. For every particular world w, $|\chi|(w)$ will be an object of type τ. This object is called the *extension* or the *denotation* of X in w. We will say that X *expresses* its intension $|\chi|$ and that X *denotes* its extension $|\chi|(w)$ in each particular world w.

For example, [John] expresses the individual concept $|\lambda w.\textsc{john}(w)|$ and in each particular world w', it denotes the individual $|\lambda w.\textsc{john}(w)|(w')$. An intransitive verb like [sing] expresses the property concept $|\lambda w.\lambda x.\textsc{sing}(w)(x)|$, and in each particular world w' it denotes the property $|\lambda w.\lambda x.\textsc{sing}(w)(x)|(w')$. This terminology extends in a natural way to the other terminal nodes and, as we will see right below, also to all non-terminal nodes.

The Translation of Non-Terminal Nodes. The composition rules in table 1.6 specify how the translation of a non-terminal LF node can be constructed from the translations of its daughter nodes. Notice that the composition rules assign to every non-terminal node X a translation χ of type $s\tau$. Thus, the meaning associated with a non-terminal node is always an *intensional* object. As in the case of terminal nodes, $|\chi|$ is called the intension of X, and in every particular world w, $|\chi|(w)$ is called the denotation or the extension of X in w.

Let me go through a few examples to illustrate how the composition rules work. First, consider the logical form in (1.10).

(1.10) $[_{S}[_{NP}$ Mary$]$ $[_{VP}[_{IV}$ sings$]]]$



This example illustrates the workings of COPY and EFA (extensional function application). First, COPY tells us that the translation of $[_{VP}[_{IV}$ sings$]]$ is identical to the translation of $[_{IV}$ sings$]$, which is defined in the lexicon:

(1.11) $\lambda w.\lambda x.\textsc{sing}(w)(x)$

The translation of $[_{NP}$ Mary$]$ is also defined in the lexicon:

(1.12) $\lambda w.\textsc{mary}(w)$

Now EFA tells us that the translation of $[_{S}$ Mary sings$]$ is:

COPY
If a non-terminal node X only has one daughter node Y then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = [\![\mathrm{Y}]\!]^{\mathrm{C}}$$

EFA (extensional function application)
If a non-terminal node X has two daughters Y and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ of type $s(\tau\sigma)$ and ζ of type $s\tau$ for some τ and σ, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.\gamma(w)(\zeta(w))$$

IFA (intensional function application)
If a non-terminal node X has two daughters Y and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ of type $s((s\tau)\sigma)$ and ζ of type $s\tau$ for some τ and σ, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.\gamma(w)(\zeta)$$

QINP (quantifying in noun phrases of type se)
If a non-terminal node X has two daughters Y^{n} (notice the binder index) and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ of type se and ζ of type st, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.(\lambda x_n.\zeta(w))(\gamma(w))$$

QIGQ (quantifying in generalized quantifiers of type $s((et)t)$)
If a non-terminal node X has two daughters Y^{n} (notice the binder index) and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ of type $s((et)t)$ and ζ of type st, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.(\gamma(w))(\lambda x_n.\zeta(w))$$

QIWH (quantifying in wh-elements of type $s((et)(et)))$
If a non-terminal node X has two daughters Y^{n} (notice the binder index) and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ of type $s((et)(et))$ and ζ of type st, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.(\gamma(w))(\lambda x_n.\zeta(w))$$

PM (predicate modification)
If a non-terminal node X has two daughters Y and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ and ζ both of type $s(et)$, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.\lambda x.\gamma(w)(x) \wedge \zeta(w)(x)$$

FC (function composition)
If a non-terminal node X has two daughters Y and Z such that $[\![\mathrm{Y}]\!]^{\mathrm{C}} = \gamma$ and $[\![\mathrm{Z}]\!]^{\mathrm{C}} = \zeta$ with γ of type $s(\tau\sigma)$ and ζ of type $s(\sigma\rho)$ for some τ, σ and ρ, then:

$$[\![\mathrm{X}]\!]^{\mathrm{C}} = \lambda w.\lambda y_\tau.(\zeta(w))(\gamma(w)(y))$$

Table 1.6: Rules which determine the translation of non-terminal LF nodes.

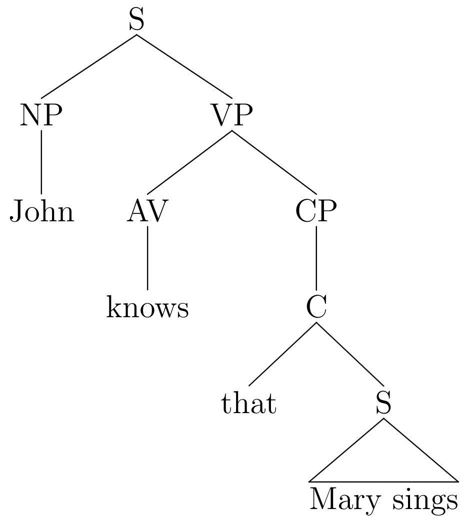
(1.13) $$\lambda w.(\overset{\text{[sings]}}{\boxed{\lambda w'.\lambda x.\textsc{sing}(w')(x)}})(w)\left((\overset{\text{[Mary]}}{\boxed{\lambda w''.\textsc{mary}(w'')}})(w)\right)$$

which can be reduced to:

(1.14) $$\lambda w.(\textsc{sing}(w)(\textsc{mary}(w)))$$

The next example, (1.15), illustrates how IFA (intensional function application) works. It also shows how the complementizer *that* and CP-embedding verbs like *know* are treated.

(1.15) $[_{S}[_{NP}$ John$]$ $[_{VP}[_{AV}$ knows$]$ $[_{CP}[_{C}$ that$]$ $[_{S}$ Mary sings$]]]]$



Let us first determine the translation of the embedded CP. Notice that the translation of $[_{S}$ Mary sings$]$ was derived above. The translation of $[_{C}$ that$]$ can be found in the lexicon:

(1.16) $$\lambda w.\lambda p.p(w)$$

Now IFA tells us how to combine the translations of $[_{C}$ that$]$ and $[_{S}$ Mary sings$]$ to get the translation of $[_{CP}$ that Mary sings$]$:

(1.17) $$\lambda w.(\overset{\text{[that]}}{\boxed{\lambda w'.\lambda p.p(w')}})(w)(\overset{\text{[Mary sings]}}{\boxed{\lambda w''.(\textsc{sing}(w'')(\textsc{mary}(w'')))}})$$

which can be reduced to:

(1.18) $$\lambda w.(\textsc{sing}(w)(\textsc{mary}(w)))$$

Notice that this is identical to the translation of $[_{S}$ Mary sings$]$. So the complementizer $[_{C}$ that$]$ has no semantic effect. Now let us determine the translation of the matrix clause. The translation of $[_{PV}$ knows$]$ can be found in the lexicon:

(1.19) $\lambda w.\lambda p.\lambda x.\text{KNOWS}(w)(p)(x)$

IFA tells how to combine this with the translation of the embedded clause to get the translation of [$_{\text{VP}}$ knows that Mary sings]:

(1.20) $\lambda w.(\overset{\text{[knows]}}{\boxed{\lambda w'.\lambda p.\lambda x.\text{KNOWS}(w')(p)(x)}})(w)(\overset{\text{[that Mary sings]}}{\boxed{\lambda w''.(\text{SING}(w'')(\text{MARY}(w'')))}})$

which can be reduced to:

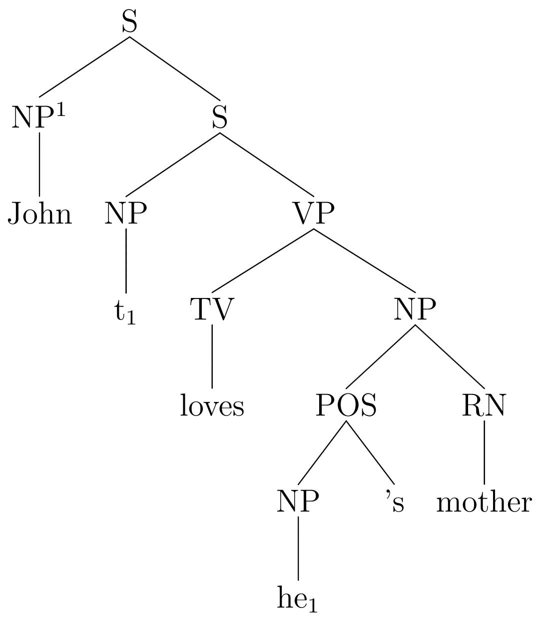
(1.21) $\lambda w.\lambda x.\text{KNOWS}(w)(\lambda w'.\text{SING}(w')(\text{MARY}(w')))(x)$

Finally, this is combined with the translation of [$_{\text{NP}}$ John], which can be found in the lexicon, to get the translation of (1.15):

(1.22) $\lambda w.\text{KNOWS}(w)(\lambda w'.\text{SING}(w')(\text{MARY}(w')))(\text{JOHN}(w))$

The next example, (1.23), illustrates how a noun phrase of type *se* is "quantified in" with the help of QINP. It also shows how possessives like [$_{\text{POS}}$ he$_1$'s] and relational nouns like [$_{\text{RN}}$ mother] are treated.

(1.23) [$_{\text{S}}$ [$_{\text{NP}}$ John]^1 [$_{\text{S}}$ [$_{\text{NP}}$ t$_1$] [$_{\text{VP}}$ [$_{\text{TV}}$ loves] [$_{\text{NP}}$ [$_{\text{POS}}$ he$_1$'s]] [$_{\text{RN}}$ mother]]]]]



Let us first derive the translation of [$_{\text{POS}}$ he$_1$'s]. The translation of its elements can be found in the lexicon and are composed using EFA to get:

(1.24) $\lambda w.\lambda R.\iota x.R(x_1)(x)$

This can be composed with the translation of [$_{\text{RN}}$ mother], again using EFA, to obtain the translation of [$_{\text{NP}}$ his$_1$ mother]:

(1.25) $\lambda w.\iota x.\textsc{mother}(w)(x_1)(x)$

Two more applications of EFA give us the translation of [$_S$ t$_1$ loves his$_1$ mother]:

(1.26) $\lambda w.\textsc{love}(w)(\iota x.\textsc{mother}(w)(x_1)(x))(x_1)$

Finally, QINP tells us how to compose this with the translation of [$_{NP}$ John]^1 to get the translation of (1.23):

$$\lambda w.(\lambda x_1.(\overset{[\mathrm{t}_1\text{ loves his}_1\text{ mother}]}{\boxed{\lambda w'.\textsc{love}(w')(\iota x.\textsc{mother}(w')(x_1)(x))(x_1)}})(w))\left((\overset{[\mathrm{John}]}{\boxed{\lambda w''.\textsc{john}(w'')}})(w)\right)$$

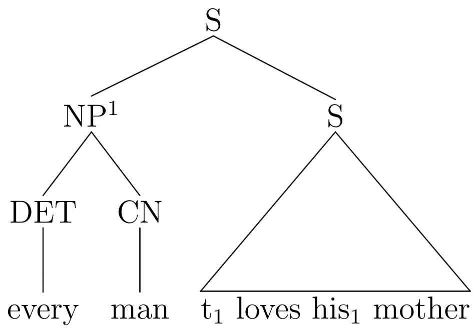
which can be reduced to:

(1.27) $\lambda w.\textsc{love}(w)(\iota x.\textsc{mother}(w)(\textsc{john}(w))(x))(\textsc{john}(w))$

If the bound pronoun [he$_1$] in (1.23) were replaced by a referential pronoun [he] with [John] as its antecedent, we would get exactly the same end result.

The example in (1.28) is very much like the one in (1.23). Only, instead of showing how noun phrases of type *se* are quantified in, it shows how *generalized quantifiers* of type $s((et)t)$ are quantified in, and also how determiners like [$_{DET}$ every] work.

(1.28) [$_S$ [$_{NP}$ [$_{DET}$ every] [$_{CN}$ man]]^1 [$_S$ t$_1$ loves his$_1$ mother]]



The translation of [$_S$ t$_1$ loves his$_1$ mother] was given in (1.26). The translation of [$_{DET}$ every] can be found in the lexicon:

(1.29) $\lambda w.\lambda P.\lambda P'.\forall x(P(x) \rightarrow P'(x))$

This can be composed with the translation of [$_{CN}$ man] using EFA to get the translation of [$_{NP}$ every man]:

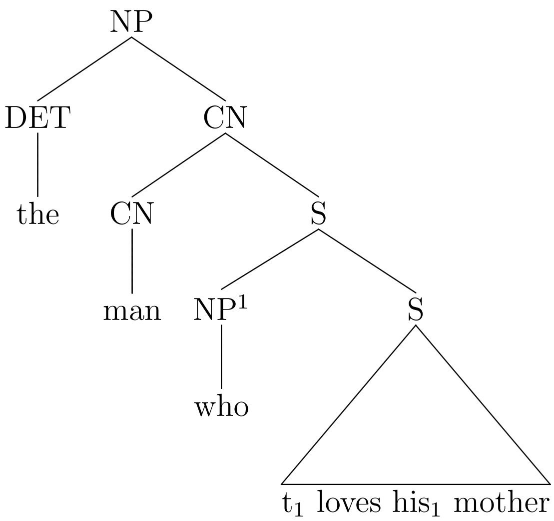
(1.30) $\lambda w.\lambda P'.\forall x(\textsc{man}(w)(x) \rightarrow P'(x))$

Now QIGQ tells us how to combine the translation of [$_{NP}$ every man]^1 with that of [$_S$ t$_1$ loves his$_1$ mother] to get the translation of (1.28):

(1.31) $\lambda w.\forall x(\textsc{man}(w)(x) \rightarrow \textsc{love}(w)(\iota x'.\textsc{mother}(w)(x)(x'))(x))$

We will do two more examples. One to illustrate how QIWH and PM deal with relative clauses, and one to show how FC deals with quantifiers in object position. First consider (1.32).

(1.32) $[_{\text{NP}}\ [_{\text{DET}}\ \text{the}]\ [_{\text{CN}}\ [_{\text{CN}}\ \text{man}]\ [_{\text{S}}\ [_{\text{NP}}\ \text{who}]^1\ [_{\text{S}}\ \text{t}_1\ \text{loves his}_1\ \text{mother}]]]]$



The translation of $[_{\text{S}}\ \text{t}_1$ loves his_1 mother] was given in (1.26). The translation of $[_{\text{NP}}$ who] can be found in the lexicon:

(1.33) $\lambda w.\lambda P.P$

Now, QIWH tells us how to compose the translation of $[_{\text{NP}}$ who] with the translation of $[_{\text{S}}\ \text{t}_1$ loves his_1 mother] to get the translation of the relative clause:

$$\lambda w.((\overset{\text{[who]}}{\boxed{\lambda w'.\lambda P.P}})(w))(\lambda x_1.(\overset{[\text{t}_1\ \text{loves his}_1\ \text{mother}]}{\boxed{\lambda w''.\textsc{love}(w'')(\iota x.\textsc{mother}(w'')(x_1)(x))(x_1)}})(w))$$

which reduces to:

(1.34) $\lambda w.\lambda x_1.\textsc{love}(w)(\iota x.\textsc{mother}(w)(x_1)(x))(x_1)$

The next step is to derive the translation of $[_{\text{CN}}$ man $\text{who}^1\ \text{t}_1$ loves his_1 mother]. The translation of $[_{\text{CN}}$ man] is given in the lexicon:

(1.35) $\lambda w.\lambda x.\textsc{man}(w)(x)$

and PM (predicate modification) tells us how to compose this with the translation of the relative clause to get:

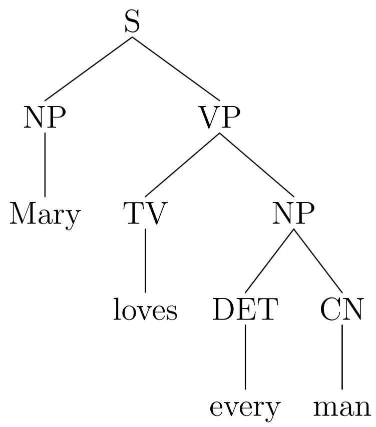
(1.36) $\lambda w.\lambda x.\textsc{man}(w)(x) \wedge \textsc{love}(w)(\iota x'.\textsc{mother}(w)(x)(x'))(x)$

Finally, EFA tells us how to compose this with the translation of $[_{\text{DET}}$ the$]$ to get the translation of (1.32):

(1.37) $\lambda w.\iota x.\textsc{man}(w)(x) \wedge \textsc{love}(w)(\iota x'.\textsc{mother}(w)(x)(x'))(x)$

The last example, (1.38), shows how FC deals with quantifiers in object position.

(1.38) $[_{\text{S}}\ [_{\text{NP}}$ Mary$]$ $[_{\text{VP}}\ [_{\text{TV}}$ loves$]$ $[_{\text{NP}}\ [_{\text{DET}}$ every$]$ $[_{\text{CN}}$ man$]]]]$



Such constructions cannot be dealt with by standard function application, because quantifiers are of type $s((et)t)$ while transitive verbs are of type $s(e(et))$. Thus, transitive verbs combine, in every world, with something of type e to yield something of type et. Quantifiers don't provide something of type e but something of type $(et)t$ in every world, so function application is impossible.

But notice that the *input type* of generalized quantifiers, et, matches the *output type* of transitive verbs. If the transitive verb could just get its input elsewhere, then the generalized quantifier would know what to do with its output. This is the idea of function composition: a function f of type $s((et)t)$ and a function f' of type $s(e(et))$ are composed into a function $\lambda w.\lambda x.f(w)(f'(w)(x))$ of type $s(et)$ which, in every world w, takes an individual of type e as its input and gives as its output the result of first applying $f'(w)$ to x and then applying $f(w)$ to $f'(w)(x)$. In our concrete example, the ingredients of function composition are the translation of $[_{\text{TV}}$ loves$]$ and the translation of $[_{\text{NP}}$ every man$]$:

(1.39) $\lambda w.\lambda x.\lambda y.\textsc{loves}(w)(x)(y)$

(1.40) $\lambda w.\lambda P.\forall x(\textsc{man}(w)(x) \rightarrow P(x))$

FC tells us how to compose these two functions in order to get the translation of $[_{\text{VP}}$ loves every man$]$:

(1.41) $\lambda w.\lambda x'.\forall x(\textsc{man}(w)(x) \rightarrow \textsc{loves}(w)(x)(x'))$

And EFA tells us how to combine (1.41) with the translation of $[_{\text{NP}}$ Mary$]$ to get the translation of (1.38):

(1.42) $\lambda w.\forall x(\textsc{man}(w)(x) \rightarrow \textsc{loves}(w)(x)(\textsc{mary}(w)))$

This concludes the illustration of the translation of non-terminal LF nodes. Let me remark that there are many alternative ways to set up the lexicon and the composition rules. For example, if we assume more complex types in the lexicon, add *type raising* to our inventory of composition rules, and/or make quantifier raising obligatory, we could possibly do without function composition and unify the rules for quantifying in (see Heim and Kratzer, 1998, chapter 7, for some discussion). However, such adaptations would, as far as I can see, not have any significant consequences for the particular issues that are to be discussed in this dissertation. For practical convenience I have chosen here to use simple types in the lexicon and a relatively large inventory of composition rules.

We have now completely filled in the picture we started out with in the beginning of this section. First, we specified $\mathcal{F}$, $\mathcal{L}$, and the interpretation function $\mathcal{I}$ which associates expressions in $\mathcal{L}$ to meanings in $\mathcal{F}$. Then we specified how logical form constituents are translated into $\mathcal{L}$-expressions. Putting everything together, we end up with a system that assigns a meaning to every logical form constituent in our fragment.

1.3 Contextual and Conventional Meaning

Let me take a step back at this point and observe that the framework laid out above allows us to make a distinction between two kinds of meaning: contextual meaning and conventional meaning. To appreciate this distinction, notice that several factors are involved in the interpretation of linguistic expressions. First of all, to interpret (a particular usage of) an expression it is necessary to assume that that expression belongs to the vocabulary of a particular language (e.g. some dialect of English) and that it is to be interpreted according to the linguistic *conventions* to which speakers of that language adhere. In the case of English, such conventions determine, for example, how words like *chair* and *sing* are interpreted.

In addition, the interpretation of an expression often depends on the *context* in which it is used (e.g. what has been said before, what is the topic of the conversation, what is the question that is being addressed, etcetera). This is especially clear in the case of referential pronouns—their interpretation is not fixed by general conventions, but depends on the context of use.

This distinction is captured by the formal machinery developed above. We may define the following two notions of meaning:

1.8. Definition. [Contextual Meaning]
The contextual meaning of a logical form constituent X in a context C is $[\![\mathrm{X}]\!]^{\mathrm{C}}$.

1.9. Definition. [Conventional Meaning]
The conventional meaning of a logical form constituent X is the function which maps every context C to the contextual meaning of X in C.

Similarly, we can define the following notions of equivalence and entailment:

1.10. Definition. [Equivalence]
Let X and Y be two logical form constituents and let C_X and C_Y be the respective contexts in which they are used. Then:

- X and Y are contextually equivalent relative to C_X and C_Y iff $[\![X]\!]^{C_X}$ and $[\![Y]\!]^{C_Y}$ are equivalent given $\mathcal{F}$ and $\mathcal{I}$;
- X and Y are conventionally equivalent iff there are two contexts C_X and C_Y such that X and Y are contextually equivalent relative to C_X and C_Y.

1.11. Definition. [Entailment]
Let X and Y be two sentential logical forms and let C_X and C_Y be the respective contexts in which they are used. Then:

- X contextually entails Y relative to C_X and C_Y iff $[\![X]\!]^{C_X}$ entails $[\![Y]\!]^{C_Y}$ given $\mathcal{F}$ and $\mathcal{I}$;
- X conventionally entails Y iff there are two contexts C_X and C_Y such that X contextually entails Y relative to C_X and C_Y.

Intuitively, X is conventionally equivalent with Y iff they are equivalent *as far as their conventional meaning is concerned*. A similar intuition holds for conventional entailment. These fine-grained notions of meaning, equivalence, and entailment will play a significant role below, especially in section 1.8.

We now turn to the formal definition of anaphoric relations such as binding and coreference.

1.4 Anaphoric Relations

The grammatical framework laid out above allows us to formally define notions such as binding and coreference. In doing so I will try to stay as close as possible to the notions that have been discussed in the literature (be it formally or informally). Let me start with binding. The definition of binding requires the definition of one auxiliary notion, namely that of c-command.

1.12. Definition. [C-command]
One node A c-commands another node B iff (i) A does not dominate B and (ii) all branching nodes that dominate A also dominate B.

1.13. Definition. [Binding]
Let X be a logical form constituent, A a noun phrase in X with a binder index, and B a pronoun or trace in X with a binding index. Then A binds B in X iff:

i A's binder index matches B's binding index;

ii A c-commands B in X;

iii A does not c-command any other NP in X which satisfies i and ii.

This notion of binding is what Heim and Kratzer (1998) and Büring (2005a) call *semantic binding* and what Reinhart (2006) calls *A-binding*. To get a feel for what the notion amounts to consider the following examples:

(1.43) [John]^1 [t_1 loves his$_1$ mother]

(1.44) [every man]^1 [t_1 thinks that he$_1$ will win]

In (1.43), [John] binds [t_1] and [his$_1$]; in (1.44), [every man] binds [t_1] and [he$_1$].
In terms of binding we may define the following notion of *cobinding*.

1.14. Definition. [Cobinding]
Two nodes A and B in a logical form constituent X are cobound iff there is a third node which binds both A and B in X.

In (1.43), [t_1] and [his$_1$] are cobound, and in (1.44), [t_1] and [he$_1$] are cobound.

Finally, consider coreference. This relation only involves *referential* noun phrases: expressions of type *se* whose translation does not contain any free variables. The notion of coreference that is generally assumed in the literature does not require that two expressions denote the same individual in *all* possible worlds, but merely that they denote the same individual in those worlds that are consistent with the speech participants' common assumptions in a given utterance context.[7] Stalnaker (1978) called this set of possible worlds the *context set*.[8] To appreciate the idea that coreference only requires denoting the same individual in each world in the context set, consider the name *Zapatero* and the description *the President of Spain*. In a conversation between two people from Madrid, the context set will probably only include worlds in which the name and the description denote exactly the same individual. As a consequence, whenever one of the

[7]This notion of coreference is sometimes called *presupposed* coreference (cf. Büring, 2005a, p.153).

[8]The term *common ground* is often used synonymously with the term *context set*. However, as Kai von Fintel pointed out to me, Stalnaker used these terms for distinct notions. The common ground, in his terminology, is a set of presupposed *propositions*, whereas the context set is a set of *possible worlds* recognized by the speaker to be the "live options" relevant to the conversation (Stalnaker, 1978, p.84–85).

speech participants uses the name, he may just as well have used the description to convey the same message. Thus, intuitively, the name and the description corefer in such a context.

In a conversation between two people from Melbourne, the context set will probably include worlds in which *Zapatero* and *the President of Spain* do not denote the same individual. Even if both speech participants know that Zapatero is the President of Spain, they may not take for granted that their interlocutor knows this as well. Therefore, if one of them uses the description, he cannot be sure that using the name instead would convey the same message. Thus, intuitively, the name and the description do not corefer in such a context. This idea can be formalized as follows:

1.15. Definition. [Coreference]
Let C be a context, and let S_C be the context set in C. Then, two referential noun phrases A and B corefer in C iff for every $w \in S_C$, $[\![\mathrm{A}]\!]^{\mathrm{C}}(w)$ is equivalent to $[\![\mathrm{A}]\!]^{\mathrm{C}}(w)$ given $\mathcal{F}$ and $\mathcal{I}$.

This concludes the definition of anaphoric relations. Let us now return to the examples discussed at the very beginning of this chapter, repeated here:

(1.3) MAX called his mother.

(1.4) Only MAX called his mother.

(1.5) Max called his mother and Bob did too.

It was suggested that the distinction between bound and referential pronouns would yield a natural explanation of the ambiguities exhibited by these examples. We are almost ready to spell out this explanation in detail. The final ingredient we need is a basic theory of focus.

1.5 Focus

Theories of focus (cf. Rooth, 1985) generally assume that constituents may or may not be F-marked at surface structure. For example, the surface structure in (1.45) has an F-marked subject NP.

(1.45) [the dog]$_{\mathrm{F}}$ [destroyed the vase]

F-marking is interpreted both phonologically (at PF) and semantically (at LF). The phonological interpretation of F-features consists in *accenting* certain syllables within each F-marked constituent. For example, the surface structure in (1.45) will be pronounced as:

(1.46) the DOG destroyed the vase

Which of the syllables in an F-marked constituent are accented depends on various factors, which are not directly relevant here (cf. Büring, 2007).

The semantic import of F-features is their role in determining the *focus alternatives* of LF constituents. The focus alternatives of an LF constituent X are obtained from X by replacing its F-marked sub-constituents with contextually salient alternatives. For example, some of the focus alternatives of (1.45) may be:

(1.47) a. [the cat] [destroyed the vase]
b. [the burglar] [destroyed the vase]
c. [a friend of my father] [destroyed the vase]

Let us write $\textsc{alt}^{C}(X)$ for the set of focus alternatives of X in C, and let us call:

$$[\![X]\!]^{C,F} = \{[\![Y]\!]^{C} \mid Y \in \textsc{alt}^{C}(X)\}$$

the *focus value* of X in C (to avoid confusion, $[\![X]\!]^{C}$ and $[\![X]\!]^{C,F}$ are sometimes called the *ordinary* semantic value and the *focus* semantic value of X in C, respectively).[9] Focus alternatives play a role in a variety of linguistic phenomena. Notorious examples are the computation of *implicatures*, which figure in examples like (1.3) and will be discussed in section 1.6, the interpretation of focus-sensitive operators, which play a role in examples like (1.4) and will be discussed in section 1.7, and the interpretation of VP ellipsis, which is relevant for examples like (1.5) and will be discussed in section 1.8.

1.6 Implicatures

Consider the following scenario, taken from Rooth (1992b). Mats, Steve, and Paul are taking an exam, which is graded right away. When Mats comes home, his brother George asks how it went. Mats answers:

(1.48) Well, I PASSED.

Given this answer, George will probably conclude that Mats did not do better than passing, that he did not, for example, ace the exam.

Now consider another answer Mats could have given:

(1.49) Well, STEVE passed.

Given this answer, George would probably conclude that Mats and Paul did not pass. The general line of reasoning that leads to this conclusion could be the

[9]For simplicity, I assume here that focus alternatives are determined at a *syntactic* level. Rooth (1985) assumed that they are determined at a *semantic* level. For the particular phenomena to be discussed here, it does not really matter which of these assumptions is adopted. I have adopted the first just to keep things as simple as possible.

following: Mats said that Steve passed. If he or Paul had passed as well, he would have said so. He didn't, so he and Paul probably didn't pass.

In the case of (1.48), George's reasoning is similar: Mats said that he passed. If he had aced he would have said so. He didn't, so he probably didn't ace.

Grice (1975) called the conclusions that arise from such reasoning patterns *implicatures*. The role of focus in the computation of implicatures is to determine the appropriate set of comparison. A given logical form LF is always compared with its focus alternatives. For example, (1.48) is compared with its focus alternatives [I failed] and [I aced]. The Gricean reasoning, then, amounts to taking every focus alternative of LF to be false, unless it is contextually entailed by LF itself. For example, (1.48) implicates that I did not ace (if I had, I would have said so).

Similarly, (1.49) is compared with its focus alternatives [nobody passed], [Mats passed], [Paul passed], [Steve and Mats passed], [Steve and Paul passed], [Mats and Paul passed], and [Steve, Mats and Paul passed]. The alternatives that are not contextually entailed by (1.49) are taken to be false, resulting in the implicature that Mats and Paul did not pass.[10]

We are now ready to consider the ambiguity in (1.3), repeated below:

(1.3) MAX called his mother.

The pronoun can either be bound or referential. Let us first take it to be referential, with [Max] as its antecedent:

(1.50) $[\text{Max}]_{\text{F}}$ [called his mother] his = Max

Now suppose that [John], [Bill], and [Fred] are the contextually salient alternatives of [Max]. Then the focus alternatives of (1.50) are:

(1.51)
a. [John] [called his mother] his = Max
b. [Bill] [called his mother] his = Max
c. [Fred] [called his mother] his = Max

These alternatives are not contextually entailed by (1.50), so they are taken to be false. In other words, (1.50) implicates that John, Bill, and Fred did not call Max's mother. This is indeed one of the possible readings of (1.3).

Now suppose that the pronoun in (1.3) is bound by [Max]:

(1.52) $[\text{Max}]^1_{\text{F}}$ [t_1 called his_1 mother]

Suppose again that [John], [Bill], and [Fred] are the contextually salient alternatives of [Max]. Then the focus alternatives of (1.52) are:

[10] This is a simplified picture of course. For a more complete story about implicatures see Davis (2005) and the references given there.

(1.53) a. [John]^1 [t_1 called his_1 mother]
b. [Bill]^1 [t_1 called his_1 mother]
c. [Fred]^1 [t_1 called his_1 mother]

These alternatives are not contextually entailed by (1.52), so they are taken to be false. In other words, (1.52) implicates that John, Bill, and Fred did not call their own mother. This is the second possible reading of (1.3).

Thus, we may conclude that the ambiguity in (1.3) is explained in a natural way if the basic framework presented here is combined with a standard theory of focus and implicature.

1.7 Only

Next, let us consider the interpretation of focus-sensitive operators. In fact, I will add one of them, *only*, to our basic fragment. But let me first illustrate why operators like *only* are called focus-sensitive. Consider the following sentences:

(1.54) John only introduced BILL to Sue.

(1.55) John only introduced Bill to SUE.

These sentences illustrate that different intonation patterns in the scope of *only* lead to different interpretations: (1.54) says that John introduced Bill, and no one else, to Sue, while (1.55) says that John introduced Bill to Sue, and to no one else. This is why *only* is called a focus-sensitive operator.

Next, let us consider the syntactic distribution of *only*.

(1.56) Bill called only SUE.

(1.57) Bill only called SUE.

(1.58) Bill only CALLED Sue.

(1.59) Only BILL called Sue.

(1.60) *Only Bill CALLED Sue.

(1.61) *Only Bill called SUE.

(1.56), (1.57), and (1.58) show that *only* can adjoin both to NP and to VP. In (1.59), *only* could be analyzed as adjoined to NP or as adjoined to S. (1.60) and (1.61) seem to suggest that *only* cannot be adjoined to S. Together with the assumption that *only* must associate with some focused element in the phrase to which it adjoins, this would explain why (1.60) and (1.61) are ungrammatical. I don't know of any alternative explanation for this fact.

However, other examples seem to suggest that it *is* possible for *only* to adjoin to S. Consider the following scenario, adapted from (Jacobson, 2007): every year, I have a large number of people over for Thanksgiving. I am very grumpy about the fact that in general people don't help out enough and don't bring enough food. I turn to you and ask:

(1.62) Do you think anyone will help out this year? Will anyone bring some extra turkey? Some salad or some wine? Or at least some extra chairs?

You answer:

(1.63) I'm afraid only SUE will bring some SALAD this year.

This sentence does not mean that Sue is the only one who will bring some salad this year, but rather that Sue will bring some salad and that nobody else will bring anything else. To accommodate such cases I will assume that *only* may adjoin to S as well as to NP and VP, and that there is an alternative explanation for the ungrammaticality of (1.60) and (1.61). Thus let us add the following rules to the syntax of our fragment:

(PS 10)	S	$\rightarrow$	only S
(PS 11)	VP	$\rightarrow$	only VP
(PS 12)	NP	$\rightarrow$	only NP

Next let us consider the semantics of *only*. First, consider the case of [only S]. Intuitively, a phrase like [only Bill_F loves Mary] is true iff [Bill_F loves Mary] is true and all the focus alternatives of [Bill_F loves Mary] are false. Formally:

(1.64) $[\![\text{only S}]\!]^{\text{C}} = \lambda w.\varphi(w) \wedge \neg\psi_1(w) \wedge \ldots \wedge \neg\psi_n(w)$

where φ is $[\![\text{S}]\!]^{\text{C}}$ and $\psi_1, \ldots, \psi_n$ are all the elements of $[\![\text{S}]\!]^{\text{C,F}}$. If the S in question is [Bill_F loves Mary], and the alternatives of [Bill] are [John] and [Fred], then:

(1.65)
a. $\varphi = \lambda w.\textsc{loves}(w)(\textsc{mary}(w))(\textsc{bill}(w))$
b. $\psi_1 = \lambda w.\textsc{loves}(w)(\textsc{mary}(w))(\textsc{john}(w))$
c. $\psi_2 = \lambda w.\textsc{loves}(w)(\textsc{mary}(w))(\textsc{fred}(w))$

This yields the following translation of [only Bill_F loves Mary]:

(1.66)
$$\begin{aligned}&\lambda w.\textsc{loves}(w)(\textsc{mary}(w))(\textsc{bill}(w))\\ &\wedge \neg\textsc{loves}(w)(\textsc{mary}(w))(\textsc{john}(w))\\ &\wedge \neg\textsc{loves}(w)(\textsc{mary}(w))(\textsc{fred}(w))\end{aligned}$$

which indeed matches our intuitions about the meaning of *only*.[11]

Now consider the case of [only VP]. Intuitively, [only likes Bill_F] expresses a property which holds of all individuals who like Bill, and no other contextually salient individuals. In other words, [only likes Bill_F] expresses a property which

[11] Again, this is of course a simplified picture. For more details on the meaning of *only* see (Ippolito, 2007; van Rooij and Schulz, 2007) and the references given there.

holds of all individuals who have the property expressed by [likes Bill$_\mathrm{F}$] and who do not have the properties expressed by the focus alternatives of [likes Bill$_\mathrm{F}$]. Formally:

(1.67) $[\![\text{only VP}]\!]^{\mathrm{C}} = \lambda w.\lambda x.\varphi(w)(x) \wedge \neg\psi_1(w)(x) \wedge \ldots \wedge \neg\psi_n(w)(x)$

where φ is $[\![\text{VP}]\!]^{\mathrm{C}}$ and $\psi_1, \ldots, \psi_n$ are all the elements of $[\![\text{VP}]\!]^{\mathrm{C,F}}$. If the VP in question is [likes Bill$_\mathrm{F}$], and the alternatives of [Bill] are [John] and [Mary] then:

(1.68) a. $\varphi = \lambda w.\lambda x.\textsc{likes}(w)(\textsc{bill}(w))(x)$
b. $\psi_1 = \lambda w.\lambda x.\textsc{likes}(w)(\textsc{john}(w))(x)$
c. $\psi_2 = \lambda w.\lambda x.\textsc{likes}(w)(\textsc{mary}(w))(x)$

which results in the following translation of [only likes Bill$_\mathrm{F}$]:

(1.69) $\lambda w.\lambda x.\textsc{likes}(w)(\textsc{bill}(w))(x)$
$\wedge \neg\textsc{likes}(w)(\textsc{john}(w))(x)$
$\wedge \neg\textsc{likes}(w)(\textsc{mary}(w))(x)$

Finally let us consider the case of [only NP]. As an example, consider [only Sue$_\mathrm{F}$ sleeps]. Intuitively, this means that Sue sleeps, and that other contextually salient individuals do not sleep. Or in other words, that the individual denoted by [Sue]$_\mathrm{F}$ sleeps, while the individuals denoted by all the focus alternatives of [Sue]$_\mathrm{F}$ do not sleep. Formally:

(1.70) $[\![\text{only NP}]\!]^{\mathrm{C}} = \lambda w.\lambda P.P(w)\varphi(w) \wedge \neg P(w)\psi_1(w) \wedge \ldots \wedge \neg P(w)\psi_n(w)$

where φ is $[\![\text{NP}]\!]^{\mathrm{C}}$ and $\psi_1, \ldots, \psi_n$ are all the elements of $[\![\text{NP}]\!]^{\mathrm{C,F}}$. Notice that the translation of [only NP] is not of type *se* but of type $s((et)t)$ (it's a generalized quantifier). If the NP in question is [Sue]$_\mathrm{F}$, and the alternatives of [Sue] are [Fred] and [Bill], then:

(1.71) a. $\varphi = \lambda w.\textsc{sue}(w)$
b. $\psi_1 = \lambda w.\textsc{fred}(w)$
c. $\psi_2 = \lambda w.\textsc{bill}(w)$

which results in the following translation of [only Sue$_\mathrm{F}$ sleeps]:

(1.72) $\lambda w.\textsc{sleeps}(w)(\textsc{sue}(w))$
$\wedge \neg\textsc{sleeps}(w)(\textsc{fred}(w))$
$\wedge \neg\textsc{sleeps}(w)(\textsc{bill}(w))$

Given this treatment of *only* we may now turn to the ambiguity in (1.4), repeated below:

(1.4) Only MAX called his mother.

Suppose that the alternatives of [Max] are [John] and [Bill]. Now, if the pronoun in (1.4) is referential, with [Max] as its antecedent, then we obtain the following translation:

(1.73) $\lambda w.\textsc{called}(w)(\iota x.\textsc{mother}(w)(\textsc{max}(w))(x))(\textsc{max}(w))$
$\wedge\neg\textsc{called}(w)(\iota x.\textsc{mother}(w)(\textsc{max}(w))(x))(\textsc{john}(w))$
$\wedge\neg\textsc{called}(w)(\iota x.\textsc{mother}(w)(\textsc{max}(w))(x))(\textsc{bill}(w))$

In words: Max called his mother, and the others did not call Max's mother. If the pronoun in (1.4) is bound by [only Max], then we obtain the following translation:

(1.74) $\lambda w.\textsc{called}(w)(\iota x.\textsc{mother}(w)(\textsc{max}(w))(x))(\textsc{max}(w))$
$\wedge\neg\textsc{called}(w)(\iota x.\textsc{mother}(w)(\textsc{john}(w))(x))(\textsc{john}(w))$
$\wedge\neg\textsc{called}(w)(\iota x.\textsc{mother}(w)(\textsc{bill}(w))(x))(\textsc{bill}(w))$

In words: Max called his mother, and the others didn't call their own mother.

Thus we may conclude that the ambiguity that arises in (1.4) is explained in a natural and straightforward way if our basic framework is combined with a simple theory of the interpretation of focus-sensitive operators like *only*.

Finally, let us turn to the ambiguity in constructions such as (1.5):

(1.5) Max called his mother and Bob did too.

In order to explain this ambiguity, we need a basic theory of VP ellipsis.

1.8 VP ellipsis

It is often assumed that ellipsis is the result of deleting certain material at PF (cf. Sag, 1976; Heim and Kratzer, 1998; Merchant, 2001). The exact conditions under which such deletion is licensed is subject to an ongoing debate. The present framework may shed some new light on this debate.

LF Identity. Sag (1976) and Williams (1977) proposed that a constituent may only be deleted at PF if it is identical to some other constituent at LF (which itself is *not* deleted at PF). This constraint, which is known as the LF Identity condition, can still be found in many textbooks (cf. Heim and Kratzer, 1998).

1.16. Definition. [LF Identity]
A constituent may be deleted at PF only if it is identical to another constituent at LF, which itself is not deleted at PF.

Semantic Identity. However, Sag and Hankamer (1984) already observed that the following examples are problematic for LF Identity:

(1.75) Do you think they will like me? - Of course they will.

(1.76) Could you come over here, please? - Of course I could.

The elided VP in (1.75) is [like you], while its antecedent is [like me]. Similarly, the elided VP in (1.76) is [come over there], while its antecedent is [come over here]. Both are legitimate cases of ellipsis, even though the LF representation of the elided VP differs from the LF representation of its antecedent VP. Sag and Hankamer concluded from this data that LF Identity is not really what is at stake. Rather, the relevant identity constraint must be *semantic*: the elided VP and the antecedent VP must be semantically equivalent.

Now recall that in the present framework, there are two notions of semantic equivalence: contextual equivalence (relative to the given context) and conventional equivalence (relative to *some* context). Thus we may define the following two Semantic Identity conditions.

1.17. Definition. [Strong Semantic Identity]
A constituent may be deleted at PF only if it is contextually equivalent to another constituent at LF, which itself is not deleted at PF.

1.18. Definition. [Weak Semantic Identity]
A constituent may be deleted at PF only if it is conventionally equivalent to another constituent at LF, which itself is not deleted at PF.

To see which of these conditions is more adequate consider the following example:

(1.77)
a. Sue: You won't believe what Sam just told me.
b. Ann: What?
c. Sue: John wants to marry his sister, and Bill does too.

Suppose that the pronoun in the antecedent VP in (1.77c) is resolved to Sam. Then the antecedent VP as a whole is interpreted as *wants to marry Sam's sister*, and the elided VP must also be interpreted as *wants to marry Sam's sister*. This is correctly predicted if the elided VP is required to be contextually equivalent with the antecedent VP (Strong Semantic Identity). If mere conventional equivalence were required (Weak Semantic Identity), then the elided VP could just as well be interpreted as *wants to marry John's sister* or *wants to marry Bill's sister*. So Strong Semantic Identity seems more adequate than Weak Semantic Identity. Or in other words, the notion of semantic equivalence relevant for VP ellipsis seems to be *contextual* equivalence rather than *conventional* equivalence.

Strong Semantic Identity accounts for Sag and Hankamer's examples if Kaplan's (1989) semantics for indexicals is adopted. It also solves another problem for LF Identity, which was discussed by Fiengo and May (1994):

(1.78) Mary loves John, and he thinks that Sally does too.

This sentence has a reading on which John thinks that Sally loves him too. Two possible LFs that would correspond to this reading are:

(1.79) a. Mary [loves John], and he thinks that Sally [loves John], too.

he = John

b. Mary [loves John], and he thinks that Sally [loves him], too.

he = him = John

LF Identity only admits (1.79a) as a possible LF of (1.78), because in (1.79b) the elided VP and its antecedent are not identical. The problem with (1.79a) is that, if the elided VP were *not* elided, then *he* could never be interpreted as referring to John.

(1.80) He thinks that Sally loves John, too. $\Rightarrow$ he $\neq$ John

It would be hard to explain how VP ellipsis suddenly makes this interpretation available. This problem does not arise for Strong Semantic Identity, which accepts (1.79b) as a possible LF of (1.78).

Focus Match. However, Strong Semantic Identity on its own is sometimes not strong enough to rule out illegitimate cases of VP ellipsis. To see this, consider the following example, adapted from Rooth (1992a):

(1.81) John's sister thinks he might have a chance, and Bill does too.

(1.82) [John]^1 [[t_1's sister] [thinks he$_1$ might have a chance]], and

[Bill]^1 [t_1 [thinks he$_1$ might have a chance]] too.

The VP in gray has an identical antecedent, so as far as Strong Semantic Identity is concerned, ellipsis is licensed. The problem is that the second conjunct of (1.81) can only be taken to mean that Bill thinks that John might have a chance, *not* that Bill thinks he himself might have a chance, which is the reading represented by (1.82). This observation can be accounted for by adopting the following additional constraint on VP ellipsis (cf. Rooth, 1992a; Tancredi, 1992; Heim, 1997; Tomioka, 1997; Fox, 1999b; Merchant, 2001):

1.19. Definition. [Focus Match]
VP ellipsis is licensed only if the elided VP is dominated by some sentential constituent S_E which focus-matches some other sentential constituent S_A (its *antecedent*). S_E focus-matches S_A iff S_A contextually entails an element of the focus value of S_E.

Intuitively, Focus Match says that an elided VP must always be contained in a clause that *contrasts appropriately* with another clause in the discourse. Let us see how this idea accounts for the fact that ellipsis is ruled out in (1.82). The first sentential constituent dominating the gray VP is [t_1 [thinks he$_1$ might have a chance]]. This phrase does not focus-match any other phrase in the discourse, so it does not license ellipsis. Another sentential constituent which dominates the gray VP is the entire second conjunct of (1.82):

(1.83) $[\text{Bill}]^1$ $[t_1$ [thinks he_1 might have a chance]]

But again, this phrase does focus-matches any other phrase in the discourse, even if we take [Bill] to be F-marked. Conclusion: Focus Match indeed predicts that ellipsis is not licensed in (1.82).

This is all very well, but still we may ask whether Rooth's example really justifies the stipulation of an additional condition on VP ellipsis such as Focus Match. Doesn't the fact that (1.81) does not have a sloppy reading simply follow from the meaning of the particle *too* (which has been ignored so far)? In fact it does: the use of a phrase [S too] roughly requires that some other phrase in the discourse contextually entails one of the focus alternatives of S. This requirement is not fulfilled in (1.82). So to account for Rooth's example we do not need to stipulate Focus Match. However, there are many parallel examples which do not involve the particle *too*:

(1.84) John's sister thinks he might have a chance. Bill_F doesn't_F.

(1.85) John's sister thinks he might have a chance, because Bill_F does.

(1.86) John's sister talked to his coach before Bill_F did.

These sentences do not have sloppy readings either, and something like Focus Match is indeed required to account for this fact. In many of the examples below I will ignore the particle *too*, given that it is always possible to construct parallel examples without *too*.

It is important to point out that, even though Focus Match is stated here as a special condition on VP ellipsis, it should really be thought of as a *corollary* of a much more general theory about the encoding of *information structure*. In English, and in many other languages, information structure is encoded by means of intonation (especially accentuation) and by means of word order. There are also languages in which information structure is encoded by means of special morphemes (cf. Büring, 2007). How a theory of information structure should be formulated exactly, and how something like Focus Match should follow from it, is of course subject to a large ongoing debate (cf. Schwarzschild, 1999; Tomioka, 1997). Here I will abstract away from this debate and simply assume that VP ellipsis must comply with Focus Match.

Semantic Identity Reconsidered. Roughly put, the difference between Strong and Weak Semantic Identity is this: Strong Semantic Identity forces a referential pronoun in an elided VP to refer to the same individual as the corresponding pronoun in the antecedent VP; Weak Semantic Identity does not force this, it allows referential pronouns in the elided VP to "shift" their reference to another individual. Example (1.77) was meant to show that such shifts in reference are generally not allowed. This lead us to the conclusion that Strong Semantic Identity should be adopted rather than Weak Semantic Identity. But once Focus Match is adopted, this conclusion should be reconsidered: Focus Match seems to prohibit exactly those kind of shifts in reference that Weak Semantic Identity by itself wrongly permits. For example, in the case of (1.77), the problem with Weak Semantic Identity was that it allows readings like:

(1.87) John wants to marry Sam's sister, and Bill wants to marry John's sister too.

(1.88) John wants to marry Sam's sister, and Bill wants to marry Bill's sister too.

But these readings are ruled out by Focus Match, because the two conjuncts do not contrast appropriately. Thus, once Focus Match is in place, we could reconsider Weak Semantic Identity as an alternative for Strong Semantic Identity. I will not attempt to tease these two options apart. In any case, for all the examples discussed below it does not really matter whether Strong or Weak Semantic Identity is adopted alongside Focus Match. Let me therefore simply say from now on that VP ellipsis is subject to a condition called *VP Identity*, meaning that it is subject to Focus Match and either Strong or Weak Semantic Identity.

Strict and Sloppy Readings. Let us now finally turn to the ambiguity in (1.5).

(1.5) Max called his mother and Bob did too.

a. ...Bob called his own mother too. [sloppy]
b. ...Bob called Max's mother too. [strict]

This ambiguity is now straightforwardly accounted for. First notice that the pronoun in the source clause may be either bound or referential. Suppose it is bound. Then the source clause has the following LF:

(1.89) [Max]^1 [t$_1$ called his$_1$ mother]

By VP Identity, the LF of the target clause must then be:

(1.90) [Bob]^1 [t$_1$ called his$_1$ mother] too

This gives us the sloppy reading in (1.5a). Now suppose the pronoun in the source clause is referential to Max:

(1.91) [Max]^1 [t$_1$ called his mother] his = Max

Then, by VP Identity, the target clause must have either one of the following LFs:

(1.92) [Bob]^1 [t$_1$ called his mother] too his = Max

(1.93) [Bob]^1 [t_1 called Max's mother] too

Both LFs represent the strict reading in (1.5b). Thus, we may conclude that the present framework accounts for all the ambiguities that we started out with at the beginning of this chapter.

1.9 Summary

Let me briefly summarize what has been established in this chapter. We started out by motivating the distinction between bound and referential pronouns. The basic motivation was that pronouns seem to be interpreted as bound variables in constructions like:

(1.1) Every man thinks he will win.

whereas they seem to be interpreted as referential expressions in other constructions such as:

(1.2) John is in good shape. I think he will win.

Furthermore, it was suggested that a distinction between bound and referential pronouns would naturally explain the ambiguity of constructions like:

(1.3) MAX called his mother.

(1.4) Only MAX called his mother.

(1.5) Max called his mother and Bob did too.

Next, we set out to formulate a formal framework in which bound and referential pronouns are clearly distinguished: a bound pronoun comes with an index and is translated as a variable with that same index, a referential pronoun inherits its translation from its antecedent, which is determined contextually. To explain the ambiguities in (1.3)-(1.5), a basic theory of focus, implicature, focus-sensitive operators like *only*, and VP ellipsis was presented. The ambiguities were accounted for in a straightforward way, and in passing it was observed that the present framework may shed some new light on the identity condition that is supposed to govern VP ellipsis.

The next chapter will start to explore constraints on binding and coreference.

Chapter 2

Constraints on Binding and Coreference: State of the Art

Binding and coreference are constrained in interesting ways. This chapter discusses some of the most prominent existing accounts of such constraints: those of Reinhart (1983), Heim (1998), Fox (1999a), Büring (2005b), and Reinhart (2006). Before turning to the theories proper, however, let me first review the basic data that is to be accounted for.

2.1 Basic Data

Section 2.1.1 discusses Condition B effects, section 2.1.2 discusses Crossover effects, and section 2.1.3 discusses Dahl's puzzle.

2.1.1 Condition B Effects

Binding is subject to so-called *Condition B* effects:

2.1. GENERALIZATION. [Condition B effects on Binding][1]
Pronouns cannot be bound by their coarguments.

2.2. DEFINITION. [Coarguments]
Coarguments are NPs whose θ-role and/or case is assigned by the same predicate.

Here are some sentences which exhibit Condition B effects:

(2.1) a. Every girl loved her.
b. Every man sent a letter to him.

[1] Several versions of this generalization have been proposed in the literature. The present formulation, adapted from Büring (2005a, pp.55–56), is relatively theory-neutral and covers the relevant data. The term *Condition B* originates from Chomsky's (1981) binding theory.

c. Every woman believes her to be a great dancer.
d. Mary asked every boy to wash him.

Coreference is very often also subject to Condition B effects. That is, it is usually impossible for pronouns to corefer with their coarguments. For example, coreference is impossible in the following examples.

(2.2) a. Susan loved her. $\Rightarrow$ her $\neq$ Susan
b. Tom sent a letter to him. $\Rightarrow$ him $\neq$ Tom
c. Norah believes her to be a great dancer. $\Rightarrow$ her $\neq$ Norah
d. Mary asked John to wash him. $\Rightarrow$ him $\neq$ John

Early theories of pronominal anaphora such as (Chomsky, 1981) assumed that coreference is *always* subject to Condition B effects. But Reinhart (1983, p.169) pointed out that there are at least two kinds of environments in which coreference is not subject to Condition B effects. The first kind of environment involves focus-sensitive operators like *only*:

(2.3) Only Max himself voted for him.

The second kind of environment is one in which previous discourse makes it particularly clear that a coreferential interpretation is intended:[2]

(2.4) I know what John and Mary have in common.
John hates Mary and Mary hates her too.

(2.5) If everyone voted for Oscar, then certainly Oscar voted for him.

I must note here that the judgments of my informants do not always confirm those of Reinhart. Many of my informants find that coreference is very marginal in (2.3), (2.4) and (2.5), and emphasize that there are certainly much more natural ways to convey the intended messages. In the recent literature, several authors have acknowledged the controversial status of these data (cf. Schlenker, 2005; Grodzinsky, 2007; Heim, 2007). The theories to be discussed below, however, do take these data very seriously. In fact, they are in large part especially designed to deal with them. Thus, for the sake of the discussion, I will pretend throughout the first part of this dissertation that these data are undisputed. Eventually, in the second part of the dissertation, I will try to account for their borderline-status.[3]

[2] Very similar examples were used by Evans (1980) to show that Condition C effects are suppressed in certain environments, see section 2.1.2 below. Therefore, examples like (2.4) and (2.5) are often attributed to Evans, rather than to Reinhart. I will likewise refer to (2.4) and (2.5) as Evans' examples from now on.

[3] The following example, adapted from Heim (1998), is sometimes taken to instantiate a third set of examples in which coreference is insensitive to Condition B effects.

(i) How can you doubt that the speaker is Zelda? She praises her to the sky.

2.1.2 Crossover Effects

Apart from Condition B effects, binding is also subject to so-called *crossover* effects. The rough idea is this. By definition, a noun phrase A can bind a pronoun P only if A c-commands P at LF. But this does not seem to be enough. For example, the sentences in (2.6) do not have the readings represented by the logical forms in (2.7).

(2.6) a. He likes every man.
b. His mother likes every man.
c. Who does he like?
d. Who does his mother like?

(2.7) a. [every man]^1 [he$_1$ likes t$_1$]
b. [every man]^1 [his$_1$ mother likes t$_1$]
c. [who]^1 [he$_1$ likes t$_1$]
d. [who]^1 [his$_1$ mother likes t$_1$]

These cases are sometimes taken to exemplify the following more general pattern: a noun phrase A can only bind a pronoun P if A already c-commands P in its base position (cf. Koopman and Sportiche, 1982; Reinhart, 1983; Büring, 2005a). In our present terminology, this generalization can be formulated as follows.

2.3. Generalization. [Crossover Effects]
A pronoun can only be cobound with traces that c-command it.

Note that all the logical forms in (2.7) involve a pronoun which is cobound with a trace that does not c-command it. So the generalization correctly predicts that these logical forms are illicit.

However, there are also many counterexamples to the generalization. Higginbotham (1980), for instance, observed that binding is possible in sentences like the following, even though this would involve a pronoun being cobound with a non-c-commanding trace.

(2.8) a. Whose mother loves him?
b. Every senator's portrait was on his desk.
c. Somebody from every city despises its architecture.

These examples seem to show that in order for A to bind P it is not necessary for A itself to c-command P in its base position, but merely for A to *be contained in* another NP which c-commands P in its base position.

However, as Heim notes, even though *she* and *her* may be intended to refer to the same person, they are intended to do so *through different guises*. Technically, they are assigned two distinct individual concepts, which may refer to the same individual in the real world but to distinct individuals in other worlds in the context set. Thus, technically speaking, coreference does not obtain here, and the question whether or not it is subject to Condition B effects does not arise.

But even to this weaker generalization there are several counterexamples. Lasnik and Stowell (1991, p.690) for example, observed that it does not apply to pronouns which occur in *adjuncts*.

(2.9) Who did Joan say she admired in order to please him?

(2.10) Which book did you tell Bill to file without reading it?

And Reinhart (1983, p.180) reported the following type of counterexample, which she attributed to Ross:

(2.11) That people hate him disturbs every president.

So it is quite clear that generalization 2.3 cannot be right. The theories to be discussed below focus on a weaker generalization that does seem to be empirically adequate.

2.4. GENERALIZATION. [Strong Crossover Effects]
A pronoun cannot be cobound with a trace that it c-commands.

This generalization avoids the counterexamples just discussed. The downside of it is that it predicts *some*, but not all alleged crossover effects. For instance, it predicts that the logical forms in (2.7a) and (2.7c) are illicit, but it has nothing to say about the ones in (2.7b) and (2.7d). In section 3.4 I will return to this issue, but in the rest of this chapter, I will assume that generalization 2.4 is the one to account for.

Strong Crossover effects are sometimes considered to be a special case of so-called *Condition C* effects (see especially Chomsky, 1981). The generalization is supposed to be that R-expressions (traces, names, and descriptions) cannot be covalued (cobound or coreferential) with any expression that c-commands them. For example, in (2.7a) and (2.7c) the traces cannot be cobound with the pronouns that c-command them, and in (2.12) below, the name *Max* cannot corefer with the pronouns that c-command it.

(2.12) a. He loves Max. $\Rightarrow$ he $\neq$ Max
b. He called Max's mother. $\Rightarrow$ he $\neq$ Max
c. He says that Mary called Max's mother. $\Rightarrow$ he $\neq$ Max

However, there are many counterexamples to this generalization. For example, coreference is possible in:

(2.13) Whom did the candidates themselves vote for?
Not surprisingly, John voted for John and Bill voted for Bill.

(2.14) I know what John and Mary have in common:
John hates Mary, and Mary hates Mary as well. (cf. Evans, 1980)

(2.15) I think that this is exactly what happened:
Peter forced Tom to call Peter's girlfriend. (cf. Schlenker, 2004)

(2.16) He didn't give her a diamond ring because,
although he's madly in love with her,
Walter's just not ready to tie the knot. (McCray, 1980)

Some of the theories to be discussed below do offer an account of Condition C effects, but none of them deals with all these counterexamples.[4] I think that Condition C effects are the result of various mechanisms working in tandem. This point will be discussed briefly in appendix A. In the present chapter Condition C effects will be left out of consideration.

2.1.3 Dahl's Puzzle

The following type of sentence, first discussed by Östen Dahl (1974), represents a notorious puzzle concerning the interpretation of VP ellipsis.

(2.17) Max said that he called his mother. Bob did too.

a. ...Bob too said that Bob called Bob's mother.
b. ...Bob too said that Max called Max's mother.
c. ...Bob too said that Bob called Max's mother.
d. #...Bob too said that Max called Bob's mother.

The challenge is to account for the fact that (2.17a), (2.17b), and (2.17c) are possible readings of the target clause, while (2.17d) is not. The possible logical forms of the source clause, *Max said that he called his mother*, are given in (2.18) (we are only interested here in those logical forms in which both *he* and *his* are anaphorically related to *Max*). I have also indicated which reading of the target clause is associated with each of these logical forms, assuming that VP ellipsis is governed by VP Identity.

(2.18) Max said that he called his mother.

a.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}_1]^2$ $[\text{t}_2$ called his_2 mother$]]$		(2.17a)
b.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}_1]^2$ $[\text{t}_2$ called his_1 mother$]]$		(2.17a)
c.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}]^2$ $[\text{t}_2$ called his mother$]]$	he=his=Max	(2.17b)
d.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}]^2$ $[\text{t}_2$ called his_2 mother$]]$	he=Max	(2.17b)
e.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}_1]^2$ $[\text{t}_2$ called his mother$]]$	his=Max	(2.17c)
f.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}]^2$ $[\text{t}_2$ called his_1 mother$]]$	he=Max	(2.17d)

[4]Notice that (2.16) also falsifies a weaker version of Condition C, which says that R-expressions cannot be covalued with c-commanding *pronouns*. It is widely known that the original formulation of Condition C is problematic, but this weaker version is usually assumed to be valid. Counterexamples such as McCray's are not often acknowledged. I thank Anna Szabolcsi for pointing me to a manuscript by Peter Sells (1987), which pays special attention to McCray's example.

To account for Dahl's puzzle, a theory of anaphora must do two things: first, it must rule out (2.18f) as a logical form of the source clause, and thus (2.17d) as a possible reading of the target clause. Second, it must allow enough logical forms of the source clause to derive each of the legitimate readings of the target clause. In particular, it should not rule out both (2.18a) and (2.18b) (because this would make (2.17a) unavailable) or both (2.18c) and (2.18d) (this would make (2.17b) unavailable) or (2.18e) (this would make (2.17c) unavailable). This pattern cannot be accounted for in terms of Condition B effects or Strong Crossover effects, so there must be certain additional restrictions on binding and/or coreference.

2.1.4 Summary

The data can be summarized as follows. Binding is subject to Condition B and Strong Crossover effects. Coreference is usually also subject to Condition B effects, but there are some specific environments in which it is not. Finally, there must be certain additional restrictions on binding and/or coreference to account for the pattern found in Dahl's puzzle.

In the remainder of this chapter I will discuss some of the most prominent existing accounts of Condition B effects, Strong Crossover effects, and Dahl's puzzle. The first two proposals, those of Reinhart (1983) and Heim (1998), are primarily concerned with Condition B effects. Fox (1999a), Büring (2005b), and Reinhart (2006) are also concerned with Strong Crossover effects and Dahl's puzzle.

2.2 Reinhart's Coreference Rule

Tanya Reinhart (1983) argues that binding and coreference are fundamentally different. Binding relations, she assumes, are encoded in the syntax (by means of indices) and are subject to *grammatical* constraints. Coreference is not encoded in the syntax, but rather established *contextually*. Therefore, coreference cannot be subject to grammatical constraints. Rather, restrictions on coreference are of a *pragmatic* nature.

The first ingredient of Reinhart's theory of Condition B effects, then, is a grammatical constraint on binding. I will simply call this constraint *Condition B* here, and formulate it in a relatively theory-neutral way (cf. Büring, 2005a, pp.55–56):[5]

2.5. Definition. [Condition B]
Pronouns cannot be bound by their coarguments.

[5] There is an ongoing debate in the literature about how this constraint should be formulated exactly (cf. Pollard and Sag, 1992; Reinhart and Reuland, 1993; Büring, 2005a), and about whether it can be derived from more general syntactic principles (cf. Reuland, 2001, 2008). This debate is interesting in its own right, but perpendicular to the discussion here.

Next, Reinhart points out an interesting consequence of the assumption that binding is encoded in the syntax while coreference is not: it is always more *risky* for a speaker to use syntactic structures which contain referential elements than to use syntactic structures in which all the anaphoric links are already encoded. This is because referential elements always have to be resolved by the hearer, and this can go wrong. If all anaphoric relations are syntactically encoded, resolution does not come into play. Reinhart assumes that speakers generally want to avoid any risk of being misinterpreted, and thus always prefer to use syntactic structures which contain *bound* anaphoric elements rather than syntactic structures which contain *referential* anaphoric elements. Only if speakers cannot express the intended meaning using bound anaphora will they use referential elements. This idea can be implemented as follows:

2.6. DEFINITION. [Coreference Rule]
A speaker will never use a logical form LF in a context C if LF is semantically indistinguishable from one of its binding alternatives in C.

2.7. DEFINITION. [Binding Alternatives]
Let C be a context, let LF be a logical form, and let A and B be two noun phrases in LF, such that A and B corefer in C and such that A c-commands B in LF. Then the structure obtained from LF by:

- Quantifier raising A in case it has not been raised yet, and
- Replacing B with a (possibly reflexive) pronoun bound by A

is called a binding alternative of LF in C.

The idea is this: if a logical form LF containing a referential expression B expresses a certain meaning M, and that same meaning could also be expressed by a logical form LF′ which only differs from LF in that B is replaced by an element that is not coreferential with its antecedent but rather *bound* by it, then a speaker will always use LF′ rather than LF to express M.

Let us see how the Coreference Rule deals with the data discussed in section 2.1. First consider a typical Condition B effect:

(2.19) Max washed him.
a. [Max]^1 [t_1 washed him] him = Max
b. [Max]^1 [t_1 washed himself$_1$]

The Coreference Rule predicts that a speaker will never use (2.19a), in which *Max* and *him* corefer, because (2.19a) is semantically indistinguishable from its binding alternative (2.19b): both express the proposition that Max washed Max. A hearer will conclude from this that coreference cannot be intended in (2.19).

Now, let's see whether the Coreference Rule can deal with Condition B environments in which coreference is exceptionally permitted. First consider a focus construction:[6]

(2.20) Only Max himself voted for him.

a. [only] [[Max himself]^1 [t_1 voted for him]] him = Max himself
b. [only] [[Max himself]^1 [t_1 voted for himself$_1$]]

This time, the Coreference Rule does *not* rule out coreference, because the interpretation of (2.20a) *differs* from the interpretation of its binding alternative (2.20b): (2.20a) says that the others did not vote for *Max*, while (2.20b) says that the others did not vote for *themselves*. Finally, consider one of Evans' examples:

(2.21) I know what John and Mary have in common.
Mary voted for John and John voted for him too.

a. [John]^1 [t_1 voted for him] him = John
b. [John]^1 [t_1 voted for himself$_1$]

The Coreference Rule does *not* rule out coreference in (2.21a), because the interpretation of (2.21a) differs from the interpretation of its binding alternative (2.21b): (2.21a) says that John has the property of having voted for John, and this is indeed the property that John and Mary are supposed to have in common. (2.21b) on the other hand, says that John has the property of having voted for *himself*, and this is certainly *not* the property that John and Mary are supposed to have in common.

So Reinhart's Coreference Rule accounts for standard Condition B effects on coreference, and also for the exceptional coreference patterns found in focus constructions and in Evans' examples.

2.3 Heim's Exceptional Codetermination Rule

Irene Heim (1998) observed that at least three aspects of Reinhart's account need some further consideration. First, the theory does not explicitly state what it means for one LF to be *semantically indistinguishable* from another. One possibility that comes to mind immediately is that two LFs should be regarded as semantically indistinguishable if and only if they express the same proposition. But this would not work: (2.21a) and (2.21b) express the same proposition, but intuitively, at least in the context of (2.21), there is a significant semantic difference between them. So the question of when two LFs should be regarded as semantically indistinguishable is not trivial and should be addressed with care.

[6] For reasons of readability, F-marking has been suppressed here and will often be suppressed below.

Second, the Coreference Rule is not compatible with the VP Identity condition on VP ellipsis. To see this, consider the basic case of VP ellipsis in (2.22):

(2.22) Max called his mother and Bob did too.

We saw before that the pronoun in the source clause can either be bound, as in (2.23a), or referential, as in (2.23b). (2.23a) gives rise to the sloppy reading of the target clause (*Bob called Bob's mother*) while (2.23b) gives rise to the strict reading of the target clause (*Bob called Max's mother*).

(2.23) a. [Max]^1 [t_1 called his_1 mother]
b. [Max]^1 [t_1 called his mother] his = Max

But the Coreference Rule predicts that a speaker will never use (2.23b), because it is semantically indistinguishable from its binding alternative (2.23a). This means that, as long as the VP Identity condition is assumed, the Coreference Rule wrongly predicts that the target clause in (2.22) does not have a strict reading.

The third issue with the Coreference Rule is that it is only concerned with coreference. Other kinds of anaphora, such as cobinding, are just as unacceptable as coreference in typical Condition B environments, and we would like to have a rule that embodies these restrictions all in one go, rather than separate rules for coreference, cobinding, and possibly yet other kinds of anaphora. Examples (2.24) and (2.25) illustrate this point.

(2.24) [John voted for him] him = John

(2.25) [Every man]^1 [t_1 said that he_1 voted for him_1]

The Coreference Rule correctly rules out coreference between [John] and [him] in (2.24), but fails to rule out cobinding of [he_1] and [him_1] in (2.25). Intuitively, these cases are analogous and should be ruled out by one and the same mechanism.

Heim's contribution, then, is twofold: first, she refines the notion of semantic indistinguishability. Second, she proposes a new constraint which preserves all the empirical virtues of Reinhart's Coreference Rule, is compatible with the VP Identity condition, and applies not only to coreference but also to cobinding and other kinds of anaphora. Let me first discuss the new constraint.

2.3.1 The Exceptional Codetermination Rule

Heim's theory is stated in terms of *codetermination*, a notion which embraces that of binding, cobinding, and coreference (and yet other anaphoric relations as well).

2.8. Definition. [Codetermination]
Let C be a context, let LF be a logical form, and let A and B be two NPs in LF. We say that A and B are codetermined in LF/C iff:

- A binds B in LF, or
- A and B corefer in C, or
- There is a third NP which is codetermined with A and B in LF/C.

The first ingredient of Heim's theory is a revised version of Condition B, which prohibits codetermination, rather than binding.

2.9. Definition. [Heim's Condition B]
Pronouns cannot be codetermined with their coarguments.

The second ingredient of the theory is a rule which states that codetermination is sometimes exceptionally allowed.

2.10. Definition. [Exceptional Codetermination Rule]
Let LF be a logical form in which a pronoun is codetermined with, but not bound by one of its coarguments. Then, LF is (marginally) allowed, in violation of Condition B, if it is semantically distinguishable from its binding alternative in the given context.

The reader is invited to check that Heim's Condition B and her Exceptional Codetermination Rule account for the standard Condition B effects, not only involving coreference, but also involving cobinding and other kinds of codetermination. The proposal also accounts for the exceptional cases in which codetermination *is* allowed in Condition B environments. Finally, it is compatible with VP Identity: (2.23b) is no longer ruled out.

2.3.2 Semantic Indistinguishability

When should two logical forms be regarded as semantically indistinguishable? For one thing, they should express the same proposition. But Heim notes that Evans' examples, repeated in (2.26) and (2.27), show that there is more to it.

(2.26) I know what John and Mary have in common.
Mary voted for John and John voted for him too.

(2.27) If everyone voted for Oscar, then certainly Oscar voted for him.

Heim suggests that these are typical cases in which *structured meaning* matters. In (2.26), there is a certain property P that Mary and John are supposed to have in common, namely, the property $[\lambda x.\ x$ voted for John$]$ of having voted for John. If such a particular property is under discussion, then an LF which says that John has the property P is to be distinguished from an LF which says that John has the property $[\lambda x.\ x$ voted for $x]$ (even though these two LFs as a whole denote the same proposition). The same reasoning can be applied to the example in (2.27), where it is the property of having voted for Oscar that is under discussion. This intuition can be implemented by defining semantic indistinguishability not only in terms of propositional content, but also in terms of *focus* values:[7]

2.11. Definition. [Semantic Indistinguishability]
Two logical forms LF and LF′ are semantically indistinguishable iff:

1. LF and LF′ express the same proposition, and
2. LF and LF′ have the same focus value.

Let me illustrate how this works for one of Evans' examples. Consider the final clause in (2.27), given in (2.28):

(2.28) ...then certainly Oscar voted for him.

Let us assume that [Oscar] is F-marked (notice that it must be accented in the given context). Now consider the logical form in (2.29a), where [him] corefers with [Oscar], and its binding alternative in (2.29b):

(2.29) a. $[\text{Oscar}]^1_F$ [t_1 voted for him] him = Oscar
b. $[\text{Oscar}]^1_F$ [t_1 voted for himself$_1$]

(2.29a) and (2.29b) express exactly the same proposition. However, their focus values differ. For example, (2.30) is a focus alternative of (2.29a) but not of (2.29b). Thus, (2.29a) is semantically distinguishable from its binding alternative.

(2.30) $[\text{Fred}]^1$ [t_1 voted for him] him = Oscar

Interestingly, the notion of meaning explored by Groenendijk (2007) comprises both the information an expression provides (≈ its propositional content) and the alternatives it gives rise to or presupposes (≈ its focus value). So in Groenendijk's system, "being semantically indistinguishable" in the sense defined here really comes down to "having the same meaning".

The three issues that Heim raised concerning Reinhart's original proposal are now resolved. It must be noted, however, that the explanatory aspect of Reinhart's theory has been lost: Heim's theory cannot be derived from the assumption that speakers generally seek to avoid risks of being misinterpreted. Furthermore, the theory does not yet account for Strong Crossover effects and Dahl's puzzle.

[7] Heim herself does not provide a concrete implementation, but I am quite convinced that the implementation given here is in line with what she has in mind (see also Heim, 2007).

2.4 Fox's Locality Constraint

Consider Dahl's puzzle again:

(2.17) Max said that he called his mother. Bob did too.

a. ... Bob too said that Bob called Bob's mother.
b. ... Bob too said that Max called Max's mother.
c. ... Bob too said that Bob called Max's mother.
d. #... Bob too said that Max called Bob's mother.

The possible logical forms of the source clause are repeated in (2.18):

(2.18) Max said that he called his mother.

a.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his_2 mother]]		(2.17a)
b.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his_1 mother]]		(2.17a)
c.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his mother]]	he=his=Max	(2.17b)
d.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his_2 mother]]	he=Max	(2.17b)
e.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his mother]]	his=Max	(2.17c)
f.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his_1 mother]]	he=Max	(2.17d)

Heim's account does not rule out any of the logical forms in (2.18). In particular, it wrongly permits (2.18f), and thus (2.17d) as a possible reading of the target clause in (2.17). Reinhart's account does rule out (2.18f), but it also rules out (2.18c), (2.18d), and (2.18e) (all the logical forms that involve coreference), and thus wrongly predicts that (2.17a) is the only reading for the target clause in (2.17). Thus, an additional constraint is required to account for Dahl's puzzle.

Danny Fox (1998) proposed that the puzzle may be explained in terms of so-called *economy conditions*. The idea is that such economy conditions prevent the grammar from generating logical forms whose interpretation is identical to that of simpler alternative logical forms. After all, why would the grammar generate complicated LFs if the meaning they represent can just as well be represented by simpler alternatives?

If this is indeed how the grammar works, then the next question to ask is what the criteria for simplicity are. Why should one LF be considered simpler than another? Fox suggests that one natural criterion involves the length of binding dependencies: one LF is simpler than another if its binding dependencies are shorter. This idea can be implemented as follows:

2.12. Definition. [Locality]
A logical form is ruled out if it is semantically indistinguishable from one of its locality alternatives.

2.13. Definition. [Locality Alternatives]
Let LF be a logical form, and let A, B and P be three nodes in LF such that A

c-commands B and B c-commands P, and such that A, but not B, binds P. Then the structure obtained from LF by:

- Quantifier raising B in case it has not been raised yet;
- Adjusting P's binding index so that it's bound by B instead of A.

is a locality alternative of LF.

Locality rules out (2.18b) and (2.18f), and thus accounts for Dahl's puzzle ((2.18b) is ruled out because it is semantically indistinguishable from its locality alternative in (2.18a) and (2.18f) is ruled out because it is semantically indistinguishable from its locality alternative in (2.18d)).

To further support his theory, Fox (1999a, p.131) discusses two striking cases in which Dahl's puzzle is obviated. First, if we slightly change the source clause, as in (2.31), then the fourth reading of the target clause suddenly *is* available.

(2.31) Max said that his mother called him. Bob did too.
  a. ... Bob too said that Bob's mother called Bob.
  b. ... Bob too said that Max's mother called Max.
  c. ... Bob too said that Bob's mother called Max.
  d. ... Bob too said that Max's mother called Bob.

This is accounted for by Locality. In (2.17), the LF in which [he] corefers with [Max] and [his] is bound by [Max] is illegitimate, because it is semantically indistinguishable from its locality alternative, in which [his] is bound by [he] instead of [Max]. However, this argument does *not* carry over to (2.31): the LF of (2.31) in which [his] corefers with [Max] and [him] is bound by [Max] does not have a locality alternative, because [his] does not c-command [him] in (2.31).

A second case in which Dahl's puzzle is obviated, is obtained by modifying the embedded clause in the source clause with *only*:

(2.32) Max said that only he had called his mother. Bob did too.
  a. ... Bob too said that only Bob had called Bob's mother.
  b. ... Bob too said that only Max had called Max's mother.
  c. ... Bob too said that only Bob had called Max's mother.
  d. ... Bob too said that only Max had called Bob's mother.

According to Fox, (2.32d) is an available reading of (2.32).[8] This is accounted for by Locality: the crucial LF, in which [he] corefers with [Max] and [his] is bound by [Max], conveys that Max said that he, Max, was the only one with

[8] I must remark here that the judgments of my informants do not really confirm those of Fox. There was quite some variability in the judgments, but many informants reacted to (2.32) just as they reacted to Dahl's original example. I will leave this data-issue aside here. Further empirical research is needed to straighten out the facts.

the property [$\lambda x.\ x$ called Max's mother] (the others didn't call Max's mother). The locality alternative of this LF, in which [his] is bound by [he] instead of [Max], conveys that Max said that he, Max, was the only one with the property [$\lambda x.\ x$ called x's mother] (the others didn't call their own respective mothers). These two LFs express different propositions and are therefore *not* semantically indistinguishable. As a consequence, the reading in (2.32d) is not ruled out.

A third case in which Dahl's puzzle is obviated, which was not discussed by Fox, is when Dahl's sentence is uttered in the context of a particular question, such as:

(2.33)
a. Did Max call everyone's mother?
b. Well, I don't know...
c. MAX said he called his mother, and BOB did too.
d. But I haven't heard from SUE and MARY yet.

Notice that (2.33c) is identical to Dahl's original sentence. But in the given context, the target clause can easily be taken to mean that Bob said that Max called Bob's mother. Again, Locality accounts for this fact: in the context of (2.33), the crucial LF differs in focus value from its locality alternative.

Thus, Locality accounts not only for Dahl's original puzzle, but also for some cases in which the puzzle is obviated. The third case, in particular, has two important implications. First, to account for the fact that Dahl's puzzle is obviated in (2.33) it is crucial that semantic indistinguishability is not formalized in terms of propositional content only. Rather, as was already argued on independent grounds in section 2.3.2, the notion should also make reference to *focus values*. This was not made quite explicit in Fox's original proposal.

Second, cases like (2.33) are problematic for alternative accounts of Dahl's puzzle such as Fiengo and May (1994), Kehler (1993), and Schlenker (2005).

Fox (1999a, p.132) notes that Locality does not only account for Dahl's puzzle, but also for Strong Crossover effects. To see this, consider the following example:

(2.34) Who did he say we should invite?
a. [who]^1 [[he$_1$]^2 [t$_2$ said we should invite t$_1$]]
b. [who]^1 [[he$_1$]^2 [t$_2$ said we should invite t$_2$]]

The logical form in (2.34a) should be ruled out, because (2.34) cannot be used to ask who have the property [$\lambda x.x$ said we should invite x]. This is a Strong Crossover effect: the pronoun in (2.34a) is cobound with a trace that it c-commands. (2.34a) is successfully ruled out by Locality, because it is semantically indistinguishable from its locality alternative in (2.34b).

Thus, Locality accounts for Dahl's puzzle, for some striking cases in which Dahl's puzzle is obviated, and also for Strong Crossover effects.

The question that arises next is whether Locality can be combined, or even unified either with Reinhart's Coreference Rule or with Heim's Exceptional Codetermination Rule in order to obtain a theory that accounts for all the data we have seen so far.[9] This question has been addressed by Daniel Büring.

2.5 Büring: Have Local Binding!

We have seen that Heim's account resolved several problems with Reinhart's original account. So, at first sight, the most logical step to take would be to combine Fox's Locality constraint with Heim's Exceptional Codetermination Rule. But Büring (2005b) is more ambitious. He observes that Heim's reinterpretation of Reinhart's approach brought about certain complications. In particular, Condition B had to be reformulated so as to apply to *codetermination* rather than binding.[10] One of the issues that motivated this complication was that Reinhart's Coreference Rule was only concerned with coreference, and not with cobinding and other kinds of anaphora. For example, it failed to rule out cobinding in:

(2.35) [Every man]^1 [t$_1$ said that he$_1$ voted for him$_1$]

Büring observes that, once Fox's Locality constraint is adopted, cobinding in (2.35) *is* ruled out. After all, (2.35) is semantically indistinguishable from its locality alternative (2.36).

(2.36) [Every man]^1 [t$_1$ said that [he$_1$]^2 [t$_2$ voted for him$_2$]]

Thus, Büring argues, the complications proposed by Heim become superfluous. He therefore proposes to return to the simple formulation of Condition B (as in definition 2.5) and adopt Reinhart's original Coreference Rule alongside Fox's Locality constraint. Finally, he observes that the latter two constraints can actually be collapsed into one:

2.14. Definition. [Have Local Binding!]
A logical form is ruled out if it is semantically indistinguishable from one of its HLB alternatives.

2.15. Definition. [HLB Alternative]
Let LF be a logical form, let A and B be two noun phrases in LF such that A c-commands B and neither A nor any node c-commanded by A binds B. Then the logical form obtained from LF by:

- Quantifier raising A in case it has not been raised yet;

[9] That this is indeed a natural question to ask has been obscured a bit by the fact that Fox considered his Locality constraint to be a notational variant of Heim's Exceptional Codetermination Rule. But, as also noticed by Büring (2005b), the two constraints are not equivalent.

[10] I personally think that the main drawback of Heim's proposal was that it did not preserve the explanatory aspect of Reinhart's account. But this does not seem to be what motivated Büring; his own proposal does not preserve the explanatory aspect of Reinhart's account either.

- Replacing B with a pronoun bound by A.

is an HLB alternative of LF.

The reader is invited to check that this constraint (HLB) does indeed account for Condition B and Strong Crossover effects, and also for the exceptional cases in which coreference is not subject to Condition B effects.

But HLB is *not* compatible with VP Identity. That is, if VP Identity is assumed, HLB yields wrong predictions both for the simple cases of VP ellipsis that were already problematic for Reinhart's account from the start, and for the more intricate pattern found in Dahl's puzzle. To see this, first consider the simple case of VP ellipsis in (2.22), repeated here as (2.37):

(2.37) Max called his mother and Bob did too.

HLB predicts that coreference is impossible in the source clause. Thus, as long as VP Identity is assumed, HLB wrongly predicts that the target clause does not have a strict reading (Bob called Max's mother too).

Next, recall Dahl's puzzle:

(2.17) Max said that he called his mother. Bob did too.
a. ...Bob too said that Bob called Bob's mother.
b. ...Bob too said that Max called Max's mother.
c. ...Bob too said that Bob called Max's mother.
d. #...Bob too said that Max called Bob's mother.

(2.18) Max said that he called his mother.

a.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his_2 mother]]		(2.17a)
b.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his_1 mother]]		(2.17a)
c.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his mother]]	he=his=Max	(2.17b)
d.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his_2 mother]]	he=Max	(2.17b)
e.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his mother]]	his=Max	(2.17c)
f.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his_1 mother]]	he=Max	(2.17d)

HLB allows (2.18a) but rules out all the other LFs in (2.18). This means that, as long as VP Identity is assumed, HLB wrongly predicts that (2.17a) is the only possible reading of the target clause in (2.17).

So Büring is forced to depart from VP Identity. Instead, he adopts the following constraint on VP ellipsis, which is originally due to Fox (1999a):[11]

2.16. Definition. [NP Parallelism]
Corresponding noun phrases in the antecedent and elided VPs must either:

[11]This formulation of NP Parallelism is taken from (Büring, 2005a, p.132). A slightly different formulation is given in (Büring, 2005b, p.267). I assume that these two formulations are intended to be equivalent, and use the above because it is slightly more explicit.

- have the same referential value, or
- be bound in parallel in their respective conjuncts.

2.17. Definition. [Referential Value]
The referential value of a noun phrase A is:

- the individual to which A refers, or
- the referential value of the NP that binds A.

Büring does not say explicitly what it means for two noun phrases to be *bound in parallel*. If the constraint were to be evaluated seriously, this would first have to be made more precise of course. But even so, we could ask whether the generalization that seems to be embodied by NP Parallelism is empirically correct. I think it is both too weak and too strong. Let me first consider a case of *cascaded ellipsis* which shows that (what seems to be) the intended generalization is too weak to rule out certain illicit readings of elided VPs.

(2.38) Bob called his mother, and Max did too. But Tom didn't.

NP Parallelism wrongly predicts that (2.38) has a so-called *mixed reading* which can be paraphrased as follows:

(2.39) Bob called Bob's mother, and Max called Max's mother.
But Tom didn't call Max's mother.

Next, consider the following well-known example from Rooth (1992a), which shows that NP Parallelism is not only too weak, but also too strong.

(2.40) First John told Mary that I was bad-mouthing her.
Then Sue heard I was.

NP Parallelism erroneously rules out the sloppy reading of (2.40), which says that Sue heard that I was bad-mouthing her, Sue. This is because such a sloppy reading would involve "non-parallel" binding of the pronouns in the elided VP and its antecedent.

It *might* be possible, of course, to adjust the NP Parallelism constraint in such a way that it becomes strong enough to rule out the undesired readings in (2.38), while still being weak enough to allow the strict reading in (2.37) and the three possible readings in Dahl's puzzle, and even somewhat weaker than it presently is so as to allow for the sloppy reading in (2.40).

The point is that Büring's account of binding and coreference is at best compatible with an ad hoc, still to be worked out, non-standard theory of VP ellipsis.

It would be preferable to have a theory of binding and coreference that is compatible with mainstream theories of ellipsis such as those based on VP Identity.[12] This was, I think, one of the main reasons for Tanya Reinhart to eventually depart from her 1983 theory and develop a new account in the mid-1990's. This account will be discussed next.

2.6 Reinhart's Interface Rule

Reinhart's 1983 theory was based on two assumptions. First, that binding relations are encoded in syntactic structure, while coreference is not; and second, that speakers generally try to avoid risks of being misinterpreted. It follows from these two assumptions that speakers generally prefer to use bound pronouns, which explicitly encode the intended anaphoric relations, rather than referential pronouns, which may well not be resolved as intended.

In her later work, Reinhart concludes that these assumptions, as plausible as they may seem at first, eventually yield the wrong predictions and should be reconsidered. More specifically, she proposes to leave the first assumption intact (binding is encoded by syntactic structure, coreference is not) but replace the second assumption, which is about *speakers*, with an alternative assumption about *hearers*. The general assumption is that hearers *minimize interpretive options*. In the specific case of anaphora, this means that if a certain interpretation is ruled out by grammatical restrictions on binding, then a hearer will recognize that this interpretation was not intended, even if it could in principle be derived via other anaphoric mechanisms. In other words, interpretations which are ruled out by restrictions on binding cannot be *sneaked in* via other anaphoric mechanisms. Reinhart points out that the existence of such a mechanism would be extremely useful. For communication to proceed efficiently, it is crucial for a hearer to keep interpretive options to a minimum at all times.

Reinhart (2006) formalizes this idea in terms of a notion called *covaluation*.

---

[12] It should be remarked that Fox (1999a, p.117) presents a particular case of cascaded ellipsis in support of NP Parallelism:

(i) Smithers thinks that his job sucks. Homer does, too. However, Marge doesn't.

Notice that (i) is structurally analogous to (2.38). However, (i) *does* have a mixed reading, at least for people who recognize that it is about a popular American sitcom, *The Simpsons*, in which Marge is Homer's wife and does not have a job of her own. The availability of such a mixed reading would be in accordance with NP Parallelism, and not with VP Identity. However, I don't think that this really is an argument in favor of NP Parallelism. First, the availability of a mixed reading in (i) is *exceptional*. Typically, mixed readings are not available for cases of cascaded ellipsis, and this is left unexplained by NP Parallelism. Second, the fact that a mixed reading *is* exceptionally available in (i) can, I think, be explained without doing away with the essence of VP Identity. Such an explanation will be discussed in section 5.2.

2.18. Definition. [Covaluation] Let C be a context, let LF be a logical form, and let A and B be two NPs in LF. Then A and B are covalued in LF/C iff:

- A does not bind B and B does not bind A in LF, and
- A and B are cobound in LF or A and B corefer in C.

Notice that covaluation is essentially a generic term for coreference and cobinding. As such, it is more general than coreference alone, but less general than Heim's notion of codetermination, which covered other kinds of anaphora as well. Reinhart proposes the following constraint on covaluation:

2.19. Definition. [Interface Rule]
A logical form LF is ruled out if one of its binding alternatives LF′ is such that:

a. LF and LF′ are semantically indistinguishable, and

b. The transition from LF to LF′ is illicit, because:

- LF′ is ruled out by restrictions on binding (Condition B), or
- The existing binding relations in LF are not preserved in LF′, or
- LF′ is ruled out by another application of the Interface Rule.

The notion of a *binding alternative* has to be revised slightly. The only difference between definition 2.7 and definition 2.20 is that in the former A and B are supposed to *corefer*, while in the latter A and B are supposed to be *covalued.*

2.20. Definition. [Binding Alternatives]
Let C be a context, let LF be a logical form, and let A and B be two noun phrases such that A c-commands B in LF and such that A and B are covalued in LF/C. Then the structure obtained from LF by:

- Quantifier raising A in case it has not been raised yet;
- Replacing B with a pronoun or trace bound by A.

is called a binding alternative of LF.

Let us see whether the Interface Rule accounts for the data accumulated so far. Doing so will sometimes require quite some effort, because, as we will see, the workings of the Interface Rule are sometimes rather complex. Always keep in mind though, that the intuition behind it is very simple: interpretations which are ruled out by restrictions on binding cannot be sneaked in via other anaphoric mechanisms.

Let us first consider (2.19), repeated here as (2.41), which exhibits a basic Condition B effect on coreference. The Interface Rule correctly rules out (2.41a) because it is semantically indistinguishable from its binding alternative in (2.41b), and (2.41b) violates Condition B.

(2.41) Max washed him.

a. [Max]^1 [t_1 washed him] him = Max

b. [Max]^1 [t_1 washed him_1]

Now, let us see whether the Interface Rule can deal with Condition B environments in which coreference is exceptionally permitted. Consider (2.20), repeated here as (2.42):

(2.42) Only Max himself voted for him.

a. [only] [[Max himself]^1 [t_1 voted for him]] him = Max himself

b. [only] [[Max himself]^1 [t_1 voted for him_1]]

The Interface Rule does *not* rule out coreference here, because (2.42a) is not semantically indistinguishable from its binding alternative in (2.42b). Intuitively speaking, coreference is not sneaking in an interpretation that is ruled out by restrictions on binding, but rather gives rise to an interpretation that is different from what would be obtained from binding. Evans' examples are dealt with in a similar way. So the Interface Rule accounts for standard Condition B effects on coreference, and also for the cases in which coreference is exceptionally allowed in Condition B environments.

The issues which the Coreference Rule was facing and which were addressed by Heim's Exceptional Codetermination Rule are also satisfactorily dealt with by the Interface Rule. In particular, the Interface Rule does not only account for cases of illicit coreference, but also for cases of illicit cobinding, and it allows strict identity readings in VP ellipsis. Let me illustrate this with some examples. First consider a case of illicit cobinding:

(2.43) Every man said that he washed him.

a. [Every man]^1 [t_1 said that [he_1]^2 [t_2 washed him_1]]

b. [Every man]^1 [t_1 said that [he_1]^2 [t_2 washed him_2]]

The logical form in (2.43a), in which [he_1] and [him_1] are cobound, should be ruled out. The Interface Rule accounts for this: (2.43a) is ruled out because it is semantically indistinguishable from its binding alternative (2.43b), and (2.43b) violates Condition B.

Next, consider the simple case of VP ellipsis in (2.22), repeated here as (2.44):

(2.44) Max called his mother and Bob did too.

a. ...Bob called his own mother too. [sloppy]

b. ...Bob called Max's mother too. [strict]

The Coreference Rule prohibited coreference in the source clause and thus ruled out the strict reading in (2.44b). The Interface Rule, on the other hand, does *not* rule out coreference in the source clause (coreference is only ruled out if binding is, too, and binding is certainly possible here). Thus, the Interface Rule correctly admits the strict reading in (2.44b).

But the Interface Rule is not just an alternative for Heim's Exceptional Codetermination Rule. It does *more*. In particular, it accounts, at once, for Strong Crossover effects. To see this, consider example (2.34), repeated here as (2.45).

(2.45) Who did he say we should invite?

a. $[\text{who}]^1$ $[[\text{he}_1]^2$ $[\text{t}_2$ said we should invite $\text{t}_1]]$

b. $[\text{who}]^1$ $[[\text{he}_1]^2$ $[\text{t}_2$ said we should invite $\text{t}_2]]$

The logical form in (2.45a) is ruled out by the Interface Rule, because it is semantically indistinguishable from its binding alternative (2.45b), and (2.45b) does not preserve the existing binding relations in (2.45a). In particular, the trace that was bound by the wh-element in (2.45a) is bound by the pronoun in (2.45b).

Finally, Reinhart claims that the Interface Rule also accounts for Dahl's puzzle. To see whether this is indeed the case, let me briefly resume the puzzle:

(2.17) Max said that he called his mother. Bob did too.

a. ... Bob too said that Bob called Bob's mother.

b. ... Bob too said that Max called Max's mother.

c. ... Bob too said that Bob called Max's mother.

d. #... Bob too said that Max called Bob's mother.

(2.18) Max said that he called his mother.

a.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}_1]^2$ $[\text{t}_2$ called his_2 mother$]]$		(2.17a)
b.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}_1]^2$ $[\text{t}_2$ called his_1 mother$]]$		(2.17a)
c.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}]^2$ $[\text{t}_2$ called his mother$]]$	he=his=Max	(2.17b)
d.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}]^2$ $[\text{t}_2$ called his_2 mother$]]$	he=Max	(2.17b)
e.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}_1]^2$ $[\text{t}_2$ called his mother$]]$	his=Max	(2.17c)
f.	$[\text{Max}]^1$ $[\text{t}_1$ said $[\text{he}]^2$ $[\text{t}_2$ called his_1 mother$]]$	he=Max	(2.17d)

The Interface Rule is supposed to do two things: first, it is supposed to rule out (2.18f) as a logical form of the source clause, and thus (2.17d) as a possible reading of the target clause. Second, it is supposed to allow enough logical forms of the source clause to derive each of the legitimate readings of the target clause. In particular, it should not rule out both (2.18a) and (2.18b), or both (2.18c) and (2.18d), or (2.18e).

Let us see whether this is indeed established. First consider (2.18f). This LF is indeed ruled out. To see this, we have to consider the binding alternative of (2.18f), which is (2.18b). First observe that (2.18f) and (2.18b) are semantically indistinguishable. Next, observe that (2.18b) is ruled out *by another application of the Interface Rule*: (2.18b) is semantically indistinguishable from its binding alternative, (2.18a), and (2.18a) does not leave the existing binding relations in (2.18b) intact: [his] is no longer bound by [Max] in (2.18a). So (2.18b) is ruled out by the Interface Rule, and this means that (2.18f) itself is prohibited as well.

So far so good. But the problem is that exactly the same line of reasoning yields that the Interface Rule rules out (2.18e) as well: (2.18e)'s binding alternative is (2.18b), just like that of (2.18f). (2.18e) and (2.18b) are semantically indistinguishable, and we have already seen that (2.18b) is ruled out by the Interface Rule. Thus, (2.18e) must be ruled out by the Interface Rule as well. As a consequence, (2.17c) is wrongly excluded as a possible reading of the target clause.

To sum up, the Interface Rule improves on the Coreference Rule in two ways: it accounts for cases of illicit cobinding, and it avoids the strict identity problem in simple cases of ellipsis. It also improves on Heim's Exceptional Codetermination Rule: it accounts, at once, for Strong Crossover effects. However, it still does not account for the more complex case of VP ellipsis in Dahl's puzzle.

Apart from this remaining empirical problem, there are two other aspects of Reinhart's theory that call for further attention. First, there is a striking discrepancy between (the simplicity of) the intuition behind the Interface Rule and (the complexity of) its actual formulation. Recall the basic intuition: interpretations which are ruled out by restrictions on binding cannot be sneaked in via other anaphoric mechanisms. We should expect, then, that the formal statement of the rule should say something like: a logical form LF is ruled out if it is semantically indistinguishable from one of its binding alternatives LF′, and LF′ is ruled out by constraints on binding. The actual formulation of the Interface Rule is much more complicated. In particular, it additionally requires that the existing binding relations in LF are preserved in LF′ and that LF′ is not ruled out by recursive applications of the Interface Rule.

The second issue that needs further attention is that the Interface Rule is not only undesirably complex in its formulation, but also in its workings. The analysis of sentence (2.46) illustrates this:

(2.46) Max said that he washed him.

This sentence cannot be taken to mean that Max said that he washed *himself*. That is, the co-arguments of *washed* cannot be anaphorically related. A similar intuition applies to simpler examples such as:

(2.47) Max washed him.

In these simpler examples, the intuition is straightforwardly accounted for: Condition B prohibits binding and (as a consequence) the Interface Rule prohibits covaluation. We would like the Interface Rule to deal with the more complex example in (2.46) in a similar way. But this turns out not to be the case. To see this, consider the logical form in (2.48).

(2.48) $[\text{Max}]^1$ $[\text{t}_1$ said that he washed $\text{him}_1]$ he = Max

The reading represented by (2.48) is not an available reading for (2.46), so the LF should be ruled out. To see if it is, we should consider its binding alternative:

(2.49) [Max]^1 [t$_1$ said that he$_1$ washed him$_1$]

Is the transition from (2.48) to (2.49) illegitimate? Only if (2.49) is ruled out by another application of the Interface Rule. To see if it is, we must consider the binding alternative of (2.49):

(2.50) [Max]^1 [t$_1$ said that [he$_1$]^2 [t$_2$ washed him$_2$]]

The fact that [he] binds [him] in (2.50) is in conflict with Condition B. Now we can start to calculate backwards to the original LF: (2.50) is ruled out by Condition B; therefore, the Interface Rule rules out (2.49); and therefore, another application of the Interface Rule rules out (2.48). So the Interface Rule does account for the illegitimacy of (2.48), but in a roundabout way. And even more complex examples can easily be constructed of course.

In conclusion, Reinhart's Interface Rule successfully accounts for Condition B and Strong Crossover effects. Moreover, it allows for strict identity readings in VP ellipsis. But it does *not* account for Dahl's puzzle, its actual formulation is more complex than its underlying intuition, and its workings are (sometimes) undesirably complicated. This concludes my discussion of Reinhart's Interface Rule. Let me now summarize what has been established in this chapter.

2.7 Summary

Reinhart's (1983) proposal was based on two assumptions. First, binding is encoded in the syntax, while coreference is not. Second, speakers try to avoid ambiguity. It follows from these two assumptions that speakers will always prefer to use bound pronouns, which unambiguously encode the intended anaphoric relations, rather than referential pronouns, which could well be misinterpreted. This idea motivated Reinhart's Coreference Rule, which accounts for Condition B effects on coreference (given a syntactic Condition B constraint on binding) and also for cases in which pronouns are exceptionally allowed to corefer with one of their coarguments.

Heim (1998) noted that three aspects of Reinhart's proposal needed further attention. First, the Coreference Rule accounts for Condition B effects on coreference, but not for Condition B effects on cobinding and other kinds of codetermination. For example, cobinding is not ruled out in:

(2.51) Every man said that he voted for him.

Second, it rules out strict readings in cases of VP ellipsis such as:

(2.52) Max called his mother and Bob did too.

Third, it does not make precise what semantic indistinguishability means exactly.

In response to these issues, Heim does two things. First, she observes that semantic indistinguishability may be sensitive to "properties under discussion". This observation has been formalized here in terms of focus values. Second, she proposes a new rule, the Exceptional Codetermination Rule (ECR). This rule accounts for Condition B effects on cobinding and other kinds of codetermination, and it also allows for strict readings in VP ellipsis. But it does not preserve the *explanatory* aspect of Reinhart's approach, that is, it cannot be derived from the general assumptions that binding, but not coreference, is encoded in syntax, and that speakers try to avoid misinterpretation. Furthermore, it does not account for Strong Crossover effects and for Dahl's puzzle.

Fox (1999a) suggested that Strong Crossover effects and Dahl's puzzle may be derived from the general idea that syntactic derivations are subject to certain *economy* principles. The idea is that a complicated structure is not derived if there is a simpler structure with exactly the same interpretation. In particular, a structure may not be derived if there is an alternative structure in which binding relations are more local. This idea motivated Fox's Locality constraint, which accounts both for Strong Crossover effects and for Dahl's puzzle.

To enhance the empirical coverage of Fox's Locality constraint, Büring (2005b) combined it with Reinhart's Coreference Rule. In fact, he proposed a new constraint, Have Local Binding (HLB), which incorporates the effects of both the Coreference Rule and Locality. There are several problems with this proposal, however. First, it is not clear what the intuition is that underlies HLB. Reinhart's Coreference Rule and Fox's Locality constraint are derived from general ideas about the workings of grammar and the behavior of speakers in discourse. But these ideas really seem to be independent of one another. There does not seem to be an more general idea that underlies all of them. Therefore, it is unclear why the Coreference Rule and Locality should be unified, and not just be considered as two separate mechanisms working in tandem. The second problem with Büring's proposal is that it rules out strict readings in VP ellipsis, just as Reinhart's Coreference Rule did. This also has consequences for Dahl's puzzle. To make the right predictions, HLB must be combined with some ad-hoc, still to be worked out theory of VP ellipsis.

In the late 1990's, Reinhart herself departed from her 1983 theory because of the persistent problem with strict readings in VP ellipsis. The alternative view she developed, which took its final shape in (Reinhart, 2006), leaves the first assumption of her 1983 theory intact (binding is encoded in syntactic structure, coreference is not) but replaces the second assumption (speakers avoid ambiguity) with an alternative (hearers minimize interpretive options). In particular, Reinhart assumes that interpretations which are ruled out by constraints on binding are not sneaked in via other anaphoric mechanisms such as coreference and cobinding. This idea underlies the Interface Rule. We have seen that this rule accounts for Condition B effects and Strong Crossover effects, and has no trouble with strict readings in VP ellipsis. However, it does not account for Dahl's puzzle, and there is an undesirable mismatch between (the complexity of) its formulation and its workings on the one hand, and (the simplicity of) the underlying intuition on the other. These remaining issues will be addressed in the next chapter.

Chapter 3

Pushing the Limit

In this chapter I will try to resolve the issues that were left open by the theories discussed in the previous chapter. Section 3.1 provides a reformulation of Reinhart's Interface Rule that perfectly matches (the simplicity of) the underlying intuition. This rule will be referred to as *Rule S*; it literally prohibits *sneaking in* interpretations that are ruled out by restrictions on binding. It will be shown that Rule S, together with Fox's Locality constraint, straightforwardly accounts for all the data discussed in chapter 2.

However, sections 3.2–3.4 show that, upon closer inspection, Locality is both empirically and conceptually problematic. Several alternative principles will be considered. Eventually, the problems raised by Locality will not only be overcome, but the empirical coverage of the overall theory will also be broadened significantly.

3.1 A Simpler Interface Rule

Recall, again, the basic intuition underlying Reinhart's Interface Rule:

> *Interpretations which are ruled out by restrictions on binding*
> *cannot be sneaked in via other anaphoric mechanisms.*

Let me first consider, and slightly revise, Reinhart's formal rendering of *other anaphoric mechanisms*. Reinhart assumes that these mechanisms are all instances of covaluation (cobinding and coreference). But the notion of covaluation does not cover all the relevant instances of anaphoric relatedness. In particular, it does not cover the *indirect* instances of anaphoric relatedness *via third parties*, so to speak. In section 2.6, we saw that it took various recursive applications of the Interface Rule to rule out logical forms like (3.1) below, whereas (3.2) was ruled out in one simple step.

(3.1) [Max]^1 [t_1 said that he washed him$_1$] he = Max

(3.2) [Max washed him] him = Max

Essentially, this is because the indirect way in which [he] and [him] are anaphorically related in (3.1) does not count as covaluation. The complication can easily be avoided by adopting Heim's notion of *codetermination* instead of Reinhart's notion of covaluation. The definition is repeated here for convenience.

3.1. DEFINITION. [Codetermination]
Let C be a context, let LF be a logical form, and let A and B be two NPs in LF. We say that A and B are codetermined in LF/C iff:

- A binds B in LF, or
- A and B corefer in C, or
- There is a third NP which is codetermined with A and B in LF/C.

Next, I propose the following simplified interface rule:

3.2. DEFINITION. [Rule S]
A logical form is illicit if it is semantically indistinguishable from one of its binding alternatives LF′, and LF′ is ruled out by constraints on binding (Condition B).

The notion of a *binding alternative* has to be revised slightly. The only difference between definition 2.7 and definition 3.3 is that in the former A and B are supposed to *corefer*, while in the latter A and B are supposed to be *codetermined*.

3.3. DEFINITION. [Binding Alternatives]
Let C be a context, let LF be a logical form, and let A and B be two noun phrases such that A c-commands B in LF, A and B are codetermined in LF/C, but A does not bind B in LF. Then the structure obtained from LF by:

- Quantifier raising A in case it has not been raised yet;
- Replacing B with a pronoun or trace bound by A.

is called a binding alternative of LF.

Notice that Rule S is in exact correspondence with Reinhart's original intuition. Also, the workings of Rule S are as straightforward as we would like them to be. In particular, (3.1) is now dealt with just as (3.2): codetermination is ruled out in one simple step.

Rule S is weaker than the Interface Rule. In particular, it does not account for Strong Crossover effects, and it allows too *many* (instead of too *few*) readings of the target clause in Dahl's puzzle. To see this, first consider the typical Strong Crossover effect exhibited by example (2.34), repeated here as (3.3):

(3.3) Who did he say we should invite?

a. $[\text{who}]^1$ $[[\text{he}_1]^2$ $[t_2$ said we should invite $t_1]]$

b. $[\text{who}]^1$ $[[\text{he}_1]^2$ $[t_2$ said we should invite $t_2]]$

The Interface Rule correctly rules out (3.3a) because its binding alternative, (3.3b), does not preserve the binding relations that are present in (3.3a). Rule S does not require that binding relations be preserved, so it has no ground on which to rule out (3.3a) and other cases of Strong Crossover.

Next, consider Dahl's puzzle. Recall the source clause:

(3.4) Max said that he called his mother.

Rule S does not rule out *any* of the relevant logical forms of (3.4), because none of their respective binding alternatives violates Condition B. As a consequence, unlike the Interface Rule, Rule S wrongly permits the unavailable reading of the target clause. On the other hand, Rule S correctly permits all the other readings of the target clause, which the Interface Rule fails to do.

The next subsection shows that the empirical shortcomings of Rule S are naturally accounted for if it is combined with Fox's Locality constraint.

3.1.1 Rule S plus Locality

We have already seen, in section 2.4, that Locality accounts for Strong Crossover effects and for Dahl's puzzle. We only need to verify that Rule S and Locality *together* are not too strong. In particular, they should not rule out any of the first three readings in Dahl's puzzle. The relevant logical forms are repeated here:

(2.18) Max said that he called his mother.

a.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his_2 mother$]]$		(2.17a)
b.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his_1 mother$]]$		(2.17a)
c.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his mother$]]$	he=his=Max	(2.17b)
d.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his_2 mother$]]$	he=Max	(2.17b)
e.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his mother$]]$	his=Max	(2.17c)
f.	$[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his_1 mother$]]$	he=Max	(2.17d)

We have seen that Rule S does not rule out any of these logical forms and that Locality rules out (2.18b) and (2.18f). As a result, the three possible readings of the target clause are correctly permitted and the one unavailable reading is correctly ruled out (assuming that VP ellipsis is governed by VP Identity[1]). Thus, Rule S and Locality, together with Condition B, account for all the data we have seen so far and are compatible with the VP Identity condition on VP ellipsis.

[1]Incidentally, the same result would be obtained if we assumed NP Parallelism instead of VP Identity. But we are not *forced* to do so. Thus, the present proposal overcomes the main drawback of Büring's theory.

Finally, let me compare Rule S with Heim's Exceptional Codetermination Rule.

3.1.2 Rule S and Heim's ECR

It turns out that Rule S, in combination with the standard version of Condition B, is empirically equivalent with Heim's ECR together with her own, stronger version of Condition B. The standard version of Condition B is only concerned with *binding*: it rules out logical forms in which a pronoun is bound by one of its coarguments. Rule S, then, extends this restriction to *codetermination*: it rules out logical forms in which a pronoun is codetermined with one of its coarguments, *except* when such logical forms are semantically distinguishable from their binding alternatives. In Heim's system, the division of labour between Condition B and the ECR is just a little bit different. Heim's version of Condition B is directly concerned with *codetermination*: it rules out any logical form in which a pronoun is codetermined with one of its coarguments. The ECR, then, takes care of the exceptions to this rule: it says that a logical form in which a pronoun is codetermined with (but not bound by) one of its coarguments is exceptionally allowed if it is semantically distinguishable from its binding alternative. Thus, both systems make exactly the same predictions. The fact that they were arrived at via different routes is, I think, a positive sign in itself.

There are two reasons to prefer Rule S over the ECR. First, it allows us to stick to the light version of Condition B, which is only concerned with binding. This relieves syntax from a significant burden. Second, Rule S is derived from the general principle that hearers minimize interpretive options. It seems plausible that such a principle is at work in human communication, and Reinhart (2006) argues that its effects are not only exhibited by the interpretation of anaphora but also by the assignment of quantifier scope. The ECR cannot be derived from any such general principle. Thus, Rule S reconciles the empirical adequacy of Heim's ECR with the attractive explanatory outlook underlying Reinhart's Interface Rule.

At this point, all the problems that were left open in chapter 2 seem to be resolved in a satisfactory manner. The resulting theory has three components:

- Condition B: a syntactic constraint, which we may try to derive from more general syntactic mechanisms (cf. Reuland, 2001, 2008), accounts for Condition B effects on binding;

- Locality: a syntactic constraint, which instantiates the general idea that syntactic derivations are subject to *economy* principles (cf. Fox, 1999a), accounts for Strong Crossover effects and Dahl's puzzle;

- Rule S: an interpretive principle, which is derived from the general idea that hearers *minimize interpretive options* (cf. Reinhart, 2006), accounts for Condition B effects on coreference, cobinding, and more intricate kinds of codetermination.

The theory as a whole is compatible with the VP Identity condition on VP ellipsis.

It seems, then, that we may stop here. Unfortunately, however, there are certain problems with Locality that have been overlooked so far. The following sections discuss these problems and will eventually lead to the conclusion that Locality must be replaced by other principles.

3.2 Free Variable Economy

A Problem for Locality. Consider the following sentence:

(3.5) Every man said that he called his mother and that Bill did too.

This sentence has the following two readings:

(3.6) a. Every man x said that x called x's mother and that Bill called Bill's mother too. [sloppy]

b. Every man x said that x called x's mother and that Bill called x's mother too. [strict]

The problem with Locality is that it wrongly prohibits the strict reading in (3.6b). To see this, consider the logical from corresponding to this reading:

(3.7) [Every man]^1 [t$_1$ said [
[that [he$_1$]^2 [t$_2$ called his$_1$ mother]] and
[that [Bill]^2 [t$_2$ called his$_1$ mother]] too]]

This logical form has the following Locality alternative:

(3.8) [Every man]^1 [t$_1$ said [
[that [he$_1$]^2 [t$_2$ called his$_2$ mother]] and
[that [Bill]^2 [t$_2$ called his$_1$ mother]] too]]

The difference between these two logical forms is this: in (3.7), the pronoun [his] in the first conjunct of the embedded clause is bound by [every man]; in (3.8), it is bound by [he] (more locally). The two logical forms are semantically indistinguishable so Locality rules out (3.7) and therefore wrongly predicts that (3.5) does not have the strict reading in (3.6b).

Notice that (3.5) is very similar to Dahl's example. If we strip off the clauses in which ellipsis takes place, we are left with:

(3.9) a. Max said that he called his mother.
b. Every man said that he called his mother.

The only difference is that the subject of (3.9a) is a referential noun phrase, whereas the subject of (3.9b) is a quantifying noun phrase. Locality predicts that non-local binding of [his] is impossible in both sentences. In the case of (3.9a) this is a welcome prediction, as it accounts for Dahl's puzzle. But in the case of (3.9b) it is not, because it wrongly rules out the strict reading of (3.5).

It should be noted that this problem carries over to alternative accounts of Dahl's puzzle such as those of Kehler (1993), Fiengo and May (1994), and Schlenker (2005).

Possible Solution. I will try to overcome this impasse in a way that preserves the general idea that syntactic derivations are subject to economy principles. Locality was derived from this general idea by assuming that the grammar considers one logical form to be *simpler* (more economical) than another if the binding relations it encodes are more local. This particular assumption seems to be problematic, but that does not mean that the general idea must be given up. There may be other criteria for simplicity. Below, I will formulate such a criterion. This criterion is concerned with *free variables*, which are defined as follows:

3.4. Definition. [Free Variables]
The free variables in a constituent X are the indices on pronouns in X whose binder is not contained in X.

Let me give some examples:

(3.10) a. [[Max]^2 [t_2 called his_2 mother]]
b. [[Max]^2 [t_2 called his_1 mother]]
c. [[he_1]^2 [t_2 called his_1 mother]]
d. [[he_1]^2 [t_2 called his_2 mother]]

(3.10a) does not contain any free variables, because the pronoun is bound within the given constituent. (3.10b) does contain a free variable, because the pronoun [his] has a binding index, and is not bound within the given constituent. (3.10c) also contains one free variable. Notice that we are not counting *occurrences* of free variables. The constituent contains two unbound pronouns, but both have the same index, so there is only one free variable. If one of the pronouns is bound, as in (3.10d), the number of free variables does not change, it is still one.

Now, we could think of one constituent as being more *economical* than another if it contains fewer free variables, and more generally, we could think of one logical form LF as being more economical than another logical form LF′ (which is identical to LF modulo binding indices) if some constituent in LF contains fewer free variables than the corresponding constituent in LF′, and all the other constituents in LF contain at most as many free variables as the corresponding constituents in LF′. Given this notion of economy, we could formulate the following principle.

3.5. Definition. [Free Variable Economy]
A logical form is ruled out if it has a more economical, semantically indistinguishable v-alternative.

3.6. Definition. [v-alternatives]
A v-alternative of a logical form LF is a structure which is identical to LF modulo binding indices on pronouns.

Free Variable Economy accounts for Dahl's puzzle without ruling out the strict reading of (3.5). That is, it prohibits non-local binding in (3.9a) but not in (3.9b). To see this, first consider the logical form of (3.9a) that should be ruled out in order to account for Dahl's puzzle:

(3.11) [Max]^1 [t_1 said that [[he]^2 [t_2 called his$_1$ mother]]] he = Max

This logical form has the following v-alternative:

(3.12) [Max]^1 [t_1 said that [[he]^2 [t_2 called his$_2$ mother]]] he = Max

The only difference between these two logical forms is that [his] is bound by [Max] in (3.11) and by [he] in (3.12). The two are semantically indistinguishable, and, crucially, (3.12) is more economical than (3.11). To see this, consider the embedded clause in both logical forms. In (3.11), the embedded clause contains a free variable; in (3.12) it doesn't. This is enough for (3.12) to be considered more economical than (3.11), and thus for Free Variable Economy to account for Dahl's puzzle.

Now consider the logical form of (3.9b), which should *not* be ruled out by Free Variable Economy (otherwise the strict reading of (3.5) cannot be derived).

(3.13) [Every man]^1 [t_1 said that [[he$_1$]^2 [t_2 called his$_1$ mother]]

This logical form has the following v-alternative:

(3.14) [Every man]^1 [t_1 said that [[he$_1$]^2 [t_2 called his$_2$ mother]]

But this v-alternative is *not* more economical. Consider, in particular, the embedded clause. In (3.13), neither [he] nor [his] is bound within the embedded clause, but both carry the same index, so the embedded clause contains *one* free variable. In (3.14), [his] is bound within the embedded clause, but [he] is not, so the clause still contains one free variable. Thus, the embedded clause in (3.14) is *not* more economical than the one in (3.13). It can be shown that no other constituent in (3.14) is more economical than the corresponding constituent in

(3.13) either, and that the same holds for other v-alternatives of (3.13). Thus, Free Variable Economy does not rule out (3.13) and correctly derives the strict reading of (3.5).

It seems, then, that we have resolved the problem raised at the beginning of this section. Replacing Locality by Free Variable Economy, however, raises a new issue. The responsibility of Locality was more than just dealing with Dahl's puzzle. It also dealt with Strong Crossover effects. The next section shows that Free Variable Economy does not and suggests that part of Locality should therefore be preserved.

3.3 Trace Locality

Free Variable Economy does not account for Strong Crossover effects. To see this, consider example (2.34), repeated here as (3.15):

(3.15) Who did he say we should invite?

The logical form that must be ruled out is:

(3.16) $[\text{who}]^1$ $[[\text{he}_1]^2$ $[\text{t}_2$ said we should invite $\text{t}_1]]$

In order for (3.16) to be ruled out by Free Variable Economy, it must have a more economical v-alternative. The v-alternatives of (3.16) are all identical to (3.16), apart from the binding index on the pronoun. But this means that the v-alternatives all contain a free variable, while (3.16) does not. Thus, (3.16) does not have a more economical v-alternative, and therefore it is not ruled out by Free Variable Economy.

Maybe, then, part of Locality should be preserved. Notice that Locality can be decomposed into two sub-constraints: one about traces and one about pronouns. Let us call these two sub-constraints Trace Locality and Pronoun Locality, respectively. Roughly speaking, Trace Locality says that traces must be bound locally, and Pronoun Locality says that pronouns must be bound locally.

Next, observe that Trace Locality accounts for Strong Crossover effects, but has nothing to say about Dahl's puzzle. On the other hand, Pronoun Locality is concerned with Dahl's puzzle, but has nothing to say about Strong Crossover effects. So there is a strict division of labor: Trace Locality accounts for Strong Crossover effects, while Pronoun Locality is concerned with Dahl's puzzle.

We have seen above that Pronoun Locality is problematic, because it prohibits the strict reading of (3.5). But that does not mean that Trace Locality should be thrown overboard as well. We could simply adopt Free Variable Economy instead of Pronoun Locality to account for Dahl's puzzle, and maintain Trace Locality to account for Strong Crossover effects.

The definition of Trace Locality is spelled out below. Notice that this definition only differs from the definition of Locality (definition 2.12 on page 56) in that it is exclusively concerned with traces.

3.7. Definition. [Trace Locality] A logical form is ruled out if it is semantically indistinguishable from one of its trace locality alternatives.

3.8. Definition. [Trace Locality Alternatives]
Let LF be a logical form, let T be a trace in LF, and let A and B be two other nodes in LF such that A c-commands B and B c-commands T, and such that A, but not B, binds T. Then the structure obtained from LF by:

- Quantifier raising B in case it has not been raised yet;
- Adjusting T's binding index so that it's bound by B instead of A.

is a trace locality alternative of LF.

Trace Locality and Free Variable Economy account for Strong Crossover effects and Dahl's puzzle. However, it will be pointed out below that Trace Locality inherits a conceptual problem from Fox's original Locality condition. Eventually, an alternative principle will be adopted which does not only overcome the problem that Trace Locality faces, but also has a much wider empirical coverage.

3.4 Movement Economy

An objection that may be raised against Trace Locality (and against Fox's original Locality constraint as well) is that it compares logical forms to structures which may not be proper logical forms in their own right. To see what is at stake, consider again the basic strong crossover case in (2.34), repeated here as (3.17).

(3.17) Who did he say we should invite?
a. $[\text{who}]^1$ $[[\text{he}_1]^2$ $[t_2$ said we should invite $t_1]]$
b. $[\text{who}]^1$ $[[\text{he}_1]^2$ $[t_2$ said we should invite $t_2]]$

(Trace) Locality compares the logical form in (3.17a) with the alternative in (3.17b). But (3.17b) is not an independently derivable logical form. The idea that motivated economy principles like (Trace) Locality was that some syntactic structures may be ruled out because they have simpler alternatives. But why should (3.17b) be considered as an alternative of (3.17a) if it isn't a proper syntactic structure in its own right?

To overcome this problem, I will consider an alternative account of crossover effects, due to Eddy Ruys (1994), which is very close in spirit to Trace Locality, but does not refer to underivable structures. Ruys' idea is best explained by means of an example. Consider again the case of (3.17). The logical form that should be ruled out is:

(3.18) [who]^1 [he$_1$ said we should invite t$_1$]

Ruys suggests that this logical form should be compared with the following alternative:

(3.19) [who]^1 [t$_1$ said we should invite he$_1$]

Notice that (3.19) only differs from (3.18) in that the pronoun and the trace have swapped places. Moreover, note that (3.19) is a logical form in its own right, and that (3.18) and (3.19) are semantically indistinguishable.

Ruys suggests that (3.19) may be considered more economical than (3.18) because the movement operation involved in its derivation is more local. In (3.19), [who] has moved to its position at LF from the matrix subject position. In (3.18), [who] has moved to its position at LF all the way from the embedded object position. On this ground, (3.19) is more economical than (3.18) in Ruys' system, and (3.18) is ruled out. This idea can be formalized as follows.

3.9. Definition. [Movement Economy]
A pair $\langle$PF,LF$\rangle$ is ungrammatical if there is a grammatical pair $\langle$PF',LF$'\rangle$ such that (i) LF$'$ is an m-alternative of LF, (ii) LF$'$ and LF are semantically indistinguishable, and (iii) the movement involved in the derivation of LF$'$ is more local than the movement involved in the derivation of LF.

3.10. Definition. [m-alternatives]
Let LF be a logical form, and let T and P be a trace and a pronoun, respectively, which are cobound in LF. Then the logical form obtained from LF by swapping T and P is an m-alternative of LF.

Recall that Strong Crossover effects are considered to be a special case of crossover effects more generally (see section 2.1.2). We saw that it is not entirely clear when crossover effects obtain. In particular, the "textbook" generalization, which says that pronouns can only be cobound with traces that c-command them (generalization 2.3 on page 47), is plagued by many counterexamples.

Now, Ruys claims that Movement Economy does not only account for Strong Crossover cases, but also for all the other cases that are supposed to be captured by generalization 2.3, *and* for its counterexamples. To support this claim, let me go through some examples from section 2.1.2. First consider example (2.6d), repeated here as (3.20a), which was one of the examples correctly captured by generalization 2.3, even though it does not exhibit a Strong Crossover effect.

(3.20) a. Who does his mother love?
b. [who]^1 [he$_1$'s mother loves t$_1$]

(3.21) a. Whose mother loves him?
b. [who]^1 [t$_1$'s mother loves him$_1$]

Movement Economy establishes the ungrammaticality of the ⟨PF,LF⟩ pair in (3.20) by comparing it to the pair in (3.21). (3.21b) can be obtained from (3.20b) by swapping the trace and the pronoun, which are cobound in (3.20b). Thus, (3.21b) is an m-alternative of (3.20b). Moreover, (3.20b) and (3.21b) are semantically indistinguishable. Finally, (3.21b) is more economical than (3.20b) because the movement operation involved in its derivation is more local. Thus, the pair in (3.20) is correctly ruled out by Movement Economy.

Next, let us consider some of the counterexamples to generalization 2.3. Two types of counterexamples will be discussed here. Many more can be found in Ruys' paper. First consider the class of counterexamples discovered by Higginbotham (1980). One representative of this class is the pair in (3.21). Generalization 2.3 wrongly predicts that the logical form in (3.21b) should be illicit, because the pronoun in it is cobound with a non-c-commanding trace. Movement Economy does not make this prediction, because (3.21b) does not have a more economical m-alternative. Another case discussed by Higginbotham is:

(3.22) a. Every senator's portrait was on his desk.
b. [every senator]^1 [t_1 's portrait was on his$_1$ desk]

Again, generalization 2.3 wrongly predicts that the logical form in (3.22b) should be illicit. Movement Economy does not make this prediction, because (3.22b) does not have a more economical m-alternative.

(3.23) represents another type of counterexample considered in section 2.1.2 (originally discussed by Reinhart (1983) with attribution to Ross).

(3.23) a. That people hate him disturbs every president.
b. [every president]^1 [[that people hate him$_1$] disturbs t_1]

Generalization 2.3 wrongly predicts that the logical form in (3.23b) is out, because the pronoun in it is cobound with a trace that doesn't c-command it. Movement Economy doesn't make this prediction, because (3.23b) does not have a grammatical, more economical m-alternative. In particular, the m-alternative in (3.24b) is ungrammatical, because it violates constraints on movement.

(3.24) a. That people hate every president disturbs him.
b. [every president]^1 [[that people hate t_1] disturbs him$_1$]

Thus, Movement Economy does not only avoid the problem that Trace Locality runs into. It also has much wider empirical coverage: it does not only account for Strong Crossover effects, but for crossover effects in general.

Finally, I would like to point out that Movement Economy could be seen as an economy principle, as Ruys did, but it can also be taken to follow from a general assumption about speakers' behavior in discourse, similar to the one underlying Reinhart's (1983) theory of Condition B effects on coreference. The general assumption is that a speaker always tries to make the hearer's life as easy as possible. This is in the speaker's own interest, because the harder it is to interpret his utterances, the more likely it is that he will be misunderstood at some point. One way to make the hearer's life easier is to use syntactic structures which involve local movement rather than structures which involve long-distance movement. If this is indeed a principle that guides the speaker's behavior, then a hearer will never assign an interpretation M to a structure involving long-distance movement, if M could also be expressed using a structure involving local movement. After all, if the speaker had wished to convey M he would have used the latter structure instead of the former. This, then, is how the effects of Movement Economy could come about.

I will leave open at this point whether Movement Economy should be thought of as a reflex of general assumptions about speakers' behavior in discourse, or whether it should be seen as a grammatical economy principle. In any case, it is important to know that both options are, in principle, open.

3.5 Summary

In this chapter we set out to resolve the issues that were left open by the theories discussed in chapter 2. Section 3.1 introduced Rule S, a much simplified version of Reinhart's Interface Rule. Rule S is a straightforward implementation of the intuition that interpretations which are ruled out by constraints on binding cannot be sneaked in via other anaphoric mechanisms. It was shown that Rule S, together with the standard version of Condition B, is empirically equivalent with Heim's Exceptional Codetermination Rule combined with Heim's stronger version of Condition B. Rule S has two advantages over the ECR. First, it can be derived from the general principle that hearers minimize interpretive options, and second, it presupposes the standard, "light" version of Condition B (which is concerned with binding only), and not Heim's stronger version (which is concerned with codetermination). Finally, it was shown that Rule S, together with Locality, accounts for all the data discussed in chapter 2.

Section 3.2 presented additional data, which was problematic for Locality. Free Variable Economy was proposed as an alternative account of Dahl's puzzle. But section 3.3 showed that Free Variable Economy does not account for Strong Crossover effects. Thus it was suggested that a restricted version of Locality—Trace Locality—be preserved. In section 3.4, however, it was pointed out that Trace Locality inherits a conceptual problem from Locality: it compares logical forms to structures which may not be proper logical forms in their own right. Eventually, an alternative account of crossover effects, due to Ruys (1994), was considered. This account, which we called Movement Economy, avoids the problem that Trace Locality was facing, and also has a much broader empirical coverage: it does not only account for Strong Crossover effects, but for crossover effects in general. Finally, it was observed that Movement Economy could be seen as an economy principle, as Ruys did, but it could also be taken to follow from the general assumption that speakers try to minimize the processing load imposed on their audience. This means, in particular, that they avoid structures involving long-distance movement whenever possible.

Thus, we have arrived at a theory consisting of the following four components:

- Condition B: a syntactic constraint, which we may try to derive from more general syntactic mechanisms (cf. Reuland, 2001, 2008), accounts for Condition B effects on binding;
- Rule S: an interpretive principle, which is derived from the general idea that hearers *minimize interpretive options* (cf. Reinhart, 2006), accounts for Condition B effects on coreference, cobinding, and more intricate kinds of codetermination.
- Free Variable Economy: a syntactic constraint, which instantiates the general idea that syntactic derivations are subject to *economy* principles (cf. Fox, 1999a), accounts for Dahl's puzzle;
- Movement Economy: a constraint which can either be thought of as a syntactic economy principle (cf. Ruys, 1994) or as a consequence of the idea that speakers try to minimize the processing load imposed on their hearer, accounts for crossover effects.

The theory as a whole is compatible with the assumption that VP ellipsis is governed by Focus Match and a Semantic Identity condition.

Part II

Anaphora Resolved

Chapter 4

Issues

In the first part of this dissertation I examined several existing accounts of pronominal anaphora. A number of problems were identified, but these problems were all resolved without departing in any substantial way from the fundamental ideas underlying the original theories. For example, the new interface rule I proposed, Rule S, preserved Reinhart's idea that hearers minimize interpretive options, and the new economy constraint, Free Variable Economy, preserved Fox's idea that syntactic derivations are subject to economy principles. The distinction between bound and referential pronouns, and the assumption that binding is encoded in the syntax, were also preserved.

In this second part, I will consider certain objections to some of these ideas and advance an alternative theory which is based on the idea that pronouns and elided VPs always retrieve their meaning from the context in which they are used. In a slogan: *anaphora are resolved.*

4.1 Issues of Unification

The framework adopted in Part I was based on the following two assumptions about the nature of pronominal anaphora and VP ellipsis:

1. Pronouns are either bound or referential.
2. VP ellipsis consists in deleting a VP constituent at PF.

These assumptions immediately raise the following two issues:

1. Can't we treat all pronouns alike?
2. Can't we treat pronouns and VP ellipsis alike?

I will discuss each of these issues in somewhat more detail.

Bound and Referential Pronouns. In section 1.1 it was argued that not all pronouns should be treated alike, and in particular, that a distinction should be made between bound and referential pronouns. The primary reason for doing so was that pronouns seem to be interpreted as bound variables in sentences like (1.1), repeated here as (4.1), whereas they seem to function as referential expressions in sentences like (1.2), repeated here as (4.2).

(4.1) Every man thinks he will win.

(4.2) John is in good shape. I think he will win.

A secondary reason to adopt the distinction between bound and referential pronouns was that it yields an attractive explanation of the ambiguities exhibited by sentences like (1.3)-(1.5), repeated here as (4.3)-(4.5).

(4.3) MAX called his mother.

(4.4) Only MAX called his mother.

(4.5) Max called his mother and Bob did too.

Thus, there seemed to be good reasons to assume that English pronouns are in fact systematically ambiguous between a bound variable interpretation and a referential interpretation. However, if such a systematic ambiguity were real, then we would expect to find other languages in which certain pronouns are either always interpreted as bound variables or always interpreted referentially (no matter in which context they are used).

To make this point clear, let me briefly consider another case of apparent lexical ambiguity in English. The sentences in (4.6) and (4.7) exemplify two seemingly distinct usages of the verb *to know*.

(4.6) I know John McEnroe.

(4.7) I know that John McEnroe won Wimbledon in 1984.

Intuitively, in (4.6) *to know* roughly means *to be acquainted with*, whereas in (4.7) it roughly means *to have conclusive evidence for*. Evidence for the hypothesis that the verb *to know* is lexically ambiguous comes from the fact that languages other than English indeed have distinct lexical items for *being acquainted with* and *having conclusive evidence for*. In Dutch, for example, the first is translated as *kennen* and the second as *weten*. Similarly, German has *kennen* and *wissen*, Spanish has *conocer* and *saber*, and Italian has *conoscere* and *sapere*. In each of these languages, (4.6) would be translated using the first verb, and (4.7) would be translated using the second verb.

In the case of pronouns, to the best of my knowledge, no languages have been found in which there are specialized lexical items either for bound or for referential pronouns. This makes one think again. Couldn't there be a unified analysis of pronouns which would be compatible with their seemingly distinctive use in (4.1) and (4.2), and which would moreover still provide a natural explanation for the ambiguities in (4.3), (4.4), and (4.5)?

Pronouns and VP ellipsis. The idea that VP ellipsis and pronominal anaphora are fundamentally different dates back to an influential paper by Hankamer and Sag (1976). These authors argued that there are two types of anaphora: *deep* anaphora and *surface* anaphora. Pronouns were classified as deep anaphora, whereas VP ellipsis was classified as surface anaphora. The main difference between the two was taken to be that deep anaphora (pronouns) can be deictic—that is, their meaning may be retrieved from the non-linguistic context—while surface anaphora (VP ellipsis) always require a linguistic antecedent. Soon after the publication of Hankamer and Sag's paper, Schachter (1977) already presented several counterexamples to the claim that VP ellipsis always requires a linguistic antecedent:

(4.8) [John meets Mary in a bar. He points at a chair near her and says:] May I?

(4.9) [She answers:] Please do.

(4.10) [He offers her a drink, but she says:] I really shouldn't.

(4.11) [Then he invites her to dance:] Shall we?

Webber (1978) also discussed several cases of VP ellipsis which do not involve direct linguistic antecedents. One famous example is the following:

(4.12) Irv and Mary wanted to dance together, but Mary couldn't, because her husband was there.

VP ellipsis is not really deictic here, as in Schachter's examples, but the antecedent must be *inferred* from the linguistic context (*Irv and Mary wanted to dance together* entails *Mary wanted to dance with Irv*). Many authors have been convinced by Schachter's and Webber's examples that VP ellipsis does not require a direct linguistic antecedent. In fact, Sag (2006) himself provides many more cases of inference-based VP ellipsis, and abandons the viewpoint expressed in (Hankamer and Sag, 1976) and (Sag, 1976).

Still, there are also authors who have tried to keep the distinction between deep and surface anaphora alive. In particular, Hankamer (1978) argued that Schachter's counterexamples were all highly conventionalized, idiosyncratic expressions which do not involve ellipsis at all. Pullum (2000) endorses this view. But I am not convinced. I do think Schachter's counterexamples are legitimate. It is true that deictic pronouns are much more common than deictic VP ellipsis, but there are good reasons for why this is so. First, pronouns usually carry some (gender/number/person) information, which makes it easier to determine their intended referent. Elided VPs do not convey such information. Second, there are many more situations in which a particular object is the single most salient (female/singular/third person) object than there are situations in which a certain property or activity is the single most salient property or activity. Thus, the intended meaning of elided VPs is generally much harder to recover from the non-linguistic context than the intended meaning of pronouns.

In sum, I do not think that there are any good reasons to assume that pronouns and VP ellipsis are fundamentally distinct. In fact, there is significant evidence to the contrary. In particular, as discussed at length by Wasow (1972) and also emphasized by Williams (1977), pronouns and VP ellipsis have several striking properties in common. I will briefly illustrate three such properties here. First, neither pronouns nor elided VPs can c-command their antecedents. This is illustrated by the following examples from (Wasow, 1972, p.81):

(4.13) a. *John* dropped out after *he* tried LSD.
b. After *John* tried LSD, *he* dropped out.
c. After *he* tried LSD, *John* dropped out .
d. #*He* dropped out after *John* tried LSD.

(4.14) a. John *tried LSD* after Bill *did*.
b. After Bill *tried LSD*, John *did*.
c. After Bill *did*, John *tried LSD*.
d. #John *did* after Bill *tried LSD*.

Second, neither pronominal anaphora nor VP ellipsis are subject to Ross' (1967) island constraints. Example (4.15) shows this for the Complex Noun Phrase Constraint, example (4.16) for the Coordinate Structure Constraint, and example (4.17) for the Sentential Subject Constraint. All the examples are from (Wasow, 1972, pp.94-95). The relevant islands are indicated by square brackets.

(4.15) a. John believes [the prediction that he will win].
b. John didn't take LSD, but Bill believed [the claim that he did].

(4.16) a. Nixon seems to believe that [he and Agnew] had to cheat to win.
b. The public realizes that Nixon lied, although Mitchell claims that he [never would and can be trusted].
c. The public realizes that Nixon lied, although Mitchell claims that he [is an honest man and never would].

(4.17) a. Ford beliefs that [for him to resign] would be a disaster.
b. Although Ford didn't resign, [that many people wanted him to] is encouraging.

Finally, as observed by Williams (1977, pp.101–102), both pronominal anaphora and VP ellipsis may operate across sentence boundaries and even across speaker boundaries. The following examples illustrate this:

(4.18) A: John called.
B: What did he say?

(4.19) A: Who called?
B: John did.

Pronominal anaphora and VP ellipsis differ in these respects from other operations, such as Gapping, Comparative Deletion, Object Deletion, and Relative Deletion (for illustrations of this fact, see Wasow, 1972; Williams, 1977).

Thus, I think we should seek a unified treatment of pronouns and VP ellipsis. This means that the basic assumptions underlying the framework adopted in chapter 1 should be revised.

4.2 Issues of Stipulation

In the previous section, we considered the assumptions adopted in Part I as to what pronominal anaphora and VP ellipsis are. Now, let us consider the proposed account of how the interpretation of pronouns and VP ellipsis is *constrained*. The conclusion we reached in Part I was that the interpretation of pronominal anaphora is subject to the following four constraints:

1. Condition B
2. Rule S
3. Free Variable Economy
4. Movement Economy

The interpretation of VP ellipsis was assumed to be governed by Focus Match and a Semantic Identity constraint.

We saw that Rule S and Movement Economy can be derived from general ideas about how people behave in communication. In particular, Rule S is derived from the idea that hearers minimize interpretive options, and Movement Economy can be derived from the idea that speakers minimize the risk of being misinterpreted. It was also pointed out that Condition B may be derived from general syntactic principles, that Focus Match may be taken to follow from a general theory of information structure encoding, and finally, that Free Variable Economy may be derived from the general idea that syntactic derivations are subject to economy principles. The Semantic Identity constraint on VP ellipsis, however, was not shown to follow from more general ideas. This is rather dissatisfying: if VP ellipsis is indeed subject to a Semantic Identity constraint we would like to have an explanation for why this is so. Such an explanation was not given in Part I.

4.3 Pronouns Revisited

In Part I we assumed that pronouns can be bound or referential. A bound pronoun is translated as a variable; a referential pronoun takes on the meaning of its antecedent. Section 4.1 raised the issue of whether the distinction between bound and referential pronouns is really necessary. In this section, we will see that there are several usages of pronouns that cannot be classified as either bound or referential (in the above sense). The conclusion will be that the distinction between bound and referential pronouns drawn in Part I is not only conceptually undesirable, but also empirically too narrow to cover all usages of pronouns.

Deixis. Pronouns can be used deictically. That is, they can be used to refer to an object in the non-linguistic context:

(4.20) [Pointing at John] He won.

Strictly speaking, this usage does not count as referential in the terminology of Part I, because the pronoun does not take on the meaning of its antecedent (there is no antecedent). Rather, its meaning is retrieved from the context in a more liberal way.

Inferred Antecedents. Deictic pronouns are not the only pronouns which do not have an explicit linguistic antecedent. Pronouns may also have *inferred* antecedents:[1]

(4.21) If I get pregnant, I'll definitely keep it. (overheard in conversation)

(4.22) There was not a man, a woman or child within sight; only a small fishing boat, standing out to sea some distance away. Harriet waved wildly in its direction, but they either didn't see her or supposed that she was merely doing some kind of reducing exercises.
(Gundel *et al.*, 1993, quoted from a novel by Doroty Sayers)

(4.23) What's that shadow creeping up the wall? Could it be a burglar?
(Geurts, 2008)

(4.24) John bled so much it soaked through his bandage and stained his shirt.
(Tic Douloureux, 1971)

(4.25) Maxine was kidnapped but they didn't hurt her. (Bolinger, 1977)

Again, no explicit antecedent is involved in these cases.

[1] Examples (4.24) and (4.25) are taken from (Geurts, 2008), who provides several more examples of the same kind.

Verbal Antecedents. The following example shows that the meaning of a pronoun is sometimes retrieved from a verbal antecedent.

(4.26) We built the house ourselves, but it wasn't easy.

Indefinite Antecedents. Pronouns can have indefinite antecedents:

(4.27) A man came in. He sat down.

Again, the pronoun does not simply take on the meaning of its antecedent in this case: (4.27) differs in meaning from (4.28).

(4.28) A man came in. A man sat down.

Donkey Pronouns. A pronoun can also have an indefinite antecedent which does not introduce any specific individual. The following classical examples are from Geach (1962):

(4.29) Every farmer who owns a donkey beats it.

(4.30) If a farmer owns a donkey, he beats it.

Notice that there is no single man and no single donkey to which *he* and *it*, respectively, refer. Pronouns of this kind are called *donkey* pronouns, *E-type* pronouns (Evans, 1980) or *D-type* pronouns (Elbourne, 2005b).

Pronouns of Laziness. The following example from Karttunen (1969) exemplifies another usage of pronouns which is difficult to classify as either bound or coreferential:

(4.31) The man who gave his paycheck to his wife is smarter than the man who gave it to his mistress.

The pronoun *it* refers to the second man's paycheck, while its antecedent refers to the first man's paycheck. Such pronouns are called pronouns of laziness, or paycheck pronouns.

Conclusion. Many pronouns cannot be classified straightforwardly as either bound or coreferential. In particular, pronouns do not necessarily take on the meaning of an explicit linguistic antecedent. Rather, the meaning of pronouns is sometimes retrieved from the context in more liberal ways.

4.4 VP ellipsis Revisited

Similar observations can be made concerning VP ellipsis. In this section, we will see various examples which seem to disprove the idea that VP ellipsis consists in deletion under VP Identity.

Deixis. As remarked in section 4.1, VP ellipsis may be deictic. See examples (4.8)–(4.11).

Inferred Antecedents. As also remarked in section 4.1, VP ellipsis may involve inferred antecedents. See example (4.12) (from Webber, 1978) and the ones below (from Sag, 2006).

(4.32) A: I just need the impetus of someone to collaborate with.
B: Well, I'd love to.

(4.33) I'm gonna send them an email saying that Ling 1 is something they could take. I don't think that many of them will, though. (DB, Sept. 26, 2005)

(4.34) They can't come here to Akron or to any other place in America and talk to you about all the jobs that they created, because they haven't. (John Kerry, Sept. 4, 2004)

(4.35) Scott: They need reassurance that I can't give them.
Harper: Yes you can. (Boston Public, Dec. 2, 2002)

(4.36) There will be a "total evacuation of the city. We have to. The city will not be functional for two or three months," Nagin said.
(Guardian Unlimited, Aug. 31, 2005)

Nominal Antecedents. Just like pronouns can have verbal antecedents—see example (4.26) above—elided VPs can have nominal antecedents. The following three examples, all taken from Hardt (1993, p.35), show that the meaning of elided VPs can be retrieved from nouns, nominalized verbs, and gerundive nominals, respectively.

(4.37) People say that Harry is an excessive drinker at social gatherings.
Which is strange, because he never does at my parties.

(4.38) We should suggest to her that she officially appoint us as a committee and invite faculty participation. They won't, of course,...
(example attributed to Bonnie Webber)

(4.39) Seeing them did not greatly surprise Enid either, though she would wish later she hadn't. (*You Must Remember This*, Joyce Carol Oates, p.287)

Cascaded VP ellipsis. In section 2.5 we discussed the following case of cascaded VP ellipsis:

(4.40) Bob called his mother, and Max did too. But Tom didn't.

It was observed that (4.40) does not have the following mixed reading:

(4.41) Bob called Bob's mother, and
Max called Max's mother.
But Tom didn't call Max's mother.

That is, the pronoun cannot be interpreted sloppily in one elided VP and strictly in the other, and this is correctly predicted by VP Identity.

However, several authors, including Dahl (1973), Dalrymple *et al.* (1991), Fiengo and May (1994), Fox (1999a), and Schlenker (2005), have pointed out that mixed readings do obtain in slightly different cases of cascaded ellipsis, and have used this observation as an important argument against VP Identity. Consider the following example from Fox (1999a, p.117) (see footnote 12 on page 62):

(4.42) Smithers thinks that his job sucks, and Homer does, too.
But Marge doesn't.

Fox claims that (4.42) does have a mixed reading, at least for people who recognize that it is about a popular American sitcom, *The Simpsons*, in which Marge is Homer's wife and doesn't have a job of her own. The reading in question can be paraphrased as follows:

(4.43) Smithers thinks that Smithers' job sucks, and
Homer thinks that Homer's job sucks.
But Marge doesn't think that Homer's job sucks.

Here is another case of cascaded ellipsis which has been claimed to have a mixed reading (Sem, 1994, p.22):

(4.44) John didn't wash his car, but Bill did, even though Harry already had.

VP Identity cannot account for these mixed readings.

Unexpected Sloppy Readings. The idea that a pronoun can be interpreted sloppily in an elided VP if and only if it is interpreted as a bound variable in the antecedent VP cannot be quite right. Sometimes a pronoun can be interpreted sloppily in an elided VP, even though it cannot possibly be interpreted as a bound variable in the antecedent VP. The following examples are from Wescoat (1989) and Hardt (1993), respectively.

(4.45) The police officer who arrested John insulted him, and the one who arrested Bill did, too.

(4.46) If Harry has trouble at school, I will help him. But if John has trouble at school, I won't.

The pronouns in the antecedent VPs cannot be bound by *John* and *Harry*, respectively, because *John* and *Harry* cannot be raised out of the syntactic islands in which they are contained (the relative clause in (4.45) and the *if*-clause in (4.46)). This is confirmed by the fact that the pronouns in (4.47) and (4.48) cannot be interpreted as anaphorically related to *every murderer* and *every student*, respectively:

(4.47) The police officer who arrested every murderer insulted him.

(4.48) If every student has trouble at school, I will help him.

Thus, it is unexpected that the pronouns in the elided VPs in (4.45) and (4.46) can be interpreted sloppily. Notice that the source clauses in (4.45) and (4.46) are structurally analogous to Geach's donkey examples (4.29)-(4.30).

Conclusion. The above examples all seem to refute the idea that VP ellipsis consists in PF deletion under VP Identity. As in the case of pronouns, it is particularly clear that the meaning (let alone the form) of an elided VP does not necessarily correspond to the meaning of an explicit linguistic antecedent. Rather, it seems that the meaning of elided VPs is sometimes retrieved from the context in more liberal ways.

4.5 Economy Principles Revisited

Next, let me consider the hypothesis that the derivation of syntactic structures is subject to economy principles. As support for this idea, Fox (1999a) does not only consider Dahl's puzzle, in which VP ellipsis interacts with binding, but also another famous puzzle, due to Sag (1976), Williams (1977), and Hirschbühler (1982), in which VP ellipsis interacts with scope. Fox argues that these puzzles can be explained by the idea that non-local binding and non-local scope are only allowed if the resulting interpretation differs from the local binding/scope interpretation. Let me consider the phenomena in question in some detail to see whether this could be the right kind of explanation.[2]

Binding Economy. Recall Dahl's puzzle:

[2] In section 3.2, I argued that, in the case of binding, Fox's Locality principle is problematic, and suggested that it should be replaced with Free Variable Economy. But Free Variable Economy is still an economy principle. Here, I will have to conclude that Dahl's puzzle should not be explained in terms of syntactic economy principles at all.

(4.49) Max said that he called his mother, and Bob did too.

a. ... Bob too said that Bob called Bob's mother. [sloppy-sloppy]
b. ... Bob too said that Max called Max's mother. [strict-strict]
c. ... Bob too said that Bob called Max's mother. [sloppy-strict]
d. #... Bob too said that Max called Bob's mother. [strict-sloppy]

The fact that the strict-sloppy reading in (4.49d) is not a possible reading of (4.49) can be explained by means of an economy principle (either Locality or Free Variable Economy). The problem, however, is that there are contexts in which (4.49d) *is* a possible reading of (4.49). Hardt (1993, page 119) discusses such a context: Max is suspected of murdering Bob's mother. Bob has claimed that Max was visiting Bob's mother at the time of the murder. But Max has presented as his alibi that he was at home with his own mother during the night in question. When the district attorney asks where Max was, someone replies:

(4.50) Well, Max says he was visiting his mother, but Bob does too.

The preferred interpretation of the target clause in (4.50) is that Bob said that Max was visiting Bob's mother. But this reading corresponds exactly with the strict-sloppy reading in Dahl's puzzle.

Another relevant context is discussed by Reuland (2008): John used to be a fanatic gambler. His brother Bill never really liked gambling himself, but he did entrust John with his capital. When John went to Las Vegas to play big, things did not turn out so well, but of course now:

(4.51) John does not dare to admit that he lost his fortune in Vegas, and Bill doesn't either.

A natural interpretation of the target clause in (4.51) is that Bill does not dare to admit that John lost Bill's fortune in Vegas. Again, this reading corresponds exactly with the strict-sloppy reading in Dahl's puzzle.

I take these examples to show that Dahl's puzzle should not be explained in terms of grammatical economy principles, but rather in terms of interpretative preferences, which can be overruled by world knowledge and contextual information.

Scope Economy. Fox argues that economy principles also solve a long-standing puzzle concerning the interaction of scope and VP-ellipsis. The puzzle is that the scope ambiguity in (4.52a) is not present in (4.52b) (Sag, 1976; Williams, 1977) but reappears in (4.52c) (Hirschbühler, 1982). Let us call this the scope puzzle.

(4.52)

a. A boy admires every teacher. $(\exists > \forall)(\forall > \exists)$
b. A boy admires every teacher. Mary does, too. $(\exists > \forall)\cancel{(\forall > \exists)}$
c. A boy admires every teacher. A girl does, too. $(\exists > \forall)(\forall > \exists)$

Fox's economy constraint on scope is formulated as follows (notice the similarity with Binding Locality):

4.1. Definition. [Scope Locality]
A logical form is illegitimate if it is semantically indistinguishable from one of its local scope alternatives.

4.2. Definition. [Local Scope Alternatives]
Let LF be a logical form, and let Q be a scope-taking element which takes non-local scope in LF. Let LF′ be just like LF, but with Q taking local scope instead of non-local scope. Then LF′ is a local scope alternative of LF.

Fox assumes that VP ellipsis is subject to a *Parallelism* constraint, which says that the relative scope of scope-taking elements in the target clause must be parallel to those in the source clause.[3] Then, he argues that Scope Locality and Parallelism together account for the scope puzzle. The argument goes as follows. In the target clause of (4.52b), inverse scope is ruled out by Scope Locality. Therefore, by Parallelism, inverse scope is also ruled out in the source clause of (4.52b). In the target clause of (4.52c) on the other hand, Scope Locality does *not* rule out inverse scope, because the LF with inverse scope is semantically distinguishable from the LF with local scope. Therefore, by Parallelism, inverse scope is also allowed in the source clause of (4.52c). This resolves the puzzle.

However, just like Binding Locality, Scope Locality is sometimes too strong. To see this, consider the following example (compare with (4.52b)):

(4.53) In the morning, a nurse checked every patient. $(\exists > \forall)(\forall > \exists)$
In the afternoon, Doctor Jones did.

Johnson and Lappin (1997, p.311) (cited by Fox, 1999a, p.35) already discussed very similar examples:

(4.54) At least one natural number other than one divides into every prime number, and one does too. $(\exists > \forall)(\forall > \exists)$

(4.55) At least two cabinet members bear responsibility for each government department, and the Prime Minister does too. $(\exists > \forall)(\forall > \exists)$

What these examples show is that the scope puzzle, just like Dahl's puzzle, should not be explained in terms of grammatical economy principles, but rather in terms of interpretive preferences, which may be overruled by world knowledge and contextual information.

[3] This Parallelism constraint on VP ellipsis follows from Focus Match (see page 40), which in turn is supposed to follow from general theories of information structure encoding (cf. Rooth, 1992a; Tancredi, 1992; Schwarzschild, 1999). Alternatively, the Parallelism constraint on VP ellipsis can be derived from general theories of discourse coherence establishment (cf. Prüst *et al.*, 1994; Asher *et al.*, 2001).

4.6 A Closer Look at Dahl's Puzzle

It is true that, if Dahl's sentence is considered in a neutral context, the fourth reading (strict-sloppy) is strongly dispreferred. But there is something else about this sentence which many informants find even more striking, namely that the first two readings (sloppy-sloppy and strict-strict) are distinctly preferred over the second two (sloppy-strict and strict-sloppy). Let us call the first two readings *across-the-board* readings and the second two *mixed* readings. A complete analysis of Dahl's puzzle, then, should explain two things. First, it should explain why, in neutral contexts, across-the-board readings are preferred over mixed readings. And second, it should explain why the sloppy-strict reading is easier to accommodate than the strict-sloppy reading. The theory developed in Part I is far from complete in this sense. In particular, it has nothing to say about the preference for across-the-board readings over mixed readings.

4.7 Condition B Data Revisited

It was assumed in Part I, following Reinhart (1983) and many others, that there are certain environments in which coreference is not subject to Condition B effects. In particular, coreference was assumed to be possible in:

(4.56) Only Max himself voted for him.

(4.57) I know what John and Mary have in common.
John hates Mary and Mary hates her too.

(4.58) If everyone voted for Oscar, then certainly Oscar voted for him.

I remarked in section 2.1.1 that many of my informants actually find coreference very marginal in (4.56), (4.57) and (4.58), and emphasize that there are certainly much more natural ways to convey the intended messages. In the recent literature, several authors have acknowledged the controversial status of these data (cf. Schlenker, 2005; Grodzinsky, 2007; Heim, 2007). This complication was deliberately ignored in Part I, but should of course eventually be accounted for.

I think it would be best to interpret the judgments of my informants, as well as the remarks of Schlenker (2005), Grodzinsky (2007), and Heim (2007) as follows. On the one hand, the constructions in (4.56), (4.57) and (4.58) are felt to be improper ways of expressing coreference (informants often use the word *ungrammatical*). But on the other hand, there is something about these sentences which somehow gives the impression that coreference is in fact intended.

It should be remarked that it is not so uncommon for hearers to associate a sentence with a certain meaning even if that sentence does not constitute a proper way of expressing that meaning. For example, if someone says:

(4.59) John has three brother.

a hearer typically concludes that John has three brothers, even though the sentence is felt to be ungrammatical.

Thus, a theory of anaphora should explain, first, why (4.56), (4.57) and (4.58) are felt to be ungrammatical on a coreferential reading, and second, why these sentences somehow give the impression that coreference is in fact intended. Clearly, the theory presented in Part I does not provide such explanations.

4.8 Summary

In this chapter, I have considered various objections that could be raised against the theory proposed in the first part of this dissertation. First, the theory does not provide a unified account of pronominal anaphora and VP ellipsis. Second, the theory is partly stipulative. In particular, Condition B, Free Variable Economy, and VP Identity cannot be derived from general ideas about human behavior and/or cognition. Third, several usages of pronouns cannot be classified as either bound or coreferential. Fourth, several cases of VP ellipsis seem to refute the idea that VP ellipsis consists in PF deletion under VP Identity. Fifth, it seems that Dahl's puzzle should not be explained in terms of grammatical economy principles such as Locality or Free Variable Economy, but rather in terms of interpretive preferences, which may be overruled by world knowledge or contextual information. Furthermore, a complete analysis of Dahl's puzzle should not only explain why the strict-sloppy reading is so strongly dispreferred, but also the more basic observation that across-the-board readings are distinctly preferred over mixed ones. Finally, the borderline status of certain Condition B effects should be acknowledged and explained.

In the next chapter, I will present a theory that addresses these issues.

Chapter 5

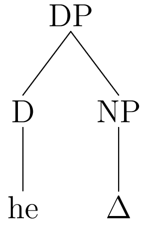
Resolution

This chapter presents a unified account of pronominal anaphora and VP ellipsis. The central assumption is that the meaning of anaphora is always retrieved from the context of use. In a slogan: anaphora are resolved. Moreover, it will be argued that constraints on the interpretation of anaphora follow from plausible assumptions concerning the resolution process and concerning the way people generally behave in communication.

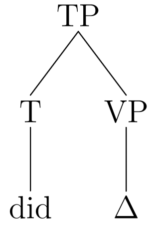
5.1 Unification

The syntactic framework assumed in Part I did not exhibit any parallel between pronominal anaphora and VP ellipsis. In particular, VP ellipsis was taken to involve deletion, whereas pronominal anaphora was not. In the recent literature, several syntactic analyses have been proposed that do treat pronominal anaphora and VP ellipsis analogously (cf. Lyons, 1999; Elbourne, 2005b, and the references given there). I will abstract away from the details of and the differences between these individual proposals, and assume an analysis which, I think, captures the essential analogy in a most perspicuous way. I will assume that a pronoun is a determiner with an empty NP complement and that VP ellipsis involves a tense auxiliary with an empty VP complement.

(5.1) The syntax of pronouns:



(5.2) The syntax of VP ellipsis:



Henceforth, I will often refer to VP ellipsis as VP anaphora and to pronominal anaphora as NP anaphora. In fact, NP anaphora does not necessarily involve a pronominal determiner; many other determiners also permit NP anaphora, as illustrated by the following examples from Elbourne (2005b, p.45):

(5.3) a. Sue only bought two books. Mary bought at least three.
b. Most movies bore Mary, but she does like some.
c. Most MIT students build robots, and all watch Star Trek.
d. Many Athenians went to Sicily, but few returned.

Below, I will focus on pronominal NP anaphora, but the proposal should in principle be applicable to NP anaphora more generally.

Semantically, I will assume that pronouns are just like definite articles, apart from the fact that they encode ϕ-features (number, person, and gender). This assumption is rather common in the literature and has received strong support from cross-linguistic studies (cf. Lyons, 1999; Elbourne, 2005b).

5.2 Resolution

The crucial idea that I want to defend here is that the meaning of the empty constituents in (5.1) and (5.2) is contextually determined. Typically, the meaning of an empty NP/VP constituent is retrieved from an antecedent in the linguistic context. But it may also be retrieved from the non-linguistic context, or from an inferred antecedent. Furthermore, it is not necessarily retrieved from a single antecedent; semantic material from several sources may be combined, as long as the result is of the right semantic type. Let me illustrate this with a few examples. First, consider a case in which a pronoun is resolved to a definite description.[1]

(5.4) The clown came in. He sat down.

The logical form of the second sentence contains an empty NP constituent Δ:

(5.5) [the clown came in] [he Δ sat down]

The meaning of Δ may be retrieved from the noun phrase [clown]. If it is, we will say that Δ is resolved to [clown] and write:

(5.6) [the clown came in] [he Δ sat down] $\Delta \rightarrow$ clown

Often, I will simply (and sloppily) say instead that [he] is resolved to [the clown] and write:

[1] Strictly speaking, I should say that this is a case in which the *complement* of a pronoun is resolved to the NP component of a definite description. This remark also applies to many cases discussed below.

(5.7) [the clown came in] [he sat down] he → the clown

Next consider a pronoun with an indefinite antecedent:

(5.8) A clown came in. He sat down.

Again, the logical form of the second sentence contains an empty NP constituent, which may be resolved to the noun phrase [clown]:

(5.9) [a clown came in] [he Δ sat down] Δ → clown

Next consider a pronoun whose antecedent is a proper name:

(5.10) a. John is in good shape. I think he will win.
b. [John is in good shape] [I think he Δ will win]

The meaning of Δ may be retrieved from [John], but we need something extra here, because [John] denotes an individual (at least so we assumed), and Δ must denote a property. There are at least two possible ways to settle this mismatch: one is to assume that proper names do in fact not denote individuals but properties (cf. Burge, 1973; Larson and Segal, 1995; Elbourne, 2005b). The other is to maintain the assumption that proper names denote individuals, and to add the assumption that the use of a proper name N does not only make salient the individual x denoted by N, but also the *property* of being identical to x. This, then, is a property to which an empty NP constituent may be resolved.

In (5.10b) for example, Δ may be resolved to the property of being John. If it is, I will simply say that [he] is resolved to [John] and write:

(5.11) [John is in good shape] [I think he will win] he → John

I will assume this latter option here (a name denotes an individual, but also makes salient the property of being identical to that individual), but I am not strongly committed to it. It may turn out that names really should be analyzed as denoting properties. This would not have far-reaching consequences for the theory proposed here. It would merely simplify it.

Besides names and descriptions, I will assume that pronouns can also have traces as their antecedents. Consider:

(5.12) a. Every man thinks he will win.
b. [every man]^1 [t_1 thinks he Δ will win]

The meaning of Δ may be retrieved from that of [t_1]. Only, [t_1] denotes an individual, whereas Δ must denote a property. To settle the mismatch, I will assume that Δ may be resolved to the property of being identical to the individual denoted by [t_1], just as in the case of proper names. If Δ is resolved to the property of being identical to the individual denoted by [t_1], I will simply say that [he] is resolved to [t_1] and write:

(5.13) [every man]^1 [t_1 thinks he will win] he $\rightarrow$ t_1

The meaning of a pronoun may also be retrieved from the non-linguistic context. For example, if I point at a certain athlete and say:

(5.14) a. He will win.
b. [he Δ will win]

Δ will generally be resolved to the property of being that athlete. There are situations in which Δ may be resolved to another property. For example, if we are watching a soccer game and I am explaining the rules of the game to you, I might point at one of the goalkeepers, say, Edwin van der Sar, and tell you:

(5.15) a. He is allowed to use his hands.
b. [he Δ is allowed to use his hands]

Then Δ is intended to be resolved to the property of being a goalkeeper and not to the property of being Edwin van der Sar.

We have also seen cases in which the meaning of a pronoun is retrieved neither from an explicit linguistic antecedent nor from the non-linguistic context, but rather from an inferred antecedent. For example, in (5.16), Δ may be resolved to the property of being a baby.

(5.16) [if I get pregnant, I will definitely keep it Δ]

Finally, let us consider donkey pronouns and paycheck pronouns. First, consider one of Geach's examples:

(5.17) a. Every farmer who owns a donkey beats it.
b. [every farmer who owns a donkey]^1 [t_1 beats it Δ]

For such cases I will assume, following Cooper (1979) and Heim and Kratzer (1998), that the meaning of Δ may be retrieved from several sources. For example, Δ may plausibly be resolved to the property of "being a donkey owned by t_1". Resolution to such a complex property may be forced here because resolution to a simpler property, such as that of being a donkey, would trigger the presupposition that there is a unique most salient donkey in the domain of discourse, which is not the case. Cooper (1979) observed that this strategy can also be applied to Karttunen's paycheck example:

(5.18) a. The man who gave his paycheck to his wife is smarter than the man who gave it to his mistress.
b. The man who gave his paycheck to his wife is smarter than the man [who]^1 [t_1 gave it Δ to his mistress]

Δ may be resolved to the property of being a paycheck of t_1. This yields exactly the intended truth-conditions.

I must remark here that a simpler treatment of paycheck pronouns may be given under the assumption that possessive DPs like [his paycheck] can have the following structure at LF:

(5.19) $[_{DP}$ $[_{D}$ the$][_{NP}$ paycheck of him$]]$

This assumption is adopted by Heim and Kratzer (1998) and Elbourne (2005b), among others. Elbourne (2005b, p.82) points out that evidence for it can be found in work of Larson and Cho (1999), who analyze the ambiguity of DPs like [John's former house]. This phrase may denote either the object John owns that was formerly a house, or the object that was formerly a house owned by John. Larson and Cho take this ambiguity to be structural, depending on the order in which the elements of the possessive DP are semantically composed. If we first compose the meaning of [house] with that of [former], and then compose the result of this with the meaning of [John's], we get the object John owns that was formerly a house. If we first compose the meaning of [John's] with that of [house], and then compose the result of this with the meaning of [former], we get the object that was formerly a house owned by John. What is relevant for the analysis of paycheck sentences is that for [John's] to compose with [house] before [former] does, it must be in a low position at LF. Thus, there is some evidence that possessive DPs like [his paycheck] may indeed have a structure like that in (5.19).

If this is the case, then the resolution of paycheck pronouns is as straightforward as can be:

(5.20) The man $[\text{who}]^1$ [t_1 gave [the [paycheck of him]] to his wife] is smarter than the man $[\text{who}]^1$ [t_1 gave [it Δ] to his mistress]

In the first clause, [him] and [his] are resolved to [t_1], and the empty NP in the second clause is then resolved to [paycheck of t_1].

So much for NP anaphora; let us now turn to VP anaphora. First, consider a simple strict/sloppy ambiguity:

(5.21) Max called his mother, and Bob did too.

The logical form of the source clause is:

(5.22) $[\text{Max}]^1$ [t_1 called his mother]

where [his] can be resolved either to [Max] or to [t_1]. The logical form of the target clause is:

(5.23) $[\text{Bob}]^1$ [t_1 did Δ too]

Δ is resolved to the VP in the source clause. If [his] was resolved to [Max] we get a strict reading; if [his] was resolved to [t_1] we get a sloppy reading.

Just like empty NP elements, empty VP elements may also be resolved deictically (see examples (4.8)–(4.11)) or to inferred antecedents (see examples (4.12) and (4.32)–(4.39)). As noted in section 4.1, deictic resolution of VP anaphora is not as common as deictic resolution of pronominal NP anaphora. I mentioned two reasons for why this may be so. First, pronouns usually carry some (gender/number/person) information, which greatly facilitates deictic resolution. VP anaphora do not convey such information. Second, there are many more situations in which a particular object is the single most salient (female/singular/third person) object than there are situations in which a certain property or activity is the single most salient property or activity. Thus, VP anaphora are generally much harder to resolve deictically than pronominal NP anaphora. Incidentally, *non-pronominal* NP anaphora seem to pattern with VP anaphora in this respect: it is quite uncommon for non-pronominal NP anaphora to be resolved deictically, and this should be expected given that non-pronominal determiners do not encode ϕ-features.

Now let us consider the resolution of anaphora to inferred antecedents in somewhat more detail. I will assume that this process imposes a higher processing load on the hearer than resolution to non-inferred antecedents. Therefore, a hearer will only consider inferred antecedents if really necessary. This could be because the context does not provide any suitable explicit antecedents, or because resolving the anaphora to any of the given explicit antecedents yields an interpretation that is incoherent or inconsistent with world knowledge and/or contextual information. As observed by Hardt (2005), such a restriction on the use of inference in resolution is necessary to explain contrasts like that between Webber's original example, repeated here as (5.24), and the variant in (5.25).

(5.24) Irv and Mary wanted to dance together, but Mary couldn't, because her husband was there.

(5.25) Irv and Mary wanted to dance together, but Tom and Sue didn't.

The elided VP in (5.24) may be resolved to the inferred antecedent [dance with Irv], because there is no suitable explicit antecedent. In (5.25), on the other hand, there is a suitable explicit antecedent, namely the verb phrase [want to dance together]. Therefore, the inferred antecedent [want to dance with Irv] does not come into play.

This line of reasoning also yields a natural treatment of cascaded VP anaphora. Recall that typical cases of cascaded VP anaphora, such as (5.26), do not allow mixed readings, whereas some special cases, such as (5.27) and (5.28), do.

(5.26) Bob called his mother, and Max did too. But Tom didn't.

(5.27) Smithers thinks his job sucks, and Homer does, too. But Marge doesn't.

(5.28) John didn't wash his car, but Bill did, even though Harry already had.

In cases like (5.26), the elided VP in the second clause is resolved to the explicit VP in the first clause, and the elided VP in the third clause is resolved either to the explicit VP in the first clause or to the copy of that VP in the second clause. In any case, we either get a strict reading for both elided VPs or we get a sloppy reading for both elided VPs. The examples in (5.27) and (5.28) are special, because they may trigger inference. Suppose, for example, that the pronoun in the first clause of (5.27) is resolved to the trace of [Smithers], and that the elided VP in the second clause is resolved, as usual, to the VP in the first clause. This means that the second clause attributes to Homer the property [$\lambda x.$ x thinks that x's job sucks]. Now consider the elided VP in the third clause. Normally, this VP would be resolved to the VP in the first clause or to the copy of that VP in the second clause. This would mean that the third clause assigns the above property to Marge. But this is inconsistent with world knowledge: Marge doesn't have a job. This triggers inference. From the information that Homer has the property [$\lambda x.$ x thinks that x's job sucks] it can be inferred that Homer has the property [$\lambda x.$ x thinks that Homer's job sucks]. This inference provides a suitable antecedent for the elided VP in the third clause, and yields exactly the attested mixed reading.

A similar story applies to (5.28). Here, a sloppy interpretation of the elided VPs in the second and third clause would give rise to an incoherent discourse. In particular, the contrast indicated by *even though* and *already* would not be established. This triggers an inference parallel to the one in (5.27), which in turn yields the attested mixed reading.

To the best of my knowledge, this is the first successful analysis of mixed readings in cascaded ellipsis. Theories based on VP Identity are all too rigid (they don't allow for mixed readings at all). Theories based on NP Parallelism (Fox, 1999a; Büring, 2005b) or unification (Dalrymple *et al.*, 1991) are too flexible (they allow mixed readings even in cases like (5.26)). The same goes for theories which assume that sloppy readings arise because pronouns in the antecedent VP may be reinterpreted at the ellipsis site (Hardt, 1999; Schlenker, 2005). Of course, this flexibility may be restricted by independently motivated constraints related, for instance, to information structure (Focus Match) or discourse structure (cf. Prüst *et al.*, 1994; Hardt and Romero, 2004). But such constraints won't be able to account for the contrast between (5.26) and (5.27), because there is no pertinent difference between these two cases as far as information structure and discourse structure are concerned. The idea that inference in anaphora resolution must be triggered really seems to be the only viable explanation.

Finally, let us turn to the unexpected sloppy readings observed by Wescoat and Hardt. Consider Wescoat's example:

(5.29) The police officer who arrested John insulted him, and
the police officer who arrested Bill did, too.

As noted above, the source clause of (5.29) is structurally analogous to one of Geach's donkey examples:

(5.30) a. Every man who owns a donkey beats it.
b. [every man who owns a donkey]^1 [t_1 beats it Δ]

I assumed above, following Cooper (1979), that Δ may be resolved to the property of "being a donkey owned by t_1" in this case. As observed by Tomioka (1999), a similar strategy can be applied to (5.29). First consider the logical form of the source clause:

(5.31) [the police officer who arrested John]^1 [t_1 insulted him Δ]

Here, Δ may be resolved to the property of "being arrested by t_1". Next, consider the logical form of the target clause:

(5.32) [the police officer who arrested Bill]^1 [t_1 did Δ]

Now Δ can just be resolved to the relevant VP in the source clause, which can be glossed as [insulted the person arrested by t_1]. This yields the attested sloppy reading.

Thus, all the cases of anaphora that were problematic for the theory proposed in Part I are now dealt with in a rather straightforward way. Moreover, there is no longer any need to stipulate a Semantic Identity constraint on VP ellipsis. The fact that the meaning of an elided VP must be identical to the meaning of its antecedent (in case it is not resolved deictically and there is no inference involved in its resolution) simply follows from the way resolution works: the meaning of the elided VP is *retrieved* from the meaning of the antecedent VP. As a result, the meaning of the two VPs must—in non-deictic, non-inferential cases—be identical.

5.3 Anaphoric Relations

The distinction between inherently bound pronouns and inherently referential pronouns has been dropped. All pronouns are assumed to be definite articles with empty NP complements, the meaning of which is contextually retrieved. However, depending on how a pronoun is resolved, we may still think of it as being bound, cobound, covalued, or coreferential with another DP, in a sense that is very much in line with the way in which these terms were used in Part I. For example, covaluation can be defined as follows:[2]

[2] Context sets, $\mathcal{F}$, $\mathcal{I}$, and equivalence given $\mathcal{F}$ and $\mathcal{I}$ are defined on pages 18, 19, and 31.

5.1. Definition. [Covaluation]
Let C be a context and let s_c be the context set of C. Then, two expressions A and B are covalued in C iff for every $w \in s_c$, $[\![A]\!]^{C}(w)$ is equivalent to $[\![B]\!]^{C}(w)$ given $\mathcal{F}$ and $\mathcal{I}$.

Coreference can be seen as a special case of covaluation, namely the one involving only referential expressions (expressions of type *se* whose translation does not contain any free variables).

5.2. Definition. [Coreference]
Two expressions corefer in a context C iff they are referential and covalued in C.

Binding may be defined as follows:

5.3. Definition. [Binding]
A moved DP always binds its own trace. Moreover, if X is a logical form constituent, A a moved DP in X, B a pronoun in X, and C a context, then A binds B in LF/C iff:

i B is covalued with the trace of A in C;

ii A c-commands B in X;

iii A does not c-command any other NP in X which satisfies i and ii.

This notion of binding is very similar to the one defined in Part I, and therefore also very similar to what Heim and Kratzer (1998) and Büring (2005a) call *semantic binding* and what Reinhart (2006) calls *A-binding.* The crucial difference is that the present notion is not defined in terms of indices, but rather in terms of covaluation between a pronoun and a trace.
Finally, cobinding can be defined in terms of binding:

5.4. Definition. [Cobinding]
If X is a logical form constituent, A and B two nodes in X, and C a context, then A and B are cobound in LF/C iff there is a third node which binds both A and B in LF/C.

Let me illustrate these notions by means of a simple example:

(5.33) $[\text{John}]^1$ $[\text{t}_1$ thinks he will win]

If [he] is resolved to [John] then [he] and [John] are covalued and even coreferential (because [John] is a referential expression). On the other hand, if [he] is resolved to $[\text{t}_1]$ then (i) [he] and $[\text{t}_1]$ are covalued (though not coreferential), (ii) [he] is bound by [John], and therefore (iii) [he] and $[\text{t}_1]$ are cobound.

Thus, the familiar notions of anaphoric relatedness can be maintained, even though pronouns are no longer assumed to be either inherently bound or inherently referential. One consequence of this is that the ambiguities in (1.3), (1.4), and (1.5) can still be explained just as they were in chapter 1. A second consequence is that the formulation of Movement Economy in section 3.4 is still valid. Thus, we don't need a new account of crossover effects. Dahl's puzzle and Condition B effects, however, do require a new analysis.

5.4 Dahl's Puzzle

In section 4.6 it was pointed out that Dahl's puzzle requires two kinds of explanations. First, it should be explained why, in neutral contexts, across-the-board readings are preferred over mixed readings. Second, it should be explained why the sloppy-strict reading is easier to accommodate than the strict-sloppy reading.

I propose that the first issue, the preference for across-the-board readings, is due to a general preference for *local* resolution. The idea that such a preference exists is plausible given the incremental nature of the interpretation process, and the limited capacity of short-term memory. To see how this explains the preference for across-the-board readings, consider the source clause of Dahl's example:

(5.34) [Max]^1 [t_1 said that [he]^2 [t_2 called his mother]]

Assuming a preference for local resolution, the pronoun [his] will preferably be resolved to [t_2] or to [he] rather than to [t_1] or [Max].[3] Thus, the preferred resolutions are:

(5.35)
a. he → Max
his → he

b. he → Max
his → t_2

c. he → t_1
his → he

d. he → t_1
his → t_2

These resolutions all give rise to across-the-board readings of the elided VP in the target clause: (5.35a) and (5.35b) yield the strict-strict reading, while (5.35c) and (5.35d) yield the sloppy-sloppy reading. Hence, the preference for local resolution explains the preference for across-the-board readings in Dahl's puzzle.

The idea that resolution is preferably local is of course reminiscent of Fox's Locality constraint. But the two are really quite different. Locality is a grammatical principle, which classifies certain logical forms as ungrammatical. The local

[3]There may also be a slight preference for [t_2] over [he] and for [t_1] over [Max], but I will assume that this preference is negligible. In general, if two possible antecedents are directly adjacent, I will assume that the difference between them is too small to induce a noticeable preference.

resolution preference is an interpretive preference, which explains why certain interpretations of a given sentence are more accessible than others. Dahl's example is one case in which the two yield different predictions. Another example which is worth highlighting is the one discussed in section 3.2:

(5.36) Every man said that he called his mother and that Bill did too.

This sentence has the following two readings:

(5.37) a. Every man x said that x called x's mother and that Bill called Bill's mother too. [sloppy]
b. Every man x said that x called x's mother and that Bill called x's mother too. [strict]

It was observed in section 3.2 that this example is problematic for Locality and many other accounts of Dahl's puzzle, because they all predict the strict reading in (5.37b) to be unavailable. This prediction does not follow from the local resolution preference. To see this, consider the following logical form of (5.36):

(5.38) [Every man]^1 [t_1 said [[that [he]^2 [t_2 called his mother]] and [that [Bill]^2 [t_2 did Δ]] too]]

We are only interested of course in readings in which [he] and [his] are anaphorically related to [every man]. Thus, [he] must be resolved to [t_1], and [his] must be resolved to [t_1], [he], or [t_2]. Now, if resolution is preferably local, [his] will preferably be resolved to [he] or to [t_2] (and not to [t_1]). These two possibilities lead exactly to the strict and the sloppy reading in (5.37a) and (5.37b). This is another case, then, in which the local resolution preference makes different, and more desirable predictions than Locality.

Now let us turn back to Dahl's puzzle. It must still be explained why one of the mixed readings is easier to accommodate than the other. Both these readings are harder to get than across-the-board readings, but many people find the sloppy-strict reading significantly more accessible than the strict-sloppy reading. I propose the following account of this contrast. First, it should be noted that people generally need some time to decide whether the mixed readings are acceptable or not (the across-the-board readings are generally judged ok without much reflection). It seems that people use this time to try and figure out a specific context in which the reading they are considering is likely to be the intended reading. We could say that people try to find a context which *supports* the reading under consideration, where a context C is defined to support a reading R of a sentence S iff R is a likely reading of S in C. Now, in the case of Dahl's puzzle, it is relatively straightforward to find a context which supports the sloppy-strict reading. For example, if the question under discussion is:

(5.39) Who called Max's mother?

then the sloppy-strict reading (Max said that Max called Max's mother and Bill said that Bill called Max's mother) is likely to be intended.

There are also contexts which support the strict-sloppy reading. One such a context was given in (2.33) in Part I, repeated here as (5.40):

(5.40)
a. Did Max call everyone's mother?
b. Well, I don't know...
c. Max said he called his mother, and Bob did too.
d. But I haven't heard from Sue and Mary yet.

Other contexts supporting the strict-sloppy reading were given in (4.50) (Hardt's lawsuit case) and (4.51) (Reuland's gambling case). However, there are good reasons to believe that these contexts are much harder to evoke than contexts that support the sloppy-strict reading. The question in (5.39) is relatively simple: its logical structure can be represented as $?x.R(x, m)$ (read: *which x stand in relation R to m?*), where R is a simple relation, namely that of calling, and m is a simple individual, namely Max's mother. The question in (5.40a) is rather more complex. First of all, it is ambiguous between the two readings given in (5.41) and (5.42):

(5.41)
a. $?\forall x.R(\text{max}, \text{mother}(x))$
b. Is it true that for every x, Max called x's mother?
c. Possible answers: *yes*, *no*.

(5.42)
a. $\forall x.?R(\text{max}, \text{mother}(x))$
b. For every x, is it true that Max called x's mother?
c. Possible answer: well, Max called Max's mother, and he called Bob's mother, but I don't know whether he called Sue's mother and Mary's mother.

Only if the quantifier takes wide scope, as in (5.42), does the question license Dahl's sentence as a possible (partial) response. Notice, however, that there is a rather strong preference for the narrow scope reading in (5.41) over the wide scope reading in (5.42).[4] Moreover, even if the quantifier *is* given wide scope, it is unlikely that someone would use Dahl's sentence as a *complete* response to (5.40a) (the discourse in (5.40) becomes very odd if (5.40b) and (5.40d) are left out). Presumably, this is because *everyone* is unlikely to quantify just over Max and Bob. In any case, the relevant observation is that the question in (5.40a), and the way in which it may support the strict-sloppy reading of Dahl's sentence, is not nearly as straightforward as the question in (5.39), and the way in which

[4]The work of Groenendijk (2007) provides an interesting explanation for this preference. Roughly speaking, less inquisitive questions are generally preferred over more inquisitive questions (just as more informative assertions are generally preferred over less informative assertions) and the question in (5.41) is indeed less inquisitive than the one in (5.42) (a complete answer to the second question always entails an answer to the first, but not the other way around).

it supports the sloppy-strict reading of Dahl's sentence. Clearly, this observation also applies to Hardt's lawsuit case and Reuland's gambling case, where the strict-sloppy reading of Dahl's sentence is supported not just by a simple question, but rather by a whole plot.

I suggest, then, that in general, the level of accessibility of a particular reading for a given sentence will correlate with the complexity of the contexts that support this reading, and that, in particular, this is what explains the contrast in accessibility between the two mixed readings in Dahl's example.

5.5 Condition B Effects

Throughout the first part of this dissertation it was assumed, following Reinhart and many others, that Condition B effects are to be accounted for by two distinct mechanisms. The first accounts for Condition B effects on binding; the second accounts for Condition B effects on other kinds of codetermination. The primary piece of evidence in favor of such a two-level approach is Reinhart's observation that coreference is sometimes exceptionally allowed in Condition B environments. The relevant cases are repeated below:

(5.43) Only Max himself voted for him.

(5.44) I know what John and Mary have in common.
John hates Mary and Mary hates her too.

(5.45) If everyone voted for Oscar, then certainly Oscar voted for him.

However, as discussed in section 4.7, this judgment is disconfirmed by many informants. (5.43), (5.44) and (5.45) are generally felt to be ungrammatical on a coreferential reading, even though it is typically considered likely that such a reading is in fact intended.

It might be possible to formulate a two-level theory which accommodates this assessment of (5.43)–(5.45). But it wouldn't make much sense to do so. Reinhart's assessment of (5.43)–(5.45) was adduced as primary evidence for a two-level approach. If this assessment turns out to be inaccurate, the motivation for the whole approach goes up in smoke. This really concerns the approach in general, not just Reinhart's or anyone else's theory in particular. If there is no significant motivation for a two-level approach[5], we may as well pursue a

[5]It must be noted here that, apart from the alleged acceptability of coreference in constructions like (5.43)–(5.45), the two-level approach has also been supported by certain findings in the acquisition literature (cf. Chien and Wexler, 1990; Grodzinsky and Reinhart, 1993). More recently, however, the validity of these findings has been disproved quite convincingly (Elbourne, 2005a; Conroy *et al.*, 2007). Thus, I take the constructions in (5.43)–(5.45) to constitute the only alleged piece of evidence for a two-level approach to Condition B effects in English (see Heim, 2007; Grodzinsky, 2007; Conroy *et al.*, 2007, for concurrent views).

simpler, "one-level" explanation of Condition B effects in English. And such an explanation can indeed be given.

I think that the essential source of Condition B effects in English is the fact that speakers of English have come to use the marker *-self* to indicate that a pronoun should be resolved to one of its coarguments, and as a consequence, hearers have come to assume that, whenever a speaker does not use such a marker, interpretations that would result from coargument resolution are not intended.

I will refer to marked pronouns such as *himself* and *herself* as *self-pronouns*. Many authors refer to *himself* and *herself* as *anaphors*, following Chomsky (1981), or *self-anaphors*, following Reinhart and Reuland (1993). I have chosen not to adopt these terms for two reasons: (i) to avoid confusion with the term *anaphora*, and (ii) to remain neutral with respect to the claim that words like *himself* in English share some essential characteristics with words like *zich* in Dutch and *sig* in Icelandic, which are also called anaphors. Other authors refer to *himself* and *herself* as *reflexive pronouns* or simply as *reflexives*. I avoid these terms because *himself* and *herself* are not only used to indicate that a reflexive interpretation is intended. They are also used as intensifiers, marking prominence and contrast, possibly among other things (cf. Baker, 1995). In fact, in the history of English, the use of self-pronouns as intensifiers preceded their use as reflexivity markers (cf. König and Siemund, 2000).

I will assume that if a self-pronoun is used to mark reflexivity, it may be resolved either to one of its coarguments, or to the trace of one of its coarguments. Thus, self-pronouns may be interpreted as bound variables, but also referentially. In the literature, it is often assumed that self-pronouns can only be interpreted as bound variables. However, this assumption is problematic: it wrongly predicts that the question-answer pair in (5.46) below is incongruent, and that the sentence in (5.47) (adapted from Dalrymple, 1991) does not have a strict reading (saying that Bill's lawyer couldn't defend Bill against the accusations). These examples clearly show that self-pronouns cannot only be interpreted as bound variables, but also referentially.[6]

(5.46) a. Who evaluated John?
b. He evaluated himself.

(5.47) Bill defended himself against the accusations because his lawyer couldn't.

[6]It should be remarked here that there are also cases of VP ellipsis involving self-pronouns which do not admit strict readings. For example:

(i) John defended himself, and Bob did too.

However, the contrast between (i) and (5.47) can be explained on independent grounds (see Kehler (2002) and the discussion on page 127 below).

I will refer to any interpretation that results from resolving a (self-)pronoun to (the trace of) one of its coarguments as a *reflexive interpretation.* Finally, I will refer to the convention that speakers always use a self-pronoun if they intend a reflexive interpretation as the Reflexivity Convention.

5.5. Definition. [Reflexivity Convention]
If a reflexive interpretation is intended, this is indicated by using a self-pronoun.

The Reflexivity Convention does not only account for Condition B effects on binding, but also for Condition B effects on other kinds of codetermination. An unmarked pronoun will never be interpreted as codetermined with one of its coarguments, because this would yield a reflexive interpretation, and such an interpretation could only be intended if the speaker had used a self-pronoun.

The idea that the Reflexivity Convention is the source of Condition B effects in English is strongly supported by the following two facts. First, in a broad range of languages, the existence of Condition B effects correlates with the existence of reflexivity markers (cf. Levinson, 2000; Huang, 2000). In particular, languages without reflexivity markers do not exhibit Condition B effects. Second, languages like English have gradually developed from an earlier stage, without reflexivity markers and without Condition B effects, to the current stage, with reflexivity markers and with Condition B effects (cf. Levinson, 2000; König and Siemund, 2000; van Gelderen, 2001; Keenan, 2002). The same development has been observed in several Creole languages (cf. Carden and Stewart, 1988; Levinson, 2000). Levinson (2000) provides a particularly attractive explanation of the crucial steps in this evolutionary process.

At an early stage, a language may not have any reflexivity markers and unmarked pronouns may freely be resolved to coarguments. This was the case, for instance, in Old English. However, there is a general tendency, even at such a stage, *not* to resolve pronouns to coarguments, for the simple reason that the agent and the patient of most actions are stereotypically distinct. Then, reflexivity markers gradually come into existence as "markers of the unusual": if a speaker intends a reflexive interpretation, he uses a marked construction (e.g., a self-pronoun) to signal to the hearer that something unusual is intended. This is an instance of what Horn (1984) called *the division of pragmatic labor*: unmarked forms are associated with stereotypical interpretations, while marked forms are associated with non-stereotypical interpretations. It should be noted that some verbs describe actions whose agent and patient are stereotypically identical (e.g., grooming verbs like *shaving* and *washing*). It should be expected, then, that a reflexive interpretation of such verbs does not necessarily involve special marking at this stage. This has indeed been observed, for example in Middle English (Faltz, 1985, p.242) and in Frisian (Reuland, 2001, p.478). Over time, though, the association between reflexive interpretations and reflexive marking becomes stronger and stronger and eventually leads to the Reflexivity Convention.

Levinson (2000) provides a wide range of cross-linguistic and diachronic data to support this hypothesis. Thus, the idea that Condition B effects in English stem from the Reflexivity Convention is well-motivated and well-supported.

Let us now return to the disputed Condition B effects in (5.43), (5.44) and (5.45). Two observations should be explained. First, these examples are generally felt to be ungrammatical on a reflexive interpretation. Second, however, informants often feel that a reflexive interpretation may nevertheless be intended.

The first observation is explained by the Reflexivity Convention. If a reflexive interpretation is intended, this should be indicated by a reflexive marker, and such a marker is not present in (5.43), (5.44) and (5.45).

There are several reasons why a reflexive interpretation may nevertheless seem to be intended in these examples. The case of (5.44) is relatively straightforward: the second sentence in (5.44) is supposed to convey what John and Mary have in common. If [her] is resolved to [Mary], there is indeed a property that is attributed to both John and Mary, namely that of hating Mary. If [her] is resolved in some other way, the sentence does not tell us which property John and Mary have in common. Therefore, it seems likely that a reflexive interpretation is intended, even though it is not properly expressed. An additional indication that this is the case is the use of the particle [too]. If [her] is resolved to [Mary], the use of this particle is justified, but if [her] is resolved in some other way, it is hard to see why [too] should have been used here.

The case of (5.45) is different but equally straightforward: only if [him] is resolved to [Oscar] does the sentence present a valid argument. If [him] is resolved in some other way, the sentence presents a nonsensical argument. Thus, a reflexive interpretation is probably intended, even though it is not properly expressed.

Example (5.43) is more subtle. I think that the crucial element here is not so much the focus-sensitive particle *only*, but rather the intensifier *himself*.[7] When confronted with examples like (5.43), informants quite often report that a reflexive interpretation is probably intended. But when confronted with examples like (5.48) (without intensifier), they don't.[8]

(5.48) Only Max voted for him.

[7]Self-pronouns can be used as intensifiers in several ways. For instance, in (5.43) and in (i) below the self-pronoun is used as an *adnominal* intensifier, while in (ii) below it is used as an *adverbial* intensifier:

(i) The President himself opened the exhibition.

(ii) The President opened the exhibition himself.

I am only concerned here with adnominal intensifiers, to which I will simply refer as intensifiers.

[8]This contrast has, to the best of my knowledge, not been noted previously, perhaps because it cannot be accounted for by any two-level theory of Condition B effects.

Thus, there must be something about intensifiers that makes the reflexive interpretation in (5.43) particularly salient. Let me try to pin down what this is.

The standard analysis of adnominal intensifiers, due to Eckardt (2001) and Hole (1999) (see also Gast, 2006; Eckardt, 2006; König and Gast, 2006), is that they adjoin to DPs and denote the identity function on the domain of individuals. Thus, the denotation of *Max himself* is obtained by applying the identity function to the denotation of *Max*. Intensifiers, then, do not make any contribution to the ordinary semantic value of a sentence. However, they do make a significant contribution to the focus semantic value: intensifiers are always in focus (accented), and therefore, just like other focused elements, evoke a set of alternatives. These alternatives are contextually determined functions, other than the identity function. For example, in (5.49) the contextually triggered alternative function is the one mapping people to their family members and in (5.50) it is the one mapping kings to the members of their court.

(5.49) John and his family are deciding where to spend their holidays.
John himself wants to go to Greece.

(5.50) The king himself opened the door.

Intensifiers interact with focus-sensitive particles like *only* just like other focused elements do. For example, (5.51) entails that John's family members do not want to go to Greece (see Eckardt (2001) for more illustrations of the fact that intensifiers behave just like other focused elements).

(5.51) John and his family are deciding where to spend their holidays.
Only John himself wants to go to Greece.

The crucial difference between intensified nominals and simply focused nominals is that the referent of an intensified nominal must be particularly *prominent* (Baker, 1995). This prominence may come from several sources. One possible source is world knowledge. For example, nominals like *the king* and *the President* refer to individuals who are prominent because of the role they play in society. Another possible source is the discourse. In particular, the prominence of a referent may be due to its being the discourse topic or the so-called subject of consciousness (the person whose perspective is taken in the discourse). In Baker's (1995) terms, the prominence of a referent may be justified either *externally* (i.e., by world knowledge) or *internally* (i.e., by the discourse).[9] The importance of this prominence-factor is illustrated by the following contrast:

[9] As observed by Baker, most examples of intensification in the linguistic literature involve nominals like *the king* and *the President*, such that prominence is justified externally. The prominence of intensified nominals "in the wild", however, is mostly justified internally, i.e., by the surrounding discourse.

(5.52) a. Eric Clapton is working on a new album with his band.
The members of the band are showing up at the studio every morning around 9am.
Clapton himself usually joins them in the afternoon.

b. Eric Clapton is working on a new album with his band.
Most members of the band are showing up at the studio every morning around 9am.
The drummer himself usually joins them in the afternoon.

What is especially relevant for examples like (5.43) is that, if a sentence is considered in isolation, and if the prominence of an intensified nominal in that sentence is not justified externally (i.e., by world knowledge), then it is supposed to be justified internally (i.e., by the (missing) surrounding discourse). Let me illustrate. Suppose the second sentence in (5.49) is considered in isolation:

(5.53) John himself would like to go to Greece.

The prominence of John is not justified externally, so it must be justified internally: the (missing) preceding discourse must be one in which John is particularly prominent, for example, one in which John figures as the discourse topic.

Now let us turn back to example (5.43), repeated here:

(5.43) Only Max himself voted for him.

This sentence tells us two things about the kind of discourse context in which it may occur. First, the use of *only* and the focus on *himself* indicate that the preceding discourse must be one in which, for some person p, the issue:

(5.54) Who voted for p?

has been raised. This is the issue, then, that (5.43) addresses.

Second, the use of the intensifier in (5.43) indicates that the preceding discourse must be one in which Max is particularly prominent. Given these two indications, the simplest assumption to make is that the discourse preceding (5.43) is one in which the following issue has been raised:

(5.55) Who voted for Max?

But if this is the issue that (5.43) addresses, then the pronoun must be resolved to Max, and this yields a reflexive interpretation. This is, I think, the reason why informants sometimes feel that a reflexive interpretation might be intended in (5.43), even though it is not properly expressed.

Thus, we have an explanation for why (5.43), (5.44) and (5.45) are felt to be ungrammatical on a reflexive interpretation (in terms of the Reflexivity Convention), but also for the fact that these sentences evoke the impression that a reflexive interpretation may nevertheless be intended.

It should be emphasized that *reflexive interpretations* are interpretations which result from resolving a pronoun to one of its coarguments *or* from resolving a pronoun to the trace of one of its coarguments. We could call these two kinds of reflexive interpretations *coreferential* and *bound*, respectively, hopefully without causing confusion. Now, when confronted with examples like (5.43) and (5.44), informants often feel that a *coreferential* reflexive interpretation may be intended, but they never feel that a *bound* reflexive interpretation may be intended. For example, (5.43) could possibly be supposed to mean that Max was the only one who voted for Max, but it can certainly not be supposed to mean that Max was the only "self-voter" (the only one with the property [$\lambda x.x$ voted for x]). This observation is accounted for by the explanations given above. In particular, the statement that Max was the only self-voter does not address the issue in (5.55), and being a self-hater cannot be the property that John and Mary have in common according to (5.44).

Finally, let me remark that the analysis of Condition B effects proposed here differs from that of Levinson (2000), even though Levinson's work provides much support for it. The crucial difference is this: the Reflexivity Convention says that whenever a reflexive interpretation is intended, this must be indicated by means of a self-pronoun. Levinson assumes that (i) a self-pronoun in argument position must be resolved to one of its coarguments, (ii) a self-pronoun is *more informative* than an unmarked pronoun, and therefore (iii) the use of an unmarked pronoun *implicates* that a reflexive interpretation is not intended (just as *some students passed the test* implicates that *not all students passed the test*).[10] The main problem with this proposal, in my view, is that Condition B effects are not *cancelable* in the way implicatures generally are. To see this, consider the contrast between (5.56) and (5.57):

(5.56) a. Some students passed the test.
b. In fact, it is possible that all of them passed.

(5.57) a. John thinks that Bill voted for him.
b. ??In fact, it is possible that John thinks that Bill voted for himself.

(5.56a) implicates that not all students passed the test. This implicature is canceled in (5.56b). It is a characteristic feature of implicatures that they are cancelable in this way. Thus, if Condition B effects are implicatures, as Levinson suggests, we should expect that they are cancelable too. Example (5.57) shows that this is not the case. Thus, although from a historical perspective it is plausible that the pragmatic inference patterns described by Levinson have played an important role in the realisation of the Reflexivity Convention, I don't think that they provide a suitable account of Condition B effects in present-day English.

[10] A summary of Levinson's proposal can be found on pp.347–348 of his (2000) book.

The issues raised in chapter 4 have now been resolved. A unified treatment of pronouns and VP ellipsis has been established. The stipulative Identity condition on VP ellipsis has been eliminated. Pronouns which could not be classified as either bound or coreferential, and instances of VP ellipsis which could not be dealt with in terms of VP Identity are no longer problematic. Dahl's puzzle has received a refined treatment. And finally, Condition B effects have been dealt with in a satisfactory way.

I now turn to a brief discussion of how the central ideas proposed here are related to previous and ongoing work of others.

5.6 Related Work

Early Ancestors. In the early days of generative grammar, Wasow (1972) proposed a theory of anaphora that has remarkably much in common with the theory defended here. In particular, Wasow argued that pronominal anaphora and VP ellipsis should be treated in a unified manner, and that anaphora involves resolution rather than deletion. Another early proponent of resolution, especially for the case of VP ellipsis, was Williams (1977).

Diversification. These ideas should have been standard ever since. But instead, much energy has been devoted to exploring several alternatives. As mentioned in section 4.1, one important reason for exploring such alternatives was the work of Hankamer and Sag (1976), who argued for a fundamental distinction between *deep anaphora* and *surface anaphora*. Pronouns were classified as deep anaphora and analyzed in terms of resolution, while VP ellipsis was classified as surface anaphora and analyzed in terms of deletion. I already pointed out that Hankamer and Sag's main argument has been refuted and that even Sag himself recently proposed that VP ellipsis should be dealt with in terms of resolution rather than deletion. Other authors who have argued for a resolution approach to VP ellipsis include Hardt (1993, 1999) and Kehler (2002). But the deletion approach is still quite widely adopted (cf. Heim and Kratzer, 1998; Merchant, 2001).

The other reason why many authors have departed from a unified theory of anaphora is the fact that Reinhart (1983) and many others have argued for a distinction between bound and referential pronouns, as discussed at length in Part I. Such a distinction does of course not permit a unified analysis of pronouns, let alone of pronouns and VP ellipsis.

Re-unification. The move I made in this chapter was to replace the idea that pronouns are *inherently* bound or referential by the alternative conception that pronouns may *end up* either as bound variables or as referential expressions, depending on how they are resolved (e.g., to a trace or to a referential antecedent).

Heim and Kratzer (1998) made a similar move. That is, they also suggested that pronouns should not be treated as inherently bound or referential, but rather end up as bound variables in some contexts and as referential expressions in others.

Heim and Kratzer's implementation of this idea, however, is different from mine. I treat all pronouns as expressions whose meaning is to be determined contextually. In particular, a pronoun is interpreted as a variable iff it is resolved to a trace. Heim and Kratzer propose that all pronouns are treated as variables. These variables, then, may end up bound, or remain free, in which case they are interpreted as referring to some contextually salient entity.

My proposal has at least three advantages over Heim and Kratzer's. First, as Heim and Kratzer (1998, chapter 11) show in detail, certain pronouns cannot be treated as plain variables (examples of such pronouns are donkey pronouns and paycheck pronouns, see (4.29)–(4.31) above). Thus, Heim and Kratzer do not establish a completely unified analysis of pronouns. Second, pronominal anaphora are treated very differently in Heim and Kratzer's system from non-pronominal NP anaphora and VP ellipsis. The theory I have proposed treats all these kinds of anaphora in a unified manner. Finally, certain cases of VP ellipsis force Heim and Kratzer (1998, p.254) to stipulate an additional constraint on logical forms: "no LF representation must contain both bound occurrences and free occurrences of the same index". This constraint does not have any independent motivation. Indeed, it only arises because referential pronouns are embodied as free variables in Heim and Kratzer's system. On my proposal, it does not have to be stipulated.

Elbourne (2005b) elaborates on Heim and Kratzer's work in order to overcome the first two problems. He analyzes pronouns as definite articles whose NP complement is either an index or a full NP which is deleted at PF under identity with some other NP in the discourse. The indexed pronouns are translated as variables, which may end up either bound or free (referential), just as in Heim and Kratzer's system. Pronouns that cannot be analyzed as bound or referential, such as donkey and paycheck pronouns, are captured by NP-deletion. Thus, Elbourne establishes a uniform account of pronouns which is very much reminiscent of—and can indeed be unified with—a deletion approach to VP ellipsis.

The crucial difference with my proposal is that on Elbourne's view, certain pronouns have an indexical complement and others have a full NP complement which is deleted at PF, while on my view, all pronouns have an empty NP complement whose meaning is contextually retrieved.

One advantage of my proposal, then, is that it does not need to postulate the existence of indices as "lexical items". Another advantage has to do with the fact that resolution provides more freedom in the interpretation process than NP-deletion does. Elbourne argued that this freedom is problematic, but I will counter Elbourne's arguments below and show that the freedom provided by resolution is indeed needed.

NP-deletion versus Resolution. Elbourne presents two arguments against resolution. The first is based on pairs of sentences like those in (5.58) and (5.59) (Elbourne, 2005b, p.64, originally from Heim, 1982, 1990).

(5.58) a. Every man who has a wife is sitting next to her.
b. #Every married man is sitting next to her.

(5.59) a. Someone who has a guitar should bring it.
b. #Some guitarist should bring it.

Elbourne's deletion theory predicts that the NP complement of a pronoun can only be deleted if there is an identical NP elsewhere in the discourse. Thus NP deletion is licensed in (5.58a) and (5.59a) but not in (5.58b) and (5.59b). Resolution is more liberal: empty NP complements can in principle be resolved to any salient property. Elbourne argues that (5.58b) and (5.59b) show that this is too unconstrained.

But I am not convinced. There are many examples which do require the freedom provided by resolution. Some such examples were discussed in section 4.3. One of these is repeated below in (5.60), and three additional examples are given in (5.61)–(5.63). (5.61) resembles (5.59b), and has already been noted in the literature at several occasions (according to Geurts, 1999, p.74, it dates back to Lakoff and Ross (1972)). (5.62) is designed to resemble (5.58b), and (5.63) is a similar 'real-life' example, taken from a website called *The Real Keys to a Happy Marriage* which, crucially, does not contain any occurrence of the word *husband*.[11]

(5.60) If I get pregnant, I'll definitely keep it. (overheard in conversation)

(5.61) John became a guitarist because he thought that it was a beautiful instrument.

(5.62) Some men have been married for more than twenty years and still don't know what her favorite breakfast is.

(5.63) If you don't know what his favorite movie is, you should plan to find out and watch it with him at the earliest convenience.

I do not have a very precise account for why resolution works much better in examples like (5.60)–(5.63) than it does in (5.58b) and (5.59b). It may be relevant that the conversational purpose of (5.58b) and (5.59b), considered in isolation, is far from clear. Thus, to make sense of (5.58b) and (5.59b) it is most naturally assumed that these sentences are part of larger discourse segments, and that the pronouns they contain refer to entities that are discussed not only in (5.58b) and (5.59b), but rather in those containing discourse segments. The conversational purpose of (5.60)–(5.63), considered as stand-alone utterances, is much clearer.

[11] `http://balancepdx.com/article_detail.php?article_id=32`

There is much room here for an improved account, but I don't think that such an account should make all too black-and-white predictions. Resolution may be much easier in some cases than in others, but there is a large gray area, with many gradations. A deletion theory such as Elbourne's appears to be much too strict in this respect.[12]

Elbourne's second argument is based on the following example:

(5.64) In this town, every farmer who owns a donkey beats it, and the priest does too.

According to Elbourne (2005b, p.69), this sentence does not have a sloppy reading (which would say that the priest also beats the donkey he owns). The analysis of donkey pronouns proposed in section 5.2, which is essentially that of Cooper (1979), predicts that the donkey pronoun in the source clause can be resolved to the property of being a donkey owned by t_1. In the source clause, t_1 is bound by [every farmer who owns a donkey]. In the target clause, it may in principle be bound by [the priest], and this would give rise to the sloppy reading that Elbourne claims not to exist.

However, I think that sloppy readings should not be ruled out by the grammar in examples like (5.64). A sloppy reading is just somewhat implausible in this particular example. In many examples that are structurally analogous to (5.64), sloppy readings are readily available:

(5.65) In this town, every farmer who has a spare room rents it out to tourists, and the priest does too.

(5.66) Most men who own a car like to show off with it. But Peter doesn't.

This view has also been voiced by Maier (2006), who uses the following example to disprove Elbourne's claim:

(5.67) Every male farmer who owns a donkey beats it, but farmer Mary doesn't.

Again, a sloppy reading is readily available here, contrary to the predictions of Elbourne's NP-deletion theory.

Very much related to these examples are Wescoat's and Hardt's examples discussed in section 4.4, repeated here:

---

[12]Jeroen Groenendijk pointed out to me that the pairs in (5.58) and (5.59) may in fact not be the right pairs to consider. The problem is that these pairs are not really *minimal.* Take the pair in (5.59). A much more minimal pair is the one in (i) below:

(i) a. #Some guitar player should bring it.
b. #Some guitarist should bring it.

Contrary to what the NP-deletion theory predicts, there is not much of a difference between (ia) and (ib), even though the former contains the NP [guitar] and the latter does not.

(5.68) The police officer who arrested John insulted him, and the one who arrested Bill did, too.

(5.69) If Harry has trouble at school, I will help him. But if John has trouble at school, I won't.

Elbourne (2005b, p.89–91) notes that his NP-deletion account of donkey pronouns wrongly prohibits sloppy readings for these examples. In reaction to this problem, he observes that sloppy readings do not only arise with pronominal NP-deletion, but also with other kinds of NP-deletion. For example, sloppy readings are available in:

(5.70) The police officer who arrested some murderers insulted at least three, and the police officer who arrested some burglars did too.

Thus, Elbourne argues, the fact that examples like (5.68) and (5.69) allow for sloppy readings does not show that his account of pronouns is wrong, but rather that such an account must rely on a theory of NP-deletion that is flexible enough to license the sloppy readings in question. However, Elbourne does not provide such a theory of NP-deletion. In fact, throughout his book he assumes a theory of NP-deletion which is based on LF identity, and the very point he wants to make when discussing examples like (5.58b), (5.59b), and (5.64) is that such a strict identity constraint on NP-deletion is necessary. How, then, would it be possible to build in the flexibility that is apparently required to account for the sloppy readings in (5.68) and (5.69)?

As mentioned in section 5.2, Tomioka (1999) already observed that an analysis of donkey pronouns à la Cooper (the one adopted here) straightforwardly explains the sloppy readings in (5.68) and (5.69). Elbourne launches an argument against such an analysis, based on the observation that the following variant of (5.68) does not have a sloppy reading:

(5.71) Every police officer who arrested a murderer insulted him,
and Officer Jones did too.

I agree that a sloppy reading is highly inaccessible in this case, but I do not think this should be explained on grammatical grounds. In fact, there is a very plausible pragmatic explanation. Namely, on a sloppy reading, the second clause of (5.71) would be completely *redundant*. It would convey information that is already conveyed by the first clause. If the example is slightly changed to avoid this redundancy, the sloppy reading reappears:

(5.72) Almost every police officer who arrested a murderer insulted him,
but Officer Jones didn't.

(5.73) Every police officer who arrested a murderer insulted him.
Even Officer Jones did.

Thus, Elbourne's arguments in favor of NP-deletion have all been countered. The flexibility provided by resolution appears to be necessary in general, although it may be constrained in certain specific cases by pragmatic factors. Elbourne's NP-deletion alternative is too strict, and there does not seem to be a straightforward way to add the necessary flexibility to it.

Complementary Theories. Of course, anaphoric mechanisms interact with many other linguistic mechanisms. Therefore, it should be expected that certain phenomena involving anaphora cannot be explained purely in terms of a theory of anaphora. Instead, a theory of anaphora must often interact with theories of other linguistic mechanisms in order to accomplish satisfactory explanations.

One important mechanism that interacts with anaphora resolution was already discussed in section 1.8, namely the encoding (and decoding) of *information structure*. The fact that Focus Match affects the resolution of VP anaphora should be seen as one particular consequence of this interaction. It is to be expected that there are many more such consequences, but these have, as far as I know, not yet been studied in much detail.

Another mechanism that interacts with anaphora resolution is the establishment of discourse coherence. This point has forcefully been made by Hobbs (1979), Prüst *et al.* (1994), Asher *et al.* (2001), and Kehler (2002), among others. To illustrate, I will consider some ways in which this interaction affects the resolution of VP anaphora, as described by Kehler (2002).

The first observation Kehler focuses on is that VP anaphora may exhibit a so-called *voice mismatch*. Sometimes, the target clause is in the passive voice, while the source clause is in the active voice, or the other way around. Kehler considers the following examples:

(5.74) In March, four fireworks manufacturers asked that the decision be revised, and on Monday the ICC did.
(from an official document originally cited by Dalrymple (1991))

(5.75) This problem was to have been looked into, but obviously nobody did.
(Vincent della Pietra, in conversation)

(5.76) Of course this theory could be expressed using DRSs, but for the sake of simplicity we have chosen not to.
(from text of Lascarides and Ahser (1993))

(5.77) Actually I have implemented the system with a manager, but it doesn't have to be.
(Steven Ketchpel, in conversation)

(5.78) Just to set the record straight, Steve asked me to send the set by courier through my company insured, and it was.
(posting on the internet)

Similar examples can be found in (Dalrymple *et al.*, 1991) and (Hardt, 1993). The fact that VP ellipsis allows voice mismatches has been used by Dalrymple *et al.*, Hardt, and others to argue against the idea that VP ellipsis consists in PF deletion under a syntactic identity constraint (à la Sag, 1976), or in copying syntactic material at LF (à la Williams, 1977). Rather, they argue the resolution of VP ellipsis involves the recovery of *semantic* material.

However, this argument is problematic, because there are also cases of VP ellipsis that do not allow voice mismatches. Kehler gives the following examples:

(5.79) #This problem was looked into by John, and Bill did too.

(5.80) #This theory was expressed using SDRSs by Smith, and Jones did too.

(5.81) #John implemented the system with a manager, but it wasn't by Fred.

Such examples could be used to argue exactly the opposite of what Dalrymple *et al.*, Hardt, and other proponents of a semantic approach argued, namely that the syntactic structure of the source clause *is* relevant for VP ellipsis.

Kehler shows us a way out of this impasse. He observes that there is a crucial difference between examples (5.74)–(5.78) on the one hand and examples (5.79)–(5.81) on the other. Namely, the kind of discourse relation between the source and target clauses in (5.74)–(5.78) is fundamentally different from the kind of discourse relation between the source and target clauses in (5.79)–(5.81). The clauses in (5.74)–(5.78) stand in a *Cause-Effect* relation, while the clauses in (5.79)–(5.81) stand in a *Resemblance* relation. Kehler argues on independent grounds that the establishment of Cause-Effect relations does not involve the reconstruction of syntactic material, while the establishment of Resemblance relations does. Thus, it is for the purpose of establishing discourse coherence (rather than for the purpose of resolving VP anaphora) that syntactic material must be reconstructed in (5.79)–(5.81) (and not in (5.74)–(5.78)). This explains why the voice mismatches in (5.79)–(5.81) are problematic, while the ones in (5.74)–(5.78) are not.

Similar observations can be made concerning VP anaphora with nominal antecedents. We have already seen some examples of this phenomenon in section 4.4. Kehler considers the following examples:

(5.82) This letter deserves a response, but before you do, ...
(attributed to Gregory Ward)

(5.83) Today there is little or no official harassment of lesbians and gays by the national government, although autonomous governments might.
(Hardt, 1993)

However, Kehler observes that VP anaphora with nominal antecedents is not always possible:

(5.84) #This letter provoked a response from Bush, and Clinton did too.

(5.85) #There is unofficial harassment of lesbians and gays by the American government, and the Canadian government does too.

Again, the contrast can be explained in terms of discourse coherence establishment. The clauses in (5.82)–(5.83) stand in a Cause-Effect relation, and the establishment of such a relation does not involve the reconstruction of syntactic material. Thus, the VP anaphora may be resolved to nominal antecedents. The clauses in (5.84)–(5.85), however, stand in a Resemblance relation, and the establishment of such a relation does involve the reconstruction of syntactic material. This is why the VP anaphora in these sentences cannot be resolved felicitously to the nominal antecedents.

Yet another manifestation of the interaction between VP anaphora resolution and discourse coherence establishment surfaces when the antecedent VP contains a self-pronoun. For example, a reflexive interpretation is forced in the target clause in examples (5.86)–(5.87), but not in examples (5.88)–(5.89).

(5.86) John defended himself, and Bob did too.

(5.87) Fred voted for himself, and Gary did too.

(5.88) John defended himself, because his lawyer couldn't.

(5.89) Fred voted for himself, even though no one else did.

The clauses in (5.86)–(5.87) stand in a Resemblance relation. The establishment of a Resemblance relation involves reconstruction of syntactic material. Hence, the self-pronoun is reconstructed and forces a reflexive interpretation in the target clause. The clauses in (5.88)–(5.89) stand in a Cause-Effect relation, the establishment of which does not involve reconstructing syntactic material. Therefore, a non-reflexive interpretation is allowed in the target clause.

Incidentally, Kehler (2002, p.58) notes that there are certain borderline cases. For example:

(5.90) The alleged murderer defended himself, and his lawyer did too.

(5.91) Bush voted for himself, and his campaign manager did too.

Kehler reports that many of his informants find a non-reflexive interpretation of the source clause in these examples at least marginally acceptable, although a majority of them report that these sentences are not completely natural.

This seems to be the same kind of judgment that my informants reported when faced with examples like (5.92) and (5.93) (see section 5.5):

(5.92) Only Max himself voted for him.

(5.93) I know what John and Mary have in common.
John hates Mary and Mary hates her too.

Kehler's assessment of (5.90) and (5.91) is indeed analogous to my assessment of (5.92) and (5.93). The fact that the elided VPs in (5.90) and (5.91) must be reconstructed in the process of discourse coherence establishment forces a reflexive interpretation of the target clause. But other factors (in this case world knowledge) strongly suggest that such a reflexive interpretation is not intended. As a result, informants generally feel that a non-reflexive interpretation is intended, even though it is not properly expressed.

It is to be expected that there is a variety of interactions between anaphora resolution and discourse coherence establishment, as well as other mechanisms, that remain to be explored. Such explorations, however, are left for future work.

Auxiliaries as proforms or pronouns as determiners? The central ideas defended in this paper are (i) that NP anaphora and VP anaphora should be analyzed in a unified manner, and (ii) that the interpretation of anaphora primarily involves resolution. I have proposed a particular implementation of these ideas, but alternative implementations are possible of course. The most important feature of the implementation proposed here is that it assimilates the case of pronominal anaphora to the case of non-pronominal NP anaphora and VP anaphora by assuming that pronouns are determiners with an empty NP complement, and that it is really this empty NP complement whose meaning is contextually determined.

The alternative is to proceed the other way around, namely to assimilate the case of VP anaphora to the case of pronouns. This would mean that neither pronouns nor auxiliaries that figure in VP ellipsis have empty NP/VP complements. Rather, the pronouns/auxiliaries themselves are resolved. Such an alternative unified (and resolution based) analysis of pronouns and VP ellipsis has been proposed by Hardt (1999).

One reason for assimilating pronouns to non-pronominal anaphora, rather than the other way around, is that much work in syntax and typology supports the idea that pronouns are definite articles, usually with empty NP complements. One relevant observation, which dates back to Postal (1966), is that English pronouns actually have overt NP-complements in some constructions:

(5.94) a. we linguists
b. you troops
c. them guys (dialect)

A comprehensive argument, which involves data from many languages other than English, can be found in (Lyons, 1999).

A second reason to treat pronouns as determiners, rather than auxiliaries as proforms, is that this allows for a unified analysis not only of pronouns and VP ellipsis, but also of non-pronominal NP anaphora. To the best of my knowledge, a proform theory of non-pronominal NP anaphora has not been proposed yet, and I find it hard to imagine one.

Sloppy readings: binding or reinterpretation? Sloppy readings arise in several constructions. In the first part of this dissertation, we concentrated on focus constructions such as (5.95) and (5.96), and on elliptical constructions such as (5.97).

(5.95) MAX called his mother.

(5.96) Only MAX called his mother.

(5.97) Max called his mother. Bob did too.

In the second part, we saw another construction in which sloppy readings arise, namely so-called paycheck sentences:

(5.98) Max spent his paycheck. Bob saved it.

Notice that (5.98) is very similar to (5.97): in both cases the sloppy reading arises because an anaphoric expression is resolved to an antecedent which itself contains another anaphoric expression. In (5.97), the empty VP in the target clause is resolved to [called his mother], which contains the pronoun [his]. In (5.98), the empty NP in the target clause is resolved to [his paycheck], which contains the pronoun [his]. For ease of reference, let me call such constructions *embedded anaphora constructions*.

There are essentially two ways to account for sloppy readings in embedded anaphora constructions. The one originally suggested by Keenan (1971) and preserved in the present proposal assumes that the embedded anaphoric expression in the source clause can be interpreted as a variable. This variable, then, is bound by one element in the source clause and by another in the target clause.

The alternative is to assume that the embedded anaphoric expression in the source clause may be *reinterpreted* in the target clause. Such an approach is pursued by Hardt (1999) and by Schlenker (2005).[13]

My main reason to account for sloppy readings in terms of binding and not in terms of reinterpretation is that sloppy readings do not only arise in embedded anaphora constructions. The focus constructions in (5.95) and (5.96), for example, are not embedded anaphora constructions, but do exhibit sloppy readings. As we have seen, a theory in which anaphoric expressions may be interpreted as bound variables immediately accounts for this fact. A reinterpretation theory of sloppy readings does not.

Another problem with reinterpretation accounts of sloppy readings is that they are, in principle, too weak to rule out certain impossible readings of embedded anaphora constructions. For example, they do not rule out mixed readings in

[13]The theories proposed by Hardt and Schlenker are cast in a *dynamic* framework. But a reinterpretation account of sloppy readings could just as well be implemented in a static framework. And vice versa, a bound variable account of sloppy readings could just as well be implemented in a dynamic framework.

cascaded ellipsis. I argued in section 5.2 that such readings should generally be ruled out (see, in particular, example (5.26)). A related problematic prediction that is made by reinterpretation theories is that the second sentence in (5.99) has readings such as (5.100) and (5.101):

(5.99) John didn't come to work. Bill called his boss, but Peter didn't.

(5.100) Bill called John's boss, but Peter didn't call Peter's boss.

(5.101) Bill called Bill's boss, but Peter didn't call John's boss.

It may be possible of course to devise additional machinery to rule out such readings. But if sloppy readings are accounted for in terms of binding rather than reinterpretation, no such additional machinery is needed. The right predictions are automatically generated.

Sloppy readings: additional support for a unified theory of anaphora. In (5.95)–(5.98), a sloppy reading arises because a pronoun is interpreted as a bound variable. If NP and VP anaphora are essentially the same, then we should also expect to find constructions in which a sloppy reading arises because an elided VP is interpreted as a bound variable. Such constructions do indeed exist. Kratzer (1991) discusses example (5.102), a VP analogue of (5.95):

(5.102) I only went to TANGLEWOOD after you did.

Schwarz (2000) considers example (5.103), a VP analogue of (5.97):

(5.103) When John had to cook he didn't want to.
When he had to clean he didn't either.

Hardt (1999) considers a similar example, originally due to Carl Pollard:

(5.104) I'll help you if you want me to.
I'll kiss you even if you don't.

Hardt also provides a VP analogue of (5.98):

(5.105) When Harry drinks, I always conceal my belief that he shouldn't.
When he gambles, I can't conceal it.

Schwarz (2000) shows that these cases are satisfactorily accounted for by assuming that VPs can be raised, just like DPs, and that VP anaphors can be bound by (in my terminology: can be resolved to the trace of) a raised VP.

I take this as further support for the view that NP and VP anaphora should be analyzed in a unified way (see also Charlow, 2008).

Pointers to collaborative work. Finally, I would like to point to two projects that I have been working on in collaboration with others, and which bear some relation to the work presented here.

First, (Nesson, Roelofsen, and Grosz, 2008) investigates the mechanisms that underly the generation and interpretation of anaphoric referring expressions in discourse. In particular, we focus on certain phenomena that have previously been considered in the light of Centering Theory (Grosz *et al.*, 1995). Very much in the spirit of this dissertation, we develop a theory which is based on very general principles about human behavior in communication. The theory explains why the main generalization embodied by Centering Theory is generally very robust, but also why it systematically fails in certain cases.

A second piece of related joint work is (Roelofsen and Aloni, 2008). This paper is concerned with a completely different empirical domain, namely that of *concealed questions*. But there is an interesting connection: the mechanism that has been argued here to play a crucial role in the interpretation of NP and VP anaphora, namely the contextual resolution of a property that is not overtly expressed, turns out to play an essential role in the interpretation of concealed questions as well.

5.7 Summary

In this chapter, a unified analysis of NP and VP anaphora has been proposed. All the issues that were raised in previous chapters have been resolved. In particular, the stipulative Semantic Identity condition on VP ellipsis has been eliminated; pronouns which could not be classified as either bound or coreferential, and instances of VP ellipsis which could not be dealt with in terms of VP Identity are no longer problematic; Dahl's puzzle has received a refined treatment; and the disputed Condition B effects have been dealt with in a satisfactory way. It has been argued that the proposed account improves on existing proposals, in particular the deletion-based account of Elbourne (2005b) and the resolution-based account of Hardt (1999). Moreover, the proposed theory neatly ties in with theories of information structure and discourse coherence establishment.

Appendix A

Resolution and C-command

I mentioned in section 2.1.2 that Strong Crossover effects are sometimes considered to be a special case of so-called Condition C effects (cf. Chomsky, 1981). Stated in our present terminology, the generalization embodied by Condition C is that traces, names, and descriptions cannot be covalued with expressions that c-command them. This generalization applies to strong crossover configurations, such as those in (A.1), but also to other constructions, such as those in (A.2). In (A.1), Condition C predicts that the pronouns cannot be resolved as indicated because, as a result of this, the traces would be cobound with a c-commanding pronoun. In (A.2), Condition C predicts that the pronouns cannot be resolved as indicated because, as a result of this, the name *Max* would corefer with a c-commanding pronoun.

(A.1)	a.	[every man]^1 [he likes t_1]	# he $\rightarrow$ t_1
	b.	[who]^1 [does he like t_1]	# he $\rightarrow$ t_1
(A.2)	a.	He loves Max.	# he $\rightarrow$ Max
	b.	He called Max's mother.	# he $\rightarrow$ Max
	c.	He says that Mary called Max's mother.	# he $\rightarrow$ Max

A unified account of (A.1) and (A.2) is attractive of course, but there are many problems with the generalization embodied by Condition C. For example, coreference is possible in:

(A.3) Whom did the candidates themselves vote for?
Not surprisingly, John voted for John and Bill voted for Bill.

(A.4) I know what John and Mary have in common:
John hates Mary, and Mary hates Mary as well. (cf. Evans, 1980)

(A.5) Only Max voted for Max. (cf. Reinhart, 1983)

(A.6) I think that this is exactly what happened:
Peter forced Tom to call Peter's girlfriend. (cf. Schlenker, 2004)

(A.7) He didn't give her a diamond ring because,
although he's madly in love with her,
Walter's just not ready to tie the knot. (McCray, 1980)

Furthermore, as various authors have observed, and as mentioned in section 4.1, the c-command restriction that is operative in constructions like (A.1) and (A.2) appears to be of a rather general nature. In particular, as illustrated by the following paradigm from (Wasow, 1972, p.81), a c-command restriction also seems to apply to VP anaphora:

(A.8) a. John tried LSD after Bill did Δ — $\Delta \rightarrow$ tried LSD
b. After Bill tried LSD, John did Δ — $\Delta \rightarrow$ tried LSD
c. After Bill did Δ, John tried LSD — $\Delta \rightarrow$ tried LSD
d. John did Δ after Bill tried LSD — # $\Delta \rightarrow$ tried LSD

Condition C does not capture this pattern, and thus seems to miss a significant generalization. As an alternative, I propose that (A.1), (A.2) and (A.8) are accounted for by the following constraint on resolution:[1]

A.1. Definition. [C-command Constraint on Resolution (CCR)]
An anaphoric element cannot be resolved to a constituent that it c-commands.

The CCR at once accounts for (A.1), (A.2) and (A.8), and avoids the counterexamples in (A.3)–(A.7). Notice that the constraint could not have been formulated in terms of binding, coreference, covaluation, codetermination, or any other derivative notion of anaphoric relatedness. Also, it could not have applied at once to (A.1), (A.2) and (A.8) if pronominal anaphora and VP ellipsis were not treated in a uniform way. Thus, we have further evidence here for the central ideas defended in chapter 5: (i) anaphora should be treated in a unified way, and (ii) anaphora are resolved.

I take the CCR to be a cognitive processing constraint. It may be language specific, but it may also be more general and affect other cognitive processes as well. A proper investigation of this question is left for future work.

Finally, I should remark that the CCR does not account for all alleged Condition C effects. And this is as it should be, I think. Here is a case in point:

(A.9) Max called Max's mother.

Condition C rules out coreference between the two occurrences of *Max*. The CCR doesn't. I think the latter prediction is correct, a coreferential interpretation of (A.9) is very well possible. It is true that the repetition of names in (A.9) is not entirely natural. But that, I think, is another issue, which is, for one thing,

[1] Wasow (1972) and Evans (1980) have proposed similar constraints. A detailed comparison is beyond the scope of this appendix.

independent of c-command. For instance, repetition of names is equally marked in constructions like (A.10) and (A.11).

(A.10) Max's mother called Max.

(A.11) Max called his mother.
Max invited her for dinner.

Moreover, speakers sometimes actually have good reasons to use a repeated name or description rather than a pronoun. In fact, this is why (A.3)–(A.6) are perceived not nearly as marked as (A.9)–(A.11).

A general theory of the use of pronouns versus names and descriptions can be found in (Nesson, Roelofsen, and Grosz, 2008). I think that this theory provides a satisfactory explanation of alleged Condition C effects which involve repeated names or descriptions (and are therefore beyond the reach of the CCR). A detailed defense of this claim, however, must be deferred to future work.

References

Andrews, P. (1986). *An Introduction to Mathematical Logic and Type Theory: to Truth Through Proof*. Academic Press, Orlando.

Asher, N., Hardt, D., and Busquets, J. (2001). Discourse parallelism, ellipsis, and ambiguity. *Journal of Philosophical Logic*, **18**(1), 1–25.

Baker, C. (1995). Contrast, discourse prominence, and intensification, with special reference to locally free reflexives in british english. *Language*, **71**, 63–102.

Beaver, D. (1997). Presupposition. In J. van Benthem and A. ter Meulen, editors, *The Handbook of Logic and Language*, pages 939–1008. Elsevier.

Bolinger, D. (1977). Pronouns and repeated nouns. Technical report, Indiana University Linguistics Club.

Burge, T. (1973). Reference and proper names. *Journal of Philosophy*, **70**, 425–439.

Büring, D. (2005a). *Binding Theory*. Cambridge University Press.

Büring, D. (2005b). Bound to bind. *Linguistic Inquiry*, **36**(2), 259–274.

Büring, D. (2007). Intonation, semantics and information structure. In G. Ramchand and C. Reiss, editors, *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press.

Carden, G. and Stewart, W. (1988). Binding theory, bioprogram, and creolization: Evidence from haitian creole. *Journal of Pidgin and Creole Languages*, **3**, 1–67.

Charlow, S. (2008). Free and bound pro-verbs: a unified treatment of anaphora. In T. Friedman and S. Ito, editors, *Semantics and Linguistic Theory (SALT 18)*.

Chien, Y. and Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, **1**, 225–295.

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Mouton de Gruyter.

Conroy, A., Takahashi, E., Lidz, J., and Phillips, C. (2007). Equal treatment for all antecedents: how children succeed with Principle B. *University of Maryland. Submitted for publication*.

Cooper, R. (1979). The interpretation of pronouns. In F. Heny and H. Schnelle, editors, *Syntax and Semantics 10: Selections from the Third Groningen Round Table*, pages 61–92. Academic Press, New York.

Dahl, O. (1973). On so-called sloppy identity. *Synthese*, **26**, 81–112.

Dahl, O. (1974). How to open a sentence: abstraction in natural language. In *Logical Grammar Reports 12*. University of Gothenburg.

Dalrymple, M. (1991). Against reconstruction in ellipsis. Technical report, SSL-91-114, Xerox.

Dalrymple, M., Shieber, S., and Pereira, F. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy*, **14**, 399–452.

Davis, W. (2005). Implicature. In *Stanford Encyclopedia of Philosophy*.

de Mulder, H., Lentz, T., Nilsen, O., and Zondervan, A. (2008). *Theoretical Validity and Psychological Reality*. Benjamins, Amsterdam.

Eckardt, R. (2001). Reanalyzing *selbst*. *Natural Language Semantics*, **9**, 371–412.

Eckardt, R. (2006). *Meaning Change in Grammaticalization*. Oxford University Press.

Elbourne, P. (2005a). On the acquisition of Principle B. *Linguistic Inquiry*, **36**, 333–365.

Elbourne, P. (2005b). *Situations and Individuals*. MIT Press.

Evans, G. (1980). Pronouns. *Linguistic Inquiry*, **11**, 337–362.

Faltz, L. (1985). *Reflexivization: a Sudy in Universal Syntax*. Foris, Dordrecht.

Fiengo, R. and May, R. (1994). *Indices and Identity*. MIT Press, Cambridge, Massachusetts.

Fox, D. (1998). Locality in variable binding. In P. Barbosa, D. Fox, P. Hagstom, M. McGinnis, and D. Pesetsky, editors, *Is the Best Good Enough?* MIT Press, Cambridge, Massachusetts. Also included in Fox (1999a).

Fox, D. (1999a). *Economy and Semantic Interpretation*. MIT Press, Cambridge, Massachusetts.

Fox, D. (1999b). Focus, parallelism and accommodation. In *Semantics and Linguistic Theory (SALT 9)*.

Gallin, D. (1975). *Intensional and Higher-Order Modal Logic*. North-Holland, Amsterdam.

Gamut, L. (1991). *Language, Logic and Meaning*. Chicago University Press.

Gast, V. (2006). *The Grammar of Identity: Intensifiers and Reflexives in Germanic Languages*. Routledge, Lodon.

Geach, P. (1962). *Reference and Generality*. Cornell University Press.

Geurts, B. (1999). *Presuppositions and Pronouns*. Elsevier, Oxford.

Geurts, B. (2008). *Piggyback anaphora: accessibility, binding, and bridging*. Manuscript, Nijmegen University.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan, editors, *Syntax and Semantics*, volume 3, pages 41–58. Academic Press, New York.

Grodzinsky, Y. (2007). *Coreference and self-ascription*. Manuscript, McGill University.

Grodzinsky, Y. and Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry*, **24**, 69–102.

Groenendijk, J. (2007). *Alternative Logical Semantics*. Manuscript, University of Amsterdam.

Grosz, B., Joshi, A., and Weinstein, S. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, **21**(2), 203–225.

Gundel, J., Hedberg, N., and Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, **69**(2), 274–307.

Hankamer, J. (1978). On the nontransformational derivation of some null VP anaphors. *Linguistic Inquiry*, **9**, 66–74.

Hankamer, J. and Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry*, **7**, 391–426.

Hardt, D. (1993). *Verb Phrase Ellipsis: Form, Meaning, and Processing*. Ph.D. thesis, University of Pennsylvania.

Hardt, D. (1999). Dynamic interpretation of verb phrase ellipsis. *Linguistics and Philosophy*, **22**, 187–221.

Hardt, D. (2005). Inference, ellipsis and deaccenting. In P. Dekker and M. Franke, editors, *Proceedings of the Fifteenth Amsterdam Colloquium*, pages 197–221.

Hardt, D. and Romero, M. (2004). Ellipsis and the structure of discourse. *Journal of Semantics*, **21**, 375–414.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts, Amherst.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy*, **13**, 137–177.

Heim, I. (1997). Predicates or formulas? Evidence from ellipsis. In A. Lawson, editor, *Proceedings of the Seventh Conference on Semantic and Linguistic Theory*, pages 197–221.

Heim, I. (1998). Anaphora and semantic interpretation: a reinterpretation of Reinhart's approach. In U. Sauerland and O. Percus, editors, *The Interpretive Tract. MIT working papers in Linguistics*. Written and distributed as a technical report at the University of Tübingen in 1992.

Heim, I. (2007). *Forks in the Road to Rule I*. Invited talk at NELS-38.

Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell Publishers.

Higginbotham, J. (1980). Logical form, binding, and nominals. *Linguistic Inquiry*, **14**, 395–420.

Hirschbühler, P. (1982). VP-deletion and across-the-board quantifier scope. In J. Pustejovsky and P. Sells, editors, *Proceedings of NELS-12*, pages 132–139.

Hobbs, J. (1979). Coherence and coreference. *Cognitive Science*, **3**, 67–90.

Hole, D. (1999). Agentive *selbst* in german. In *Proceedings of Sinn und Bedeutung VI*, pages 133–150.

Horn, L. (1984). Towards a new taxonomy of pragmatic inference: Q-based and R-based implicatures. In D. Schiffrin, editor, *Meaning, Form, and Use in Context*, pages 11–42. Georgetown University Press.

Huang, Y. (2000). *Anaphora: a Cross-linguistic Study*. Oxford University Press.

Ippolito, M. (2007). On the meaning of only. *Journal of Semantics*, pages 1–47.

Jacobson, P. (2007). *Direct compositionality and "unintepretability": a new look at "uninterpretable" features on pronouns*. Invited talk at the Amsterdam Colloquium.

Johnson, D. and Lappin, S. (1997). A critique of the minimalist program. *Linguistics and Philosophy*, **20**, 273–333.

Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*, pages 481–563. Oxford University Press. [Versions of this paper began circulating in 1971].

Karttunen, L. (1969). Pronouns and variables. In *Papers from the Fifth Regional Meeting of the Chicago Linguistics Society*, pages 108–115.

Keenan, E. (1971). Names, quantifiers, and a solution of the sloppy identity problem. *Papers in Linguistics*, **4**.

Keenan, E. (2002). Explaining the creation of reflexive pronouns in english. In D. Minkova and R. Stockwell, editors, *Studies in the History of the English Language: a Millenial Perspective*. Mouton de Gruyter, Berlin.

Kehler, A. (1993). A discourse copying algorithm for ellipsis and anaphora resolution. In *Proceedings of the Sixth European Chapter of the Association for Computational Linguistics*, pages 203–212.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. CSLI Publications, Stanford.

König, E. and Gast, V. (2006). Focused assertion of identity: a typology of intensifiers. *Linguistic Typology*, **10**, 223–276.

König, E. and Siemund, P. (2000). The development of complex reflexives and intensifiers in english. *Diachronica*, **17**, 39–84.

Koopman, H. and Sportiche, D. (1982). Variables and the bijection principle. *The Linguistic Review*, **2**, 139–160.

Kratzer, A. (1991). The representation of focus. In A. von Stechow and D. Wunderlich, editors, *Semantics: an International Handbook of Contemporary Research*. de Gruyter, Berlin.

Lakoff, G. and Ross, J. (1972). A note on anaphoric islands and causatives. *Linguistic Inquiry*, **3**(1), 121–125.

Larson, R. and Cho, S. (1999). Temporal adjectives and the structure of possessive DPs. In S. Bird, A. Carnie, J. Haugen, and P. Norquest, editors, *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL 18)*.

Larson, R. and Segal, G. (1995). *Knowledge of Meaning: an Introduction to Semantic Theory*. MIT Press.

Lascarides, A. and Ahser, N. (1993). Temporal interpretation, discourse relations, and common sense entailment. *Linguistics and Philosophy*, **16**, 437–493.

Lasnik, H. and Stowell, T. (1991). Weakest crossover. *Linguistic Inquiry*, **22**, 687–720.

Levinson, S. (2000). *Presumptive Meanings: the Theory of Generalized Conversational Implicature*. MIT Press, Cambridge, Massachusetts.

Lyons, C. (1999). *Definiteness*. Cambridge University Press.

Maier, E. (2006). *Review of Elbourne (2005b)*. `www.linguistlist.org`.

McCray, A. (1980). The semantics of backward anaphora. In J. Jensen, editor, *Proceedings of NELS 10*, pages 329–343.

Merchant, J. (2001). *The Syntax of Silence*. Oxford University Press.

Nesson, R., Roelofsen, F., and Grosz, B. (2008). Rational coordinated anaphora theory. Technical report, Harvard University. Submitted for publication, available at `ftp://ftp.deas.harvard.edu/techreports/tr-2008.html`.

Partee, B. (1978). Bound variables and other anaphors. In D. Waltz, editor, *Theoretical Issues in Natural Language Processing*, pages 79–85. Reprinted in Partee (2004).

Partee, B. (2004). *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee*. Blackwell, Oxford.

Pollard, C. and Sag, I. (1992). Anaphors in English and the scope of binding theory. *Linguistic Inquiry*, **23**, 261–303.

Postal, P. (1966). On so-called "pronouns" in English. In F. Dinneen, editor, *Report on the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, pages 177–206, Washington DC. Georgetown University Press.

Prüst, H., Scha, R., and van den Berg, M. (1994). Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy*, **17**, 261–327.

Pullum, G. (2000). *Hankamer was!* Jorge Hankamer Webfest, available at `http://ling.ucsc.edu/Jorge/`.

Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Croom Helm, London.

Reinhart, T. (2006). *Interface Strategies*. MIT Press.

Reinhart, T. and Reuland, E. (1993). Reflexivity. *Linguistic Enquiry*, **24**, 657–720.

Reuland, E. (2001). Primitives of binding. *Linguistic Inquiry*, **32**, 439–492.

Reuland, E. (2008). *Minimal versus not so minimal pronouns: feature transmission, feature deletion, and the role of economy in the language system*. Manuscript, Utrecht University. To appear in de Mulder *et al.* (2008).

Roelofsen, F. and Aloni, M. (2008). Perspectives on concealed questions. In T. Friedman and S. Ito, editors, *Semantics and Linguistic Theory (SALT 18)*.

Rooth, M. (1985). *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst.

Rooth, M. (1992a). Ellipsis redundancy and reduction redundancy. In S. Berman and A. Hestvik, editors, *Proceedings of the Stuttgart Workshop on Ellipsis*.

Rooth, M. (1992b). A theory of focus interpretation. *Natural Language Semantics*, **1**(1), 75–116.

Ross, J. (1967). *Constraints on Variables in Syntax*. Ph.D. thesis, MIT.

Ruys, E. (1994). A global economy analysis of weak crossover. In R. Bok-Bennema and C. Cremers, editors, *Linguistics in the Netherlands*, pages 223–234. John Benjamins, Amsterdam.

Sag, I. (1976). *Deletion and Logical Form*. Ph.D. thesis, MIT.

Sag, I. (2006). *What's LF got to do with it?*

Sag, I. and Hankamer, J. (1984). Toward a theory of anaphoric processing. *Linguistics and Philosophy*, **7**, 325–345.

Schachter, P. (1977). Does she or doesn't she? *Linguistic Inquiry*, **8**, 763–767.

Schlenker, P. (2004). Minimize restrictors! (Notes on definite descriptions, Condition C and epithets). In C. Bary, J. Huitink, and E. Maier, editors, *Proceedings of Sinn und Bedeutung*, pages 385–416.

Schlenker, P. (2005). Non-redundancy: towards a semantic reinterpretation of binding theory. *Natural Language Semantics*, **13**, 1–92.

Schwarz, B. (2000). *Topics in Ellipsis*. Ph.D. thesis, University of Massachusetts, Amherst.

Schwarzschild, R. (1999). Givenness, avoidF, and other constraints on the placement of acccent. *Natural Language Semantics*, **7**, 141–177.

Sells, P. (1987). Backwards anaphora and discourse structure: some considerations. Technical report, CSLI report 114.

Sem, H. (1994). VP-ellipsis and DRT. Technical report, Dynamic Interpretation of Natural Language, ESPRIT Basic Research Project Deliverable R2.2.B.

Stalnaker, R. C. (1978). Assertion. In P. Cole, editor, *Syntax and Semantics*, volume 9: Pragmatics, pages 315–332. Academic Press, New York.

Tancredi, C. (1992). *Deletion, de-accenting, and presupposition*. Ph.D. thesis, MIT.

Tic Douloureux, P. (1971). A note on one's privates. In A. Zwicky, P. Salus, R. Binnick, and A. Vanek, editors, *Studies out in left field: defamatory essays presented to James D. McCawley on the occasion of his 33rd or 34th birthday*, pages 45–51. Indiana University Linguistics Club.

Tomioka, S. (1997). *Focussing Effects and NP Interpretation in VP-ellipsis*. Ph.D. thesis, University of Massachusetts, Amherst.

Tomioka, S. (1999). A sloppy identity puzzle. *Natural Language Semantics*, **7**, 217–241.

van Gelderen, E. (2001). *A History of English Reflexive Pronouns*. Benjamins, Amsterdam.

van Rooij, R. and Schulz, K. (2007). Only: meaning and implicature. In M. Aloni, A. Butler, and P. Dekker, editors, *Questions in Dynamic Semantics*, pages 199–224.

Wasow, T. (1972). *Anaphoric Relations in English*. Ph.D. thesis, MIT.

Webber, B. (1978). *A Formal Approach to Discourse Anaphora*. Ph.D. thesis, Harvard University.

Wescoat, M. (1989). *Sloppy Readings with Embedded Antecedents*. Manuscript, Stanford University.

Williams, E. (1977). Discourse and logical form. *Linguistic Inquiry*, **8**, 101–139.

Samenvatting

Dit proefschrift gaat over pronominale en proverbale anaphorische expressies in het Engels.[2] Het eerste deel van het proefschrift evalueert, vergelijkt, en herziet een aantal van de meest prominente theorieën over pronominale anaphorische expressies die ontwikkeld zijn binnen de Generatieve Grammatica (Reinhart, 1983; Heim, 1998; Fox, 1999a; Büring, 2005b; Reinhart, 2006). Er wordt aangetoond dat geen van deze theorieën alle relevante feiten verklaart. De theorieën van Reinhart (2006) en Fox (1999a) worden herzien, gedeeltelijk voortbouwend op werk van Ruys (1994). De empirische problemen worden op die manier opgelost, terwijl de fundamentele aannames van de originele theorieën in tact worden gelaten.

Echter, het tweede deel van het proefschrift werpt een aantal bezwaren op tegen deze aannames. Uiteindelijk wordt er een alternatieve theorie voorgesteld, die ervan uitgaat dat de betekenis van anaphorische expressies altijd contextueel bepaald wordt, en niet in sommige gevallen syntactisch gecodeerd is, zoals Reinhart en vele anderen veronderstellen. Een geünificeerde analyze van pronominale en proverbale anaphorische expressies wordt uiteengezet, en een nieuwe analyze van cascaded ellipsis, de Dahl puzzle, en Conditie B effecten wordt beschreven. Er wordt beargumenteerd dat de theorie vooruitgang boekt ten opzichte van andere recente voorstellen (cf. Hardt, 1999; Elbourne, 2005b; Schlenker, 2005) en dat zij goed te combineren is met theorieën over informatie structuur en discourse coherentie (cf. Rooth, 1992a; Schwarzschild, 1999; Kehler, 2002).

[2]Deze korte samenvatting is bedoeld voor specialisten in het vak. Het voorwoord bevat een meer toegankelijke introductie, en elk afzonderlijk hoofdstuk is voorzien van een gedetailleerde samenvatting.

Abstract

This dissertation is about pronominal and verb phrase anaphora in English.[3] The first part of the dissertation evaluates, compares, and refines some of the most prominent theories of pronominal anaphora that have been developed within the framework of Generative Grammar (Reinhart, 1983; Heim, 1998; Fox, 1999a; Büring, 2005b; Reinhart, 2006). It is pointed out that none of these theories alone accounts for all the relevant data in a satisfactory manner. The theories of Reinhart (2006) and Fox (1999a) are refined, partly drawing on work by Ruys (1994). These refinements overcome the empirical problems, while keeping the fundamental assumptions of the original theories intact.

The second part of the dissertation, however, raises some objections against these assumptions. Eventually, an alternative theory is proposed, whose main premise is that anaphora are always contextually resolved, i.e., their meaning is always contextually determined, and not sometimes syntactically encoded, as Reinhart and many others assume. A unified analysis of pronominal and verb phrase anaphora is presented. A novel account of cascaded ellipsis, of Dahl's puzzle and of Condition B effects is proposed. The theory is argued to improve on other recent proposals (cf. Hardt, 1999; Elbourne, 2005b; Schlenker, 2005) and is shown to tie in neatly with theories of information structure and discourse coherence establishment (cf. Rooth, 1992a; Schwarzschild, 1999; Kehler, 2002).

[3]This short abstract is intended for specialists in the field. The preface provides a more gentle introduction, and each individual chapter comes with a summary which provides somewhat more detail than can be given here.

Curriculum Vitae

Floris Roelofsen (1980) obtained a BSc degree in Applied Mathematics and an MSc degree in Computer Science at the University of Twente, and another MSc degree in Logic at the ILLC, University of Amsterdam (all cum laude). He also briefly studied film and documentary at the AKI Academy for Visual Arts in Enschede, and contemporary ballet at the Academy for Performing Arts in Amsterdam, and traveled around the world as a fashion model for several years.

Floris was a visiting scholar at the University of Otago, New Zealand, at the Institute for Scientific and Technological Research in Trento, Italy, at the Rey Juan Carlos University in Madrid, and at Harvard University. At Harvard, he also taught a graduate course on computational linguistics, together with Barbara Grosz. His publications cover a variety of topics, ranging from linguistics to artificial intelligence and logic.

Besides studying, Floris is very fond of sports. As a teenager, he played soccer for AZ, one of the top professional teams in the Netherlands. Later, he tried windsurfing, dancing, and finally became a dedicated cyclist. He made several trips on his bike, through New Zealand, Southern Europe, Northern California, and most recently through Patagonia.

Floris lives in Amsterdam with his girlfriend Ana, who is a PhD student in linguistics at Utrecht University, and his sister Eva, who is a musical artist.

Titles in the ILLC Dissertation Series:

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability

ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations

ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs

ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software

ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance

ILLC DS-2002-02: **Marco Aiello**
Spatial Reasoning: Theory and Practice

ILLC DS-2002-03: **Yuri Engelhardt**
The Language of Graphics

ILLC DS-2002-04: **Willem Klaas van Dam**
On Quantum Computation Theory

ILLC DS-2002-05: **Rosella Gennari**
Mapping Inferences: Constraint Propagation and Diamond Satisfaction

ILLC DS-2002-06: **Ivar Vermeulen**
A Logical Approach to Competition in Industries

ILLC DS-2003-01: **Barteld Kooi**
Knowledge, chance, and change

ILLC DS-2003-02: **Elisabeth Catherine Brouwer**
Imagining Metaphors: Cognitive Representation in Interpretation and Understanding

ILLC DS-2003-03: **Juan Heguiabehere**
Building Logic Toolboxes

ILLC DS-2003-04: **Christof Monz**
From Document Retrieval to Question Answering

ILLC DS-2004-01: **Hein Philipp Röhrig**
Quantum Query Complexity and Distributed Computing

ILLC DS-2004-02: **Sebastian Brand**
Rule-based Constraint Propagation: Theory and Applications

ILLC DS-2004-03: **Boudewijn de Bruin**
Explaining Games. On the Logic of Game Theoretic Explanations

ILLC DS-2005-01: **Balder David ten Cate**
Model theory for extended modal languages

ILLC DS-2005-02: **Willem-Jan van Hoeve**
Operations Research Techniques in Constraint Programming

ILLC DS-2005-03: **Rosja Mastop**
What can you do? Imperative mood in Semantic Theory

ILLC DS-2005-04: **Anna Pilatova**
A User's Guide to Proper names: Their Pragmatics and Semanics

ILLC DS-2005-05: **Sieuwert van Otterloo**
A Strategic Analysis of Multi-agent Protocols

ILLC DS-2006-01: **Troy Lee**
Kolmogorov complexity and formula size lower bounds

ILLC DS-2006-02: **Nick Bezhanishvili**
Lattices of intermediate and cylindric modal logics

ILLC DS-2006-03: **Clemens Kupke**
Finitary coalgebraic logics

ILLC DS-2006-04: **Robert Špalek**
Quantum Algorithms, Lower Bounds, and Time-Space Tradeoffs

ILLC DS-2006-05: **Aline Honingh**
The Origin and Well-Formedness of Tonal Pitch Structures

ILLC DS-2006-06: **Merlijn Sevenster**
Branches of imperfect information: logic, games, and computation

ILLC DS-2006-07: **Marie Nilsenova**
Rises and Falls. Studies in the Semantics and Pragmatics of Intonation

ILLC DS-2006-08: **Darko Sarenac**
Products of Topological Modal Logics

ILLC DS-2007-01: **Rudi Cilibrasi**
Statistical Inference Through Data Compression

ILLC DS-2007-02: **Neta Spiro**
What contributes to the perception of musical phrases in western classical music?

ILLC DS-2007-03: **Darrin Hindsill**
It's a Process and an Event: Perspectives in Event Semantics

ILLC DS-2007-04: **Katrin Schulz**
Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity, and Conditionals

ILLC DS-2007-05: **Yoav Seginer**
Learning Syntactic Structure

ILLC DS-2008-01: **Stephanie Wehner**
Cryptography in a Quantum World

ILLC DS-2008-02: **Fenrong Liu**
Changing for the Better: Preference Dynamics and Agent Diversity

ILLC DS-2008-03: **Olivier Roy**
Thinking before Acting: Intentions, Logic, Rational Choice

ILLC DS-2008-04: **Patrick Girard**
Modal Logic for Belief and Preference Change

ILLC DS-2008-05: **Erik Rietveld**
Unreflective Action: A Philosophical Contribution to Integrative Neuroscience

ILLC DS-2008-06: **Falk Unger**
Noise in Quantum and Classical Computation and Non-locality

ILLC DS-2008-07: **Steven de Rooij**
Minimum Description Length Model Selection: Problems and Extensions

ILLC DS-2008-08: **Fabrice Nauze**
Modality in Typological Perspective

ILLC DS-2008-09: **Floris Roelofsen**
Anaphora Resolved