BEYOND IDENTITY:

TOPICS IN PRONOMINAL AND RECIPROCAL ANAPHORA

Alexis Dimitriadis

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2000

Anthony Kroch, Supervisor of Dissertation	Maribel Romero, Supervisor of Dissertation
Donald A. Ringe, Graduate Group Chair	Robin Clark, Committee Member

Acknowledgements

Like many a dissertation author before me, I owe the greatest debt to my advisors; and since I have two of them, I am twice as grateful for their contributions. If I could acknowledge every point on which they have had a substantial impact, their names would appear on every page of this dissertation.

Tony Kroch has been my advisor throughout the long gestation of this dissertation. His experience and advice have helped me define my topic, understand how to approach it, and ask the right questions. Throughout my time at Penn, I have been guided by his sense for what is important and valuable in linguistics, what is questionable, and especially, what is valuable even though questionable. Were it not for his example I might long ago have despaired of ever making useful sense of something as complicated as human language.

Maribel Romero arrived at Penn just in time to guide me in navigating the subtleties of formal semantics. Her tutelage steered me toward a sense of perspective that I could not have gained from written materials alone, and she has always been happy to share her time and advice not only on matters of content, but also of style and presentation. There is much in this dissertation that bears her direct imprint.

Robin Clark, the third member of my committee, was always generous with his time and has directed my attention to numerous issues that would have otherwise remained beyond my grasp.

My dissertation has also benefited from suggestions by many others at Penn and elsewhere. I particularly wish to acknowledge the contributions of Ellen Prince, Uli Sauerland, Roger Schwarzschild, Yael Sharvit, Christine Brisson, and Gerhard Jaeger. I am also grateful to Amanda Seidl for being a good friend during the writing of this dissertation, and for proofreading drafts of it and actually enjoying it!

In sharing credit for the long process that led to the completion of this dissertation, I must begin by thanking my good friend Daniel Elbaum, who was the first to draw me to the magic of linguistics; and my teacher at Portland State University, Beatrice Oshika, whose expert guidance through the graduate application process was absolutely indispensable.

While at the University of Pennsylvania I have been privileged to work closely with many inspiring instructors. In addition to the members of my committee, I am especially grateful for the opportunity to study with Ellen Prince and Sabine Iatridou. My studies were also enriched by the presence of fellow linguists and friends, including Spyridoula Varlokosta, Heidi Harley, Brian Teaman, Nadia Biassou, Young-Suk Lee, Carmen Fought, Mike Schultz, Ümit Turan, Roumyana Izvorski, Miriam Meyerhoff, Amanda Seidl, Hikyoung Lee, Chung-hye Han, Susan Garrett, Patricia Chow, Alan Lee, Na-Rae Han, Clarissa Surek-Clark, and Elsi Kaiser.

My non-linguist friends in Philadelphia might not always understand what I was doing or why, but they helped me remember that there's more to life than linguistics. Among those I will remember most fondly are Diane Carle, Rich Garella, Cheryl Gilbert, Yannis Drambas, Panos Grekos, Marian Thomas, Margie Politzer, and Art Torrance. Most of them no longer live in Philadelphia, but I will always think of them when I think of this city as my home.

I also wish to thank my parents and family, who have always encouraged me in all my endeavors; and Amy, who is helping me work out what happens next.

ABSTRACT

BEYOND IDENTITY:

TOPICS IN PRONOMINAL AND RECIPROCAL ANAPHORA

Alexis Dimitriadis

Supervisors: Anthony Kroch and Maribel Romero

This dissertation examines problems in the semantics of reciprocals and pronouns bound by non-quantificational NPs, and investigates their semantic analysis as individual-valued functions rather than as simple variables. The main concern is the analysis of so-called "long-distance" reciprocals, in which the local antecedent of the reciprocal is a pronoun dependent on a higher antecedent:

(i) John and Mary think they like each other.

The standard "scopal" analysis of such constructions (Heim et al. 1991a, and many others) is to let the reciprocal be bound by a distributive operator that also binds the local antecedent of the reciprocal. It is shown here that because the binder of the reciprocal determines its range, the scopal analysis cannot account for new examples in which the local antecedent of the reciprocal is not bound by a coreferring, c-commanding antecedent. It is argued that the correct semantics for such constructions is always determined by the local antecedent of the reciprocal.

I propose that all dependent pronouns should be translated as functions in the manner of Engdahl (1986), and argue for an enriched representation that includes domains. The range of a reciprocal is then obtained by applying a maximality operator to the restricted function representing its local antecedent.

This analysis always interprets reciprocals locally, eliminating the need for a “scopal” treatment.

The Variable-Free Semantics of Jacobson (1999a), which I adopt in the final chapter, is particularly well-suited to this analysis since it allows direct access to the reference function of pronouns. It also makes possible a uniform treatment of examples involving a variety of local antecedent types, including “paycheck” pronouns and complex NPs that contain a dependent pronoun.

I discuss the interaction of the scope issue with other aspects of reciprocal semantics, including weak reciprocity, collective action, types of reciprocal relations, and exceptions (non-maximality).

Contents

Acknowledgements **ii**

Abstract **iv**

Contents **vi**

1 Introduction **1**

1.1 The big picture 1

1.2 The basic semantic framework 4

1.2.1 Notes on notation 5

1.3 The semantics of plurals 6

1.3.1 A menagerie of *part* relations 8

1.3.2 Plural predication 9

1.4 Types of distributivity and reciprocity 10

2 The Scopal Analysis of Reciprocals **14**

2.1 The Heim, Lasnik and May (1991a) account 15

2.1.1 A “puzzle of grain” 15

2.1.2 A “puzzle of scope” 17

2.1.3 A sketch of the Heim et al. analysis 18

2.1.4 Distributivity 19

2.1.4.1 The position of the distributor 23
2.1.5 Reciprocals 24
2.1.5.1 Reciprocals in subject position 28
2.1.6 "Long-distance" reciprocals 29
2.1.7 Some details of the analysis 31
2.1.7.1 The semantics of [*e other*] 31
2.1.7.2 The syntax of the trace *e* 32
2.1.7.3 Multiple reciprocals: "Absorption" 33
2.2 Problems, and a revised proposal (Heim et al. 1991b) 34
2.2.1 On *each*-movement 35
2.2.2 More on the "*each*-binding" account 38
2.2.2.1 The translation of distributors and "floated" *each* 40
2.2.2.2 On quantifier-raising the NP-distributor complex 42
2.3 Additional problems 44
2.3.1 On NP-adjoined distributors 45
2.3.1.1 C-command effects 46
2.3.1.2 VP conjunction effects 47
2.3.1.3 Adjunction to VP and the "relational" nature of distributivity 49
2.3.2 Distributing without c-command 51
2.3.2.1 Non-reciprocal examples 52
2.3.2.2 Dependent reciprocals without c-command 58
2.3.3 Chained multiple reciprocals 60
2.3.4 The determination of reciprocal range 65

3 Readings of dependent pronouns **67**

3.1 Dependent pronouns and binding 68

3.1.1 Towards a solution 72

3.1.2 More types of dependent pronouns 73

3.2 Kamp and Reyle (1993): Plurals in DRT 75

3.2.1 A quick introduction to DRT 75

3.2.2 Plurals and dependent pronouns 79

3.2.3 Plural-marked referents 82

3.2.4 Dependent plural NPs 84

3.3 Are dependent pronouns real? 88

3.3.1 The cumulative alternative 88

3.3.2 Context restrictors and strongest meaning 91

3.3.3 The importance of being a pronoun 95

3.3.4 The fixed reading 97

3.3.5 Final comments on the cumulative alternative 99

3.4 Functional relative clauses? 101

3.5 The functional analysis of dependent pronouns 104

3.6 Split dependent plurals 107

4 A non-scopal analysis of reciprocals **113**

4.1 A short review of the scopal analysis 114

4.2 Dependent anaphora without identity 116

4.3 The grammaticality of dependent reciprocals 118

4.4 Dependent pronouns as restricted functions 122

4.5 A nonscopal semantics for reciprocals 125

4.5.1 The combinatorics of the proposed analysis 128

4.5.2 Larger dependent NPs . . . 129
4.6 Split-dependent reciprocals . . . 132
4.7 Reciprocals and collective action . . . 135
4.7.1 Reciprocation without distributors . . . 138
4.8 The distinctness condition . . . 141
4.9 Conclusions . . . 144

5 Other treatments of reciprocals **147**
5.1 The non-scopal view: Williams (1991) . . . 148
5.1.1 Weak and strict distribution, and floated *each* . . . 150
5.1.2 The linking theory . . . 152
5.1.2.1 Theta links . . . 155
5.1.3 Dependent reciprocals . . . 157
5.1.4 Linking multiple reciprocals . . . 159
5.1.5 Some generalizations . . . 161
5.2 Cumulation and weak reciprocity: Sternefeld (1998) . . . 163
5.2.1 Reciprocals and weak reciprocity . . . 167
5.3 Covers for distribution: Schwarzschild (1996) . . . 169
5.3.1 Plurals and covers . . . 169
5.3.2 Schwarzschild's semantics for reciprocals . . . 174
5.3.3 Dependent reciprocals . . . 176
5.3.4 The benefits of the covers approach . . . 178
5.4 Putting it together: Dependent reciprocals and covers . . . 181
5.4.1 Chained reciprocals . . . 184

6 The Variable-Free Approach **186**
6.1 A short review: The range of dependent reciprocals . . . 186

6.1.1 Deriving the range argument 190
6.2 Variable-free semantics 192
6.2.1 Dependent pronouns in larger NPs 196
6.2.2 Paycheck pronouns 198
6.3 Reciprocals in the variable-free system 200
6.3.1 The remainder of the derivation 203
6.4 Using the range 205
6.4.1 Deriving the dependent reading 206
6.4.2 Paycheck pronouns 208
6.4.3 What about non-dependent reciprocals? 210
6.5 Conclusions 216

7 Summary and further directions **218**

Bibliography **222**

Chapter 1

Introduction

1.1 The big picture

The focal concern of this dissertation is the analysis of so-called "long-distance" reciprocals, such as the reading of sentence (1.1a) given in (b):

(1.1) a. Mary and John think they like each other.

b. Mary thinks "I like John", and John thinks "I like Mary".

In a simple reciprocal sentence like (1.2) the reciprocal has an antecedent, necessarily plural, that is interpreted distributively and is "local" in the sense of Binding Principle A. This antecedent provides the reciprocal's *range argument,* the set of elements that can appear in the object position of the reciprocal predicate.

(1.2) The children saw each other.

In "long-distance" reciprocals like (1.1a), however, the local antecedent is a bound pronoun that is translated as a singular variable; this leaves the reciprocal without the plural antecedent it requires. The standard solution is to have the reciprocal look for its antecedent outside the embedded clause. This is the analysis adopted by Heim, Lasnik, and May (1991a,b), who claim that the reciprocal in (1.1a) can be bound, non-locally, by a distributor adjoined to the matrix subject—hence the name, *long-distance* reciprocal.

This "scopal" analysis, which treats the reading in question as involving non-local raising or binding of the reciprocal, has been implicit or explicit in almost all treatments of this phenomenon, including those of Lebeaux (1983), Higginbotham (1983), Moltmann (1992), Sauerland (1994), Schwarzschild (1996), and Sternefeld (1998). The notable exception is Williams (1991), who argues that the reciprocal's range argument is determined via reference to its local antecedent only. In this dissertation I examine a range of "long-distance" reciprocal constructions, which I refer to more neutrally as *dependent reciprocals.* I conclude that the long-distance analysis is inadequate, and that the data point us toward an analysis somewhat in the spirit of Williams (1991). However, Williams's own system does not include a viable proposal on how to deal with the central formal puzzle of dependent reciprocals, the extraction of a plural range argument from a singular bound variable. Instead I develop an analysis that relies on translating pronouns as functions, in the fashion of Engdahl's (1986) treatment of "paycheck" pronouns, and derive the range argument of the reciprocal from the domain of the pronoun function representing the reciprocal's local antecedent.

Getting from here to there involves several steps and auxiliary topics. Chapter 2 provides a detailed critical look at the scopal analysis of reciprocals as proposed by Heim et al. (1991a,b), and the shortcomings of their approach. Chapter 3 shifts attention away from reciprocals: it examines dependent pronouns in the scope of distributors, and shows that pronouns apparently bound without c-command should be analyzed as paycheck pronouns in the style of Engdahl's (1986) functional adaptation of Cooper's (1979) treatment. Chapter 4 applies this analysis to dependent reciprocals without apparent c-command, and shows that even with the paycheck analysis, scopal treatments give the wrong semantics for dependent reciprocals. A revised, non-scopal treatment is then proposed which derives the range argument of the reciprocal from the domain of its local antecedent.

Chapter 5 reviews other analyses of reciprocals, including the non-scopal analysis of Williams (1991). Although recent scopal treatments have made great progress with respect to a number of other aspects of reciprocal interpretation, it is shown that they are uniformly unable to handle the full range of dependent reciprocal constructions raised in chapter 2. But the solution proposed in chapter 4 is not essentially dependent on the Heim et al. analysis. Section 5.4 shows how Schwarzschild's (1996) treatment of reciprocals can be revised along the lines proposed in chapter 4.

The analysis developed in chapter 4 has the shortcoming that the paycheck function is only part of the pronoun's translation (the other part being an open variable that serves as the function's argument), and hence is not directly accessible to the reciprocal. The proposed solution is a poor fit with examples that do *not* involve a paycheck pronoun. Chapter 6 explores Jacobson's (1999a, 1999b) Variable Free Semantics, in which pronouns are translated as functions without a saturating argument. In this system it is possible for the reciprocal to access the reference function of its antecedent directly, making possible a more straightforward and more powerful variant of the analysis developed in chapter 4.

A short final chapter summarizes the issues that have been addressed in this work, and the major questions, new or old, that it leaves open.

The chapters are intended to be more or less self-contained. Inevitably there is some overlap and repetition, although the summaries tend to get shorter each time around. The remainder of the present chapter gives a summary of the semantic framework I assume throughout. Section 1.2 briefly reviews the basics, while section 1.3 provides a synopsis of Link's (1983) semantics of plurals. Section 1.4 presents a short review of several issues pertaining to distributivity and reciprocity that will be relevant to the discussion.

1.2 The basic semantic framework

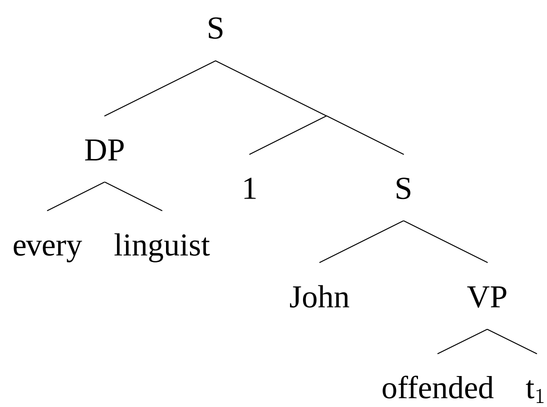
This section makes explicit some of the notational and theoretical assumptions behind the body of this dissertation. There ought to be no surprises here for anyone familiar with some formal semantics, Link's (1983) semantics of plurals, and the general Government and Binding or Minimalist approach to syntax. The reader familiar with these topics could skip this section altogether.

I assume a Montague-style semantic framework, which is for the most part represented informally. For example, assignment functions are suppressed, and only extensional translations are shown. In general, I follow the basic framework formulated by Heim and Kratzer (1998).

One aspect that merits particular mention is their handling of movement traces. Heim and Kratzer assume that a moved constituent leaves behind an indexed trace, and also causes the insertion, as a separate constituent, of a lambda abstractor just below its adjunction site. For example, quantifier-raising (QR) of *every linguist* in (1.3a) gives rise to the structure in (1.3b).

(1.3) a. John offended every linguist.

b.



The abstractor *1* combines with its sister via *Predicate Abstraction*, defined as follows:

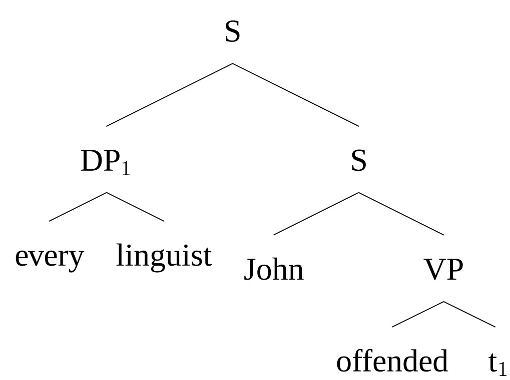
(1.4) **Predicate Abstraction Rule (PA)** (Heim and Kratzer 1998:p. 186)

Let α be a branching node with daughters β and γ, where β dominates only a numerical index i. Then, for any variable assignment g, $[\![\alpha]\!]^g = \lambda x\, [\![\gamma]\!]^{g^{x/i}}$.

In the above example, the lower S constituent is translated as the assignment-dependent proposition *offended*$($*John*$, g(1))$. The abstractor converts this into λx *offended*$($*John*$, x)$. The raised object will then immediately supply the argument x, in effect binding its movement trace t_1.

Syntacticians do not usually place the index of a moved constituent on a separate branch. Heim and Kratzer suggest that the more familiar structure (1.5), in which the index appears on the moved constituent, can be taken as an abbreviation for (1.3b).

(1.5)



1.2.1 Notes on notation

In various chapters I employ two different notations for the arguments of multi-place functions, depending on whether linguistic meaning or formal compositionality is of more immediate interest. In most of this dissertation, arguments are given in the usual English order in a single set of parentheses; for example, *visit(John, Paris)* means “John visits Paris.” It follows that the verb *visit* must be represented as $\lambda x\, \lambda y$ *visit*(y, x). In chapter 6, however, arguments are given in separate parentheses, with the thematically lower argument closest to the verb. (This is the notation used by Jacobson (1999a,b), whose work is central to that chapter). In this notation, our example would be written as *visit(Paris)(John)*, and the verb itself as $\lambda x\, \lambda y\, \textit{visit}(x)(y)$. This allows the result of supplying the lower arguments to be treated just like an atomic predicate, for example, after combining *visit* with *Paris* we get $\lambda y\, [\textit{visit(Paris)}](y)$. It is hoped that the benefits of using the most convenient notation in each chapter outweigh the potential for confusion due to the switch.

Quantifiers are often written with domains. Either $\forall x \in A\ P(x)$ or $\forall x(x \in A)\, P(x)$ are equivalent to $\forall x[x \in A \rightarrow P(x)]$.

I have been generally informal with notation, leaving out subscripts and diacritics where there is no ambiguity. Hopefully this will enhance rather than impede clarity of presentation.

1.3 The semantics of plurals

I assume the plural semantics of Link (1983). His system is well-known, and I will not attempt a complete exposition here. Rather, I sketch it via a simplified version that Schwarzschild (1992:p. 6ff) calls the "union theory" of plurals. It will hopefully enable the reader who is unfamiliar with Link's system to follow along without recourse to other sources. The union theory does not represent the part of Link's system that deals with mass terms; this will not be needed for our purposes.

I will refer to singular entities in our model as *atoms*. (They are also commonly called *individuals*, but the name has a different sense in Link's framework). The essential features of Link's system are, first, that singular and plural objects are of the same type, and second, that the result of combining, for example, an object consisting of three atoms with one consisting of two atoms is simply an object consisting of five atoms (or fewer, in the case of overlap), with no intermediate structure.[1]

Let A be the set of atoms that we wish to model, and let $D = \mathcal{P}(A)$ be the set of all subsets of A. The set D forms a well-known lattice under the relation of set inclusion. This is the *lattice of individuals* that models our singular and plural entities. Its elements are called *individuals*, regardless of whether they represent a single entity (atom) or a collection of several. (The bottom element of the lattice, corresponding to the empty set, is special and is not considered an individual). In this model, every element x of the set of atoms A is represented as the singleton set $\{x\}$, while larger individuals are represented as simple sets of atoms. There are no higher-order sets whose members are themselves sets.

On the lattice of individuals we speak of the *part* relation, which I write $\amalg$. An element x is part of y whenever $x \subseteq y$, viewed as sets. The *sum* of two individuals x and y, written as $x \oplus y$, is simply the lattice element corresponding to the set $x \cup y$. We also define the *meet* "$\wedge$" of two individuals, the individual consisting of their common parts; it is given by the intersection of the corresponding sets. If two individuals have no part in common their meet is the bottom lattice element, which I write as **0**.

This model represents atomic individuals as singleton sets, a property that many linguists find cumbersome. It is convenient to follow *Quine's innovation*, a modified set theory that identifies a singleton set containing an atomic individual with its (unique) element.[2] This makes it possible to write either $j \subseteq x$ or $j \in x$ if j is an atomic individual.

[1]See Schwarzschild (1996:pp. 1–16) for an informative exposition of the major alternative approaches to plurals.

[2]Quine's innovation is summarized by Schwarzschild (1992:p. 199). It can be visualized by letting each atomic individual be enclosed by an infinite number of set brackets, so that adding or subtracting a pair gives us the same set. If John is represented by the lattice element j "defined" as in (i), the relations in (ii) and (iii) hold.

(i) $j = \{\{\{\ldots\{\{\text{John}\}\}\ldots\}\}\}$
(ii) $j \in j$
(iii) $j = \{j\}$

It should be noted that in Quine's system a non-singleton set is not identified with the set containing it (which *is* a singleton set, of a different kind).

The system I sketched is just one possible model for a lattice of individuals. Link (1983) defines his theory of plurals axiomatically, starting with a "complete join semilattice" under the *part* relation and defining sums and meets on it. Such lattices can model domains in which there are no atomic elements, such as liquids or the material out of which objects are made. (Link's semantics, contrary to physical reality, pretends that the objects denoted by mass terms can be subdivided ad infinitum). If we restrict our attention to collections of individuals that are built up from atomic individuals, we obtain a sublattice that is isomorphic to the set model I have sketched. Kamp and Reyle (1993:p. 404) define a mapping from the lattice to the union model, and prove its validity.[3] I will follow the informal practice of treating the two representations as notational variants; Mary $\oplus$ John is the same as the set, or plural individual, $\{$Mary, John$\}$. By virtue of these shortcuts it is possible to be very informal in our discussion of individuals, speaking of sums, subsets and parts depending on which makes the most intuitive sense.

1.3.1 A menagerie of *part* relations

Link (1983) defines the basic *part* relation via the following lattice-theoretic definition:

(1.6) a is part of b iff $a \oplus b = b$

In the union representation of plurals this is equivalent to the condition $a \cup b = b$, which is true if and only if $a \subseteq b$. Note that this makes any individual a part of itself. In linguistic studies of plurals, this most general relation is commonly restricted by additional conditions: whether a is atomic, and whether a is a *proper* part of b, i.e., whether $a \neq b$. Link defines two part relation symbols, Π and $\cdot\Pi$, respectively representing the *part-of* and *atomic-part-of* relations. Unfortunately Heim et al. (1991a) use the same symbols for the

[3]Schwarzschild (1992:p. 18) also discusses this equivalence, for which he refers to Lasersohn's dissertation (Lasersohn 1988).

proper-part-of and *proper-atomic-part-of* relations, respectively. To minimize confusion, in this dissertation I use the symbols Π and ·Π in the sense of Heim et al. only, i.e., to mean *proper-part-of* and *proper-atomic-part-of*, respectively. For the unqualified *part-of* relation (Link's Π relation) I adopt the symbol ∐ . Finally, I write the *atomic-part-of* relation as $\in$ on the few occasions that it is needed.

1.3.2 Plural predication

Consider a predicate such as *lift the couch*, which describes an act that might be performed by a single person or by several persons acting collectively. It is true of an individual, say John $\oplus$ Mary, if and only if John and Mary together are the agents in an act of couch-lifting. If John and Mary together lifted the couch, and Bill lifted the couch by himself, then the extension of *lift the couch* is the set $\{$ *John* $\oplus$ *Mary, Bill* $\}$. Note that neither John nor Mary by themselves are in its extension. Other predicates, such as *child*, are only true of atomic individuals: John is a child and Mary is a child, but the individual John $\oplus$ Mary is not a child. Link calls such predicates *distributive;* elsewhere they have been referred to as *D-based* predicates.

From any predicate P we can construct another predicate $*P$, which is true of any individual that is the sum of individuals of which P is true. (Precisely: the extension of $*P$ is defined as the complete join semilattice generated by the extension of P). If P is the predicate *lift the couch* from the above example, the extension of $*P$ is the set $\{$ *John* $\oplus$ *Mary, Bill, John* $\oplus$ *Mary* $\oplus$ *Bill* $\}$. If Ann and Bob are children, *child* is true of Ann and of Bob but not of Ann $\oplus$ Bob, but **child* is true of Ann, Bob, and Ann $\oplus$ Bob. The predicate **child* corresponds to the denotation of the plural noun *children* (strictly, *child or children*).

If $*P$ is the plural form of a D-based predicate P, the following inference is licensed:

(1.7) $*P(a)\ \&\ b\ ∐\ a\ \&\ \mathrm{Atom}(b) \rightarrow P(b)$

For example, from *Ann and Bob are children* we infer *Bob is a child.* Such downward entailments are not licensed in general; for example, in the *lift the couch* example we may not distribute from *John* $\oplus$ *Mary* down to *Mary,* since Mary by herself did not lift the couch.

1.4 Types of distributivity and reciprocity

The issues discussed in this section are largely orthogonal to the main goals of this dissertation. Although I do not attempt to contribute to their resolution, they are integral to the study of reciprocals and arise at numerous occasions. Hence, in this section I collect some concepts and definitions for the reader's convenience. Most of them can be traced to the work of Langendoen (1978) and Scha (1984).

As a first approximation, sentences involving plural NPs can be translated using universal quantification over the atomic parts of these NPs. Translating a distributive plural sentence with universal quantification over the atomic parts of the subject gives the semantics of *strong distributivity.* For example, sentence (1.8a) is translated as in (b).

(1.8) a. The boys sleep.

b. $(\forall x \in \text{Boy}')\,\text{sleep}(x)$

This approach to the treatment of plurals is inadequate in (at least) three respects. The first type of problem is the possibility of collective action: For predicates that are not D-based, universal quantification results in truth conditions that are too strong. For example, in the couch-lifting situation of the previous section sentence (1.9a) should be regarded as true; but strong distributivity yields the proposition in (b), which is false since neither John nor Mary by themselves lifted the couch. (I use the symbol $\in$ to denote the relation *atomic-part-of*).

(1.9) a. John, Mary and Bill lifted the couch.

b. $(\forall x \in J \oplus M \oplus B)$ lift-couch(x)

We can derive the correct truth conditions for this sentence by giving it the semantics of *weak distributivity,* which can be defined as follows:[4]

(1.10) $(\forall x \in A)(\exists y \amalg A)\, x \in y \;\&\; P(y)$

This formula says that every atomic part x of the subject is *part* of some individual y, possibly plural, to which the predicate applies. (Since this definition does not exclude the possibility that $x = y$, it always holds when strong distributivity holds).

A second problem involves sentences with two or more plurals. Sentence (1.11) can be truthfully used even if it is not true that each boy rented all three movies. All that is required under this reading, which I will refer to as the *cumulative* reading, is that each boy rented one or more of the movies, and each movie was rented by at least one boy. The truth conditions of the cumulative reading are given by expression (1.12).[5]

(1.11) The boys rented *Babe, Citizen Kane,* and *Animal House.*

(1.12) $[(\forall x \in B)(\exists y \in M)\, xRy] \;\&\; [(\forall w \in M)(\exists z \in B)\, zRw]$

The final problem is the possibility of exceptions, or *non-maximality* as Brisson (1998) calls it. For example, we accept sentence (1.13) as true even if some villagers are awake.

(1.13) The villagers are asleep.

In contrast with the other two problems, we do not yet have a clear formal understanding of non-maximality. There are several ways to approach it. Roberts (1989) defines the

[4]This definition is consistent with, but not identical to, the usage of the term by Williams (1991). It should not be confused with the definition of "weak distributivity" given by Sternefeld (1998), which is actually the condition on two-place relations that I refer to as *cumulativity.*

[5]Sternefeld (1998) calls this condition *weak distributivity.*

operator ENOUGH, which allows a predicate to be considered true of a plural subject if it is true of a sufficiently large part of it. Kroch (1979) argues that non-maximality of distributive (but not collective) predicates is a pragmatic effect, and should be addressed by simply leaving the exceptions out of the domain of discourse. Brisson (1998) develops a more formalized version of Kroch's approach; her analysis, which is briefly presented in section 5.3.4, effectively excludes exceptions from the truth conditions of a predicate by means of *ill-fitting covers* based on the system of Schwarzschild (1996).

As discussed in depth by Langendoen (1978), we encounter similar issues in the semantics of reciprocals. Heim et al. (1991a) give reciprocals the semantics of *strong reciprocity*, shown in (1.14), which require that every atomic part of the subject of a reciprocal predicate be related to every other atomic part. This translation ignores the possibility of collective action, but even when this is not an issue, the conditions it imposes on the relation R are too strong. For example, the truth conditions it assigns to example (1.15) are only satisfied if each child was kicking all the other children.

(1.14) $(\forall x \in A)(\forall y \in A)\, x \neq y \rightarrow xRy$

(1.15) The children were kicking each other.

Heim et al. were content to use this overly strong translation because reciprocity type was not germane to their immediate concerns, as it is not to mine. But again, it is desirable to replace this with a translation involving weaker truth conditions. Langendoen defines *weak reciprocity* as follows:

(1.16) $(\forall x \in A)(\exists y, z \in A)\, x \neq y \mathbin{\&} x \neq z \mathbin{\&} xRy \mathbin{\&} zRx$

Langendoen shows that this expression has a natural correspondence with the cumulative condition (1.12).

The issue of non-maximality also arises: as Williams (1991) notes, sentence (1.15) is compatible with a situation in which there are some non-kickers. In this example the existence of non-kickers may be seen as incidental and somehow ignorable; but we also find reciprocals used in situations where some elements are *necessarily* exempt. For example, example (1.17a) is judged true even though one child necessarily followed nobody, and another was followed by nobody. The same applies to example (b).

(1.17) a. The three children followed each other into the room.

b. The plates are stacked on top of each other.

Clearly, it would be preferable if a general solution would be provided that addresses exceptions to distributive and reciprocal sentences alike; on the other hand, it is not clear whether examples such as (1.17) should be treated as exceptions, or as a separate type of reciprocal relation on a par with strong and weak reciprocity.

Chapter 5 discusses the approaches of Sternefeld (1998), Schwarzschild (1992, 1996) and Brisson (1998), who address many of the issues raised in this section. However, these issues are largely orthogonal to the concerns of this dissertation. Accordingly, in other chapters the discussion is based on the simpler analysis of Heim et al. (1991a,b), which involves the semantics of strong reciprocity with all its attendant shortcomings. In section 5.4, I show how my proposals can be incorporated in an analysis along the lines of Schwarzschild's (1996) treatment.

Chapter 2

The Scopal Analysis of Reciprocals

Our point of departure is with the following sentences, based on examples first discussed in print by Higginbotham (1985) (who cites an unpublished ms. by Dan Finer, as well as other work by himself).

(2.1) a. John and Mary told each other that they should leave.

b. John and Mary think they like each other.

These examples pose problems for a straightforward treatment of coindexation as identity. For example, straightforward coindexation of (2.1a) only makes available the interpretation in which *they* refers to *John and Mary.* But there are at least two other readings, one in which John said that he should leave (and Mary said that *she* should leave), and one in which John said that *Mary* should leave and vice versa.

Higginbotham's own solution was a theory based on directed, non-transitive links between anaphoric elements and antecedents; but his non-standard semantics did not catch on, and the standard treatment of these examples is that of Heim, Lasnik, and May (1991a). As they note, there are two ways around the problem: enrich the notion of coindexation so it can capture this reading, or posit a more complex structure that makes it possible to account for the reading anyway. Higginbotham's analysis, of course, follows the first approach, while their own follows the second. Given the overall success of Heim et al. (1991a) in accounting for this kind of examples within the standard semantic framework, there is no motivation to abandon it in pursuit of a linking semantics.

Section 2.1 presents the issues and the solution proposed by Heim et al. (1991a), followed in section 2.2 by some criticism and an alternative proposal by Williams (1991). Williams's point of view has not been very popular in the literature, perhaps because his analysis is based on a Higginbotham-style linking theory with no obvious translation into model-theoretic semantics. But as I will show in more detail, some of his criticisms are well-founded and are only partially addressed by Heim et al. in their (1991b) reply.

2.1 The Heim, Lasnik and May (1991a) account

2.1.1 A "puzzle of grain"

Heim et al. (1991a) describe the problem posed by sentence (2.1a), repeated below, as "a puzzle of grain." The problem, in short, is that there is only one way to straightforwardly coindex the matrix subject, the reciprocal *each other*, and the pronoun *they* in sentence (2.1a): to assign the same index to all of them. But along with readings in which *they* refers to some third party (or parties), sentence (2.1a) has the three readings given by (2.2a), (b) and (c):

(2.1a) John and Mary told each other that they should leave.

(2.2) a. John told Mary that *he* should leave, and vice versa. ("I" reading)

b. John told Mary that *she* should leave, and vice versa. ("you")

c. John told Mary, and Mary told John, "We should leave". ("we")

Heim et al. (1991a) call these the "I", "you", and "we" readings, respectively. The terminology has the advantage of being theory neutral as well as easy to remember, and I will adopt it for the present discussion.

Informally speaking the "you" reading, at least, can be distinguished from the others by letting *they* take as its antecedent the object of *told*, instead of its subject. But there is only one straightforward indexing of sentence (2.1a), and it coindexes all the NPs in it:

(2.3) [John and Mary]$_i$ told [each other]$_i$ that they$_i$ should leave.

It is therefore impossible to distinguish the three readings by variations in indexing. This is the puzzle of grain.

The "I" and "you" readings have the common characteristic that, once we take pronoun reference into account, John and Mary said different things: one of them expressed a proposition about John leaving, the other about Mary leaving. In contrast, under the "we" reading the propositions expressed by John and Mary are identical, being in each case about both of them.[1] I will refer to the first type of readings as *dependent* readings, since the content of what is said depends on who said it; and to the second type as *independent* readings.[2,3]

[1]Philosophers of language are divided on the question of whether a sentence containing the pronoun "I" should be seen as expressing different propositions depending on the speaker, or the same "indexical" proposition or proposition-like entity. For example, see Perry (1993) and the discussion of diagonal propositions in Stalnaker (1999:pp. 12ff). For our present purposes, at least, it can be demonstrated that the distinction being made here is linguistically real. See section 3.3.4 for details.

[2]In the dependent readings, the pronoun behaves like a variable bound by a quantifier. I reserve the term "bound pronoun" for pronouns bound by an explicit quantifier, because quantified and implicitly distributed NPs differ with respect to, among other things, their ability to serve as the antecedents of reciprocals.

[3]Beyond the readings just discussed, the "we" reading is further ambiguous between the senses of John and Mary leaving together and their leaving separately. This is the standard ambiguity between distributive and collective readings commonly found with plurals. Example (i) may mean that every member of the group denoted by *we* voted against the amendment (the distributive reading), or it may mean that a vote taken by the group came out against the amendment, perhaps not unanimously (the collective reading).

(i) We voted against the amendment.

This distinction is not relevant to the matter at hand, although it is touched on in section 4.7.

2.1.2 A "puzzle of scope"

The second "puzzle" considered by Heim et al. (1991a), again attributed to Higginbotham, is exemplified by sentence (2.1b) (below), which has the readings given in (2.4).

(2.1b) John and Mary think they like each other.

(2.4) a. John and Mary think "We like each other". ("we"/narrow/independent)

b. John thinks he likes Mary, and vice versa. ("I"/broad/dependent)

The readings here clearly correspond to the "we" and "I" readings of (2.1a); but Heim et al. follow Higginbotham and others in referring to them in scope terms, as the "narrow" and "broad" readings, respectively. The terminology, as Heim et al. say, is "loaded": The idea is that *each other,* which requires a distributed antecedent, can either be construed with respect to the pronoun *they* (the narrow reading), whose antecedent must then be the entire matrix subject; or it can be construed with respect to the matrix subject (the broad reading), to which the embedded pronoun is also bound as a variable, resulting in the "I" reading. The construals can be diagrammed as follows:

(2.5) a. John and Mary think they like each other. ("we"/narrow/independent)

b. John and Mary think they like each other. ("I"/broad/dependent)

This understanding, which Heim et al. share with earlier discussions such as those of Higginbotham (1985) and Lebeaux (1983), is central to their analysis. In turn, their treatment has served as the basis for most subsequent accounts, such as those of Moltmann (1992), Sauerland (1995b), Schwarzschild (1996), Sternefeld (1998), etc. Williams (1986, 1991) is virtually the only dissenter, arguing for a non-scopal approach to such contrasts.

It is instructive to highlight the problem that the analysis of Heim et al. solves, in order to understand why it has been so influential and what would be required of an alternative:

It is easy enough to come up with a treatment of distributivity that allows the pronoun *they* in sentence (2.1b) the option of being bound, that is, of iterating over the parts of *John and Mary;* but any straightforward way of doing so turns *they* into a singular bound variable, as indicated in the partial translation (2.6). Since the semantics of reciprocals requires them to have a plural antecedent, the tricky part is to formulate an analysis that treats *they* as a bound variable and still provides a suitable plural antecedent for *each other.*

(2.6)	John and Mary		think	they	like	each other
	$\forall x \cdot \Pi\ J \oplus M$	x	thinks	[x	likes	???]

The now-standard solution adopted by Heim et al. is to have *each other* look elsewhere for a plural antecedent, namely, to the NP that binds the iterated pronoun. Williams (1991) opts for treating *they* as the only antecedent of the reciprocal, but allowing two different types of distributivity, only one of which creates a singular variable. My own solution, which is presented in chapter 4, enriches the representation of the embedded pronoun so that it can serve as an intermediary for the plural antecedent of the reciprocal. The resulting analysis retains much of the semantics of the Heim et al. account, but follows Williams in treating the reciprocal as having only one *direct* antecedent, the local binder.

2.1.3 A sketch of the Heim et al. analysis

The following is a brief glimpse of the Heim et al. (1991a) analysis, in case the reader has not already come across it. It is developed, in copious detail, in the remainder of section 2.1.

The central claim of the Heim et al. analysis is that the *each* part of the reciprocal raises to adjoin to the matrix subject, where it functions as a distributive operator. This is shown in (2.7):

(2.7) a. The men saw each other $\Rightarrow$

b. $[_S\ [_{NP}\ [_{NP}$ the men $]_1$ each$_2$ $]$ $[_{VP}$ saw $[_{NP}$ e$_2$ other $]_3$ $]$ $]$

In this system, reciprocals are complex entities: the *each* part is an ordinary distributor, which raises to adjoin to the antecedent of the reciprocal; its movement trace e is responsible for the Principle A behavior of reciprocals, while the remnant [*e other*] is syntactically an R-expression, translated into a quantifier that ranges over values disjoint from those of the variable bound by *each*.

Distributors, overt or covert, are translated as universal quantifiers. They raise at LF to adjoin to a subject NP, and quantify over its atomic parts. Dependent pronouns are translated as variables bound by such distributors. A simple reciprocal sentence like (2.7a) is then translated as in (2.8). The symbol $\cdot\Pi$ denotes the relation *proper-atomic-part-of*. The higher quantifier is contributed by the *each* part (the distributor), the lower one by the [*e other*] part of the reciprocal. The open variable X_i is the *range argument* of the reciprocal, and is required to be coindexed with the NP to which *each* has adjoined, *men*$_i$.

(2.8) $\forall x_j(x_j \cdot\Pi\ \mathrm{men}_i)\,\forall x_k(x_k \cdot\Pi\ \mathrm{X}_i\ \&\ x_k \neq x_j)\ \mathrm{saw}(x_j, x_k)$

Finally, so-called "long-distance" reciprocals are derived by raising the distributor *each* out of the embedded clause and adjoining it to the matrix subject, from where it binds the embedded dependent pronoun. An example is shown in (2.9).

(2.9) [John and Mary$_1$ each$_2$] think [that they$_2$ like [e$_2$ other]$_3$]

= John thinks "I like Mary", and Mary thinks "I like John".

2.1.4 Distributivity

Heim et al. (1991a) set themselves the task of using reciprocals to study the interpretation of plurals, and their treatment can be seen as addressing two conceptually separate issues: First, an analysis of plurals, including dependent plural pronouns, that can handle the group-individual ambiguities seen not only in reciprocal sentences like those in (2.1), but also in non-reciprocal examples like the following:

(2.10) a. John and Mary solved the homework problem.

b. John and Mary won $100.

Second, an account of reciprocal interpretation that explains the "puzzles" encapsulated by examples (2.1):

(2.1) a. John and Mary told each other that they should leave.

b. John and Mary think they like each other.

In this section I present the Heim et al. treatment of distributivity, which is conceptually independent of reciprocals (although, of course, it is largely motivated by the reciprocal facts). We return to reciprocals in section 2.1.5.

Heim et al. adopt a basic theory of plurals along the lines of the lattice theory of Link (1983), described in section 1.3. In other words their domain D of "individuals" consists of a subset A of *atomic* individuals, plus *sum* individuals corresponding to combinations of the elements of A. Distributivity makes use of the relation $\cdot\Pi$ (meaning, *proper-atomic-part-of*): if $a \cdot\Pi\ b$, then a must be an atom and a part of b, but cannot be equal to b. It follows that since atomic individuals do not have a *proper* subset that is an individual, b cannot be atomic.[4]

[4]Note a conflict between the notations of Link (1983) and Heim et al. (1991a,b): Link uses the symbol $\cdot\Pi$ to mean *proper-part-of,* while Heim et al. use it with the meaning *proper*-atomic-*part-of.* (Similarly, Π means *part-of* for Link, but *proper-part-of* for Heim et al). I use these symbols only in the sense of Heim et al.

The main distinction to be captured is between collective and individual properties; for example, sentence (2.11) is ambiguous between a distributive reading, according to which John and Mary each solved the problem without help, and a reading according to which they solved it together. Under the first reading, the sentence *John solved the homework problem* is true; under the second one, it is false.

(2.11) John and Mary solved the homework problem.

Heim et al. express the distributive reading by adjoining a covert distributor D to subject at LF; the distributor combines first with the subject NP and then with the predicate to give a distributively interpreted proposition. D is translated as in (2.12a), and a distributed NP applied to a predicate φ is interpreted as shown in (b).[5] The definition of $\cdot\Pi$ guarantees that the distributor cannot be adjoined to a singular NP.

(2.12) a. $\mathrm{D}_{<\mathrm{e},<\mathrm{et},\mathrm{t}>>} \equiv \lambda N\, \lambda\varphi\, \forall x_j (x_j \cdot\Pi\ N)\, \varphi(x_j)$

b. $[\mathrm{NP}_i\ \mathrm{D}_j]\varphi \Rightarrow \forall x_j (x_j \cdot\Pi\ \mathrm{NP}'_i)\, \varphi'(x_j)$

Thus, distributivity is treated as universal quantification. Heim et al. admit that this is an oversimplification; to sidestep the problems it encounters, they restrict their attention to plural subjects consisting of two members only.[6]

[5]The formulas in (2.12) are not those used by Heim et al.; their notation is not always explicit about the semantic arguments to elements of their translation, for instance presenting (2.12b) as follows:

(i) $[\mathrm{NP}_i D_j]\varphi \Rightarrow \forall x_j (x_j \cdot\Pi\ \mathrm{NP}_i)\varphi'$

Here the translation φ' of φ must be understood to contain an open instance of the variable x_j, which is bound by the universal quantifier. This notation follows Montague's conventions, but is too cryptic for my purposes. In the semantic calculus assumed here (see section 1.2), open variables may only be bound through lambda abstraction (which can be freely applied if it is not already licensed by movement traces etc.) Thus φ' will be of the form $\lambda x_j\, \psi(x_j)$. For clarity, I translate all formulas used by Heim et al. into equivalent closed formulas throughout this discussion.

[6]Heim et al. (1991a:p. 70) acknowledge that their treatment of distribution, among other things, "does not readily generalize to examples where the NP to which [the distributor] attaches contains a determiner other than the definite article, or where it co-occurs with a quantificational adverb." The problem is that the

Distributed NPs are said to have a *range* and a *distribution* index (i and j in (2.12b), respectively). The range index identifies the denotation of the entire NP, while the distribution index is contributed by the distributor and ranges over the atomic parts of the NP's denotation. In this way Heim et al. address the "puzzle of grain": by making available two indices associated with a distributed NP, they provide a way to distinguish between the collective and the distributive interpretations. For example, the collective and distributive readings of (2.13) are represented as shown in (2.14a) and (b), respectively.

(2.13) Mary and Sally introduced themselves to Max.

(2.14) a. Mary and Sally$_1$ introduced themselves$_1$ to Max.

b. [Mary and Sally$_1$ D$_2$] introduced themselves$_2$ to Max.

The collective reading (a) says that Sally and Mary, together, introduced themselves to Max, while the distributive (b) says that Sally introduced herself to Max, and Mary introduced *her*self. The anaphor *themselves* is in each case translated as a variable coindexed with a c-commanding antecedent, i.e., as a bound variable. Under the distributive reading, the anaphor ranges over atomic individuals. Heim et al. allow plural pronouns to freely refer to

translation of the distributor assumes that its complement NP is of type <e>, but non-definite NPs require a higher-order translation.

Another problem mentioned by Heim et al. is raised by one reading of the sentence "My grandparents dislike each other," in which the relation of disliking can be taken to be between the members of two groups, e.g., if my paternal grandparents dislike my maternal grandparents and vice versa. A similar problem arises because distributive sentences are considered true in situations where the predicate involved is neither collective nor true of all individual parts of its subject. For example, in a situation where the children built rafts in teams of five, sentence (i) is true but sentences (ii) and (iii) are false. The theory of Schwarzschild (1992, 1996), which I present in section 5.3, addresses both of these problems.

(i) The children built rafts.
(ii) Each child built rafts/a raft.
(iii) Mary built rafts/a raft.

Reciprocal sentences may exhibit *weak reciprocity*. Sentence (iv) is judged true if every participant hit someone and was hit by someone, but the semantics of Heim et al. require that everyone hit everyone else. All of these problems are sidestepped by restricting attention to situations involving just two individuals.

(iv) They were hitting each other.

atomic individuals; this seems a bit too liberal since a plural pronoun cannot *always* be used in place of a singular one (for example, pronouns bound by the quantificational NP *every woman* must be singular), but something of the sort is necessary in any account, given that dependent plural pronouns like the one in (2.14b) are understood as ranging over singular entities.[7,8]

2.1.4.1 The position of the distributor

The NP-distributor complex is headed by the distributor and inherits its index; hence in the distributive configuration (2.14b), the range index *1* of *Mary and Sally* does not c-command the anaphor *themselves.* Accordingly, Heim et al. (1991a) predicted that the following configuration is impossible:

(2.15) [Mary and Sally_1 D_2] introduced themselves_1 to Max.

The relevant reading would mean that Mary and Sally, separately, introduced the two of them to Max. This reading does appear to be absent. The presence of a matrix distributor, then, is predicted to always block a plural interpretation of the anaphor *themselves* by causing the matrix subject to become too deeply embedded to c-command into the embedded clause. This prediction is in fact inaccurate; it was criticized by Williams (1991) and revised, in a way that still leaves much to be desired, by Heim et al. (1991b). The details of this are taken up in section 2.3.1.

[7]This discussion is not intended to include uses of the pronoun *they* as a gender-neutral, singular pronoun, as in "Every gambler thinks they have only themself/themselves to blame."

[8]A more accurate characterization is provided by Kamp and Reyle (1993:p. 346): a pronoun is morphologically plural whenever its antecedent is introduced by a morphologically plural NP (excluding partitives). This applies whether the pronoun is coreferential with the antecedent or bound by it. Their findings are discussed in more detail in section 3.2.

2.1.5 Reciprocals

In the Heim et al. (1991a) analysis part of the reciprocal (the *each* part) raises to adjoin to the matrix subject, where it plays the role of a distributor as shown in (2.16):

(2.16) a. The men saw each other $\Rightarrow$

b. $[_S\ [_{NP}\ [_{NP}$ the men $]_1$ each$_2$ $]$ $[_{VP}$ saw $[_{NP}$ e$_2$ other $]_3$ $]$ $]$

In their system, reciprocals are complex entities: the *each* part is an ordinary distributor, which raises to adjoin to the antecedent of the reciprocal; its movement trace e is stipulated to be an anaphor, and is responsible for the Principle A behavior of reciprocals.

The remnant [*e other*] is translated into a quantifier that ranges over the values, drawn from the reciprocal's "range argument", that are disjoint from those of the variable bound by the raised *each* (known as the "contrast argument"). The semantics of [*e other*] are based on those of ordinary "pronominal" *other,*[9] which is given its own semantic translation as the following ternary relation:

(2.17) other $\Rightarrow$ $\lambda x\,\lambda y\,\lambda z\,(z \cdot\Pi\; y \;\&\; z \neq x)$

Here x is the *contrast argument,* specifying what an element z must differ from; and y is the *range argument,* of which z must be a part. The relation between the three arguments is something like "z is a part of y other than x."

In a reciprocal sentence like (2.16a), the remnant [*e other*] left after raising *each* is a quantifier phrase that raises to adjoin to the VP *saw e.* The result is a one-place predicate with the structure given in (2.18a), whose meaning is roughly "saw everyone other than

[9] So-called "pronominal" *other* is the form appearing in NPs like *the other, another,* etc. (It might be more properly termed "nominal" *other*). It is contrasted with "adjectival" *other,* which modifies a noun as in the NP *the other children.*

oneself." (The indices are consistent with those in (2.16b)). The complex [*e other*] is given the translation shown in (2.18b):[10]

(2.18) a. $[_{VP}$ $[_{NP}$ e$_j$ other(i) $]_k$ $[_{VP}$ saw e$_k$]]

b. [e_j other(i) $]_k \Rightarrow\ \lambda P\, \lambda y\, \forall x_k (x_k \cdot\!\Pi\ X_i\ \&\ x_k \neq x_j)\, P(y, x_k)$

This expression involves a plethora of variables, whose function is probably unclear at first inspection. The open variable X_i is not bound, but is coindexed with the syntactic antecedent of the reciprocal, the plural NP that the raised distributor *each*$_j$ adjoined to. It is the *range argument* of the reciprocal, and denotes the set of entities that the reciprocal ranges over. The variable x_k ranges over the parts of X_i that are distinct from the *contrast argument* of the reciprocal, x_j, which is bound by the distributive quantifier *each*$_j$.

In the reciprocated-over predicate, the original syntactic position of the reciprocal (i.e., usually object position) is occupied by the variable x_k, while the subject of the predicate is supplied by the external argument λy . In the canonical case, however, the subject of the reciprocal predicate is also the contrast argument. For example, each man in (2.16a) must have seen all men other than himself. In other constructions, the contrast argument of the reciprocal is still some higher argument of some constituent. In fact, as we see in more detail in section 2.1.7.2, the syntactic properties of the reciprocal lead Heim et al. to stipulate binding requirements which guarantee that the subject of the reciprocal predicate will be the contrast argument. Hence I will identify the contrast argument x_j with the subject of the reciprocal predicate y, simplifying formula (2.18b) into the more perspicuous (2.19):

[10]Once again, formula (2.18b) is my interpretation of the one provided by Heim et al. (1991a). In this case their intent is a bit less clear than usual. The formula they actually give (their (18)) is as follows:

(i) [e_j other $]_k\ \zeta \Rightarrow\ \lambda y\, \forall x_k (x_k \cdot\!\Pi\ X_i\ \&\ x_k \neq x_j)\, \zeta'(y)$

I have renamed their indices to make them consistent with those in other formulas, but the meaning of (i) is exactly as they present it. This formula does not indicate the role played by x_k: since the trace left after quantifier-raising [*e other*] out of the VP ζ bears the index k, the variable x_k represents the object of the verb in ζ. I have rewritten their formula as (2.18b) in order to make this dependence explicit.

As discussed in the body of the text, additional adjustments to this formula are desirable.

(2.19) $[\ e_j$ other(i) $]_k \Rightarrow\ \lambda P\, \lambda y\, \forall x_k (x_k \cdot\!\Pi\ X_i\ \&\ x_k \neq y)\, P(y, x_k)$

Although I am not aware of any cases where the contrast argument should not be the same as the local antecedent, the two functions of the reciprocal antecedent will occasionally need to be distinguished for purposes of discussion.

The quantificational force of reciprocal *other,* as Heim et al. acknowledge, is not quite that of a universal quantifier; but they choose not to tackle the thorny problem of its proper identification.[11] Instead, they use universal quantification and limit their examples to simple two-element plurals, for which reciprocal meaning is expressed equally adequately by universal or existential quantification.

Heim et al. assume QR of all quantifiers, whether or not this is necessary for interpretive reasons; when [*e other*] raises out of a reciprocal VP, it leaves behind a trace e_k which is translated as the variable x_k. The remnant VP is then abstracted into the two-place predicate $\lambda x\, \lambda y\, P(y, x)$ before combining with the reciprocal operator (2.19). Combinatorially, the same effects could have been attained by leaving [*e other*] in situ; the claim of uniform QR was motivated by the syntactic perspectives of the time.

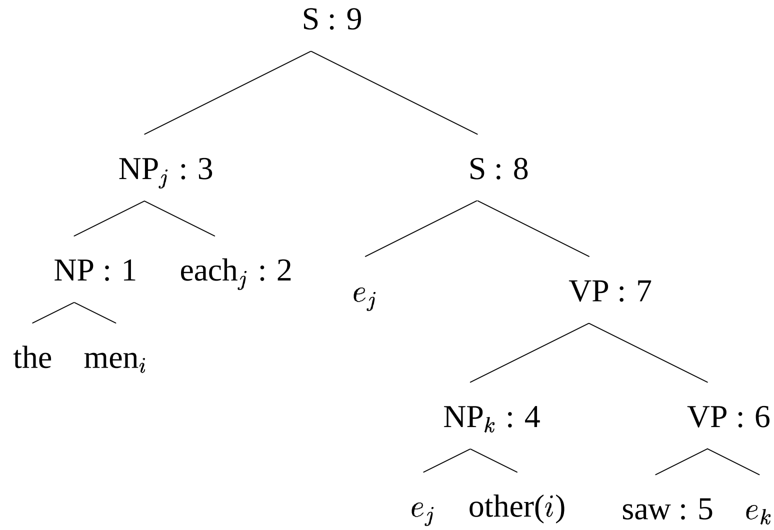
A simple reciprocal sentence like (2.16a) may be translated as shown below. First the *each* part of the reciprocal raises to adjoin to the subject NP, where it functions as a distributor; then both the subject-distributor complex and the remnant [*e other*] undergo QR, to give the structure given in (2.20a). The free variable giving the range argument of *other* is represented by the index i. After translating lexical items as discussed above and combining compositionally, we end up with formula (2.20b). The step-by-step derivation is given in (2.21), cross-referenced by the labels in the nodes of (2.20a).[12]

[11] A brief review of this issue was provided in section 1.4.

[12] There are many possible ways to formalize the binding of semantic variables. In this derivation I assume the mechanism of Heim and Kratzer (1998:p. 186) (see section 1.2 for a summary), in which a raised constituent causes the constituent to which it adjoins to undergo *Predicate Abstraction,* i.e., lambda abstraction over the index of the moved constituent. For simplicity, the indices are not represented as separate constituents in this tree.

(2.16a) The men saw each other.

(2.20) a.



b. $\forall x_j (x_j \cdot\Pi\ \mathbf{m})\, \forall x_k (x_k \cdot\Pi\ \mathbf{m}\ \&\ x_k \neq x_j)\, \text{saw}(x_j, x_k)$

(2.21) 1: ⟦ NP ⟧ = **m** (Represented as having type <e>)

2: ⟦ each ⟧ = $\lambda N\, \lambda\varphi\, \forall x_j (x_j \cdot\Pi\ N)\, \varphi(x_j)$ (= (2.12a))

3: ⟦ The men each ⟧ = $\lambda\varphi\, \forall x_j (x_j \cdot\Pi\ \mathbf{m})\, \varphi(x_j)$

4: ⟦ e$_j$ other ⟧ = $\lambda P\, \lambda y\, \forall x_k (x_k \cdot\Pi\ X_i\ \&\ x_k \neq y)\, P(y, x_k)$ (= (2.19))

5: ⟦ saw ⟧ = $\lambda x\, \lambda y\, \text{see}(y, x)$

6: ⟦ saw e$_k$ ⟧ = $\lambda y\, \text{see}(y, k)$

$\to \lambda k\, \lambda y\, \text{see}(y, k)$ (By PA on k)

7: ⟦ VP ⟧ = ⟦ e$_j$ other saw e$_k$ ⟧ = $\lambda y\, \forall x_k (x_k \cdot\Pi\ X_i\ \&\ x_k \neq y)\, \text{see}(y, x_k)$

8: ⟦ S ⟧ = ⟦ e$_j$ (e$_j$ other saw e$_k$) ⟧ = $\forall x_k (x_k \cdot\Pi\ X_i\ \&\ x_k \neq j)\, \text{see}(j, x_k)$

$\to \lambda j\, \forall x_k (x_k \cdot\Pi\ X_i\ \&\ x_k \neq j)\, \text{see}(j, x_k)$ (By PA on j)

9: ⟦ S ⟧ = $\forall x_j (x_j \cdot\Pi\ \mathbf{m})\, \forall x_k (x_k \cdot\Pi\ X_i\ \&\ x_k \neq x_j)\, \text{see}(x_j, x_k)$

$\to \forall x_j (x_j \cdot\Pi\ \mathbf{m})\, \forall x_k (x_k \cdot\Pi\ \mathbf{m}\ \&\ x_k \neq x_j)\, \text{see}(x_j, x_k)$

(by the range condition for reciprocals)

2.1.5.1 Reciprocals in subject position

What about reciprocal constructions other than the canonical reciprocal sentences we just saw? Reciprocals are anaphors, and as such are subject to Binding Principle A.[13] Besides dependent reciprocals (to which we return in the next section), the syntax of anaphors allows a number of other configurations that are worthy of mention. Exceptional Case Marking (ECM) constructions allow a reciprocal to appear in subject position:

(2.22) John and Mary believe [each other to be honest]

The analysis of Heim et al. extends to such examples if we assume that the reciprocal, which receives Case from the matrix verb, can raise out of the subordinate clause and apply to the entire matrix predicate.[14] Heim et al. (1991a) claim that the [*e other*] part of the reciprocal raises and adjoins to the VP that contained it, so that the VP becomes its complement; such movement is semantically unnecessary when the reciprocal appears in object position, but in sentence (2.22) it allows the reciprocal to modify the two-place predicate x-believe-y-to-be-honest.[15]

[13]Anaphors are not a uniform class; see Lebeaux (1983) for some differences between the binding properties of reciprocals and reflexives.

[14]This is unrelated to the “scopal” aspects of the Heim et al. analysis. Here the antecedent of the reciprocal is still its syntactic antecedent, which is sufficiently local to satisfy Binding Principle A.

[15]Reciprocals in prepositional complements raise similar issues; since quantifier-raising of the entire reciprocal in (i) is not as well motivated, it might be necessary here to allow the reciprocal to apply to a unary predicate.

(i) They showed me pictures of each other.

Example (ii) raises a different kind of issue. The embedded clause has an expletive subject, which allows the reciprocal to be bound by the matrix subject. This is something that our simple locality conditions on the antecedent of the reciprocal have not taken into account. Clearly this is a syntactic matter; given the assumption that the *each* part of the reciprocal can adjoin to the matrix subject, any semantic analysis of (i) should also apply to (ii).

(ii) They wanted there to be pictures of each other on the wall.

2.1.6 "Long-distance" reciprocals

We finally return to what Heim et al. (1991a) called the "broad" reading of sentences like (2.1b), repeated here with its two readings of interest:

(2.1b) John and Mary think they like each other.

(2.4) a. John and Mary think "We like each other". (narrow/independent)

b. John thinks he likes Mary, and vice versa. (broad/dependent)

As we saw in section 2.1.2, both readings construe the pronoun *they* as referring to *John and Mary:* the independent (or "narrow") reading attributes to each of John and Mary the thought "we like each other," and the dependent (or "broad") reading attributes to John the belief "I like Mary" and to Mary the belief "I like John." As already explained, Heim et al. treat the contrast in scope terms: the independent reading is generated according to the account given in the preceding sections, and involves raising of the distributor *each* to the embedded subject, as in (2.23a). The dependent reading involves something new: the distributor is said to raise out of the embedded clause and adjoin to the matrix subject, as shown in (b).

(2.23) a. [[John and Mary]$_1$ D] think that [[they$_1$ each$_2$]$_2$ like [e$_2$ other]$_3$]

= John thinks "we like each other", and Mary thinks the same.

b. [John and Mary$_1$ each$_2$] think [that they$_2$ like [e$_2$ other]$_3$]

= John thinks "I like Mary", and Mary thinks "I like John".

Since reciprocals are anaphors subject to Principle A, non-local movement should be disallowed. Heim et al. resolve the issue by allowing the local antecedent (in this example, the embedded pronoun *they*) to satisfy the Principle A requirement of the reciprocal. This part of their account runs into several problems, discussed in section 2.2.1.

In the dependent reading, the embedded pronoun *they* is a variable bound by the distributor along with the trace e_2, and takes atomic (singular) values. The long-distance reciprocal configuration provides the reciprocal with the plural antecedent it requires. The dependent reading of (2.1b) is translated as shown in (2.24b). For comparison, the translation of the independent reading is also given, as (a). The arrows indicate the antecedent of the range argument of the reciprocal; as noted, this is required to be the sister of the distributor that binds the reciprocal's contrast argument (x_2 in these examples).

(2.24) a. $\forall x(x \cdot\Pi\ J{\oplus}M_1)$ think$(x, {}^\wedge[\forall x_2(x_2 \cdot\Pi\ J{\oplus}M_1)\, \forall x_3(x_3 \cdot\Pi\ X_1)\, x_2 \neq x_3 \Rightarrow$ like$(x_2, x_3)])$ (Independent)

b. $\forall$x$_2(x_2 \cdot\Pi\ J{\oplus}M_1)$ think$(x_2, {}^\wedge[\forall x_3(x_3 \cdot\Pi\ X_1)\, x_2 \neq x_3 \Rightarrow$ like$(x_2, x_3)])$ (Dependent)

The preceding account glosses over the fact that the embedded pronoun *they* has, *in principle,* two indexing options: it can be dependent (bound by the distributor) or independent (coindexed with the antecedent *John and Mary*), giving a total of four possible combinations with the position of *each.* But only two of these four are actually licit, one for each position of *each:* if the latter raises to the matrix subject as in (2.23b), the embedded pronoun must be bound by the distributor in order to satisfy the Principle A requirement of the reciprocal;[16] if *each* adjoins to the embedded pronoun instead, as in (2.23a), the pronoun must be plural in order to allow distribution, i.e., it must have the independent interpretation.

[16]The missing reading would involve a wide-scope reciprocal combined with a plural embedded pronoun; it would cause sentence (2.1b) to say that John thinks that "we like Mary", and Mary thinks "we like John."

2.1.7 Some details of the analysis

So far we have discussed the core aspects of the Heim et al. (1991a) analysis. The empirical data on reciprocals make necessary a number of additional details, the most important of which are sketched here for completeness.

2.1.7.1 The semantics of [*e other*]

Heim et al. (1991a) intended their semantics of *each other* to be compositionally derived from the semantics of the quantifier *each* (translated as a distributor) plus the semantics of the element *other.* Unfortunately, the properties of *other* as part of a reciprocal do not entirely match those of stand-alone "pronominal" *other.*

In reciprocals, there is no freedom in selecting the values of the arguments x and y: the contrast argument x is bound by the distributor corresponding to the *each* part of the reciprocal, while the range argument y is unbound, but necessarily coreferential with the entire antecedent of the reciprocal (the sister of the distributor that binds the contrast argument). Accordingly, Heim et al. (1991a:fn. 3) must stipulate that "once the contrast index is determined, the choice of the range index is fixed as well: it is always the index of the sister of the contrast argument's A′-binder," i.e., the sister the distributor.

On the other hand the range argument of non-reciprocal *other* is determined more flexibly. Heim et al. (1991a) cite the following example, attributed to Mats Rooth (personal communication): the non-reciprocal sentence (2.25a) has a reading according to which each of the youngest three women gave a lecture to all women other than herself, but the reciprocal (2.25b) can only be understood as saying that the lecture was given to the other two of the three youngest women. (Sentence (a) also has this reading).

(2.25) a. The youngest three of the women each gave a lecture to the others.

b. The youngest three of the women gave lectures to each other.

Thus, although the goal of Heim et al. was to account for the semantics of (2.25b) by attributing to it an LF-structure equivalent to that of (2.25a), they acknowledge that the semantics of *each other* cannot be taken to be simply the sum of the properties of "floated" *each* and of pronominal *other.*

2.1.7.2 The syntax of the trace *e*

Any hope for a strictly compositional treatment of *each other* as the sum of its component words is thwarted by its status as an anaphor, subject to Principle A. Since neither *each* nor *other* is an anaphor when used on its own, it is necessary for Heim et al. (1991a) to explicitly associate the Principle A properties of the reciprocal with some part of it. They choose the trace e left after raising the distributor *each* out of the reciprocal, of which they stipulate: "e of *each* is an anaphor" (p. 73).

In instances of "long-distance" reciprocals, such as example (2.26a) under the dependent reading, the distributor is claimed to raise beyond the local domain, resulting in the structure shown in (b). In such sentences the Principle A requirement is satisfied by the presence of the coindexed pronoun *they,* as can be seen by the ungrammaticality of the otherwise similar (c).[17]

(2.26) a. John and Mary think they like each other.

(Dependent reading: John thinks he likes Mary, and vice versa).

b. [[John and Mary]$_1$ each$_2$] think they$_2$ like [e_2 other].

c. * John and Mary think I like each other.

I find that the association of the movement trace with the reciprocal's status as an ana-

[17]The coindexation involved is satisfying the Principle A requirement forces the local antecedent, which is the subject of the reciprocal predicate, to be identical with the contrast argument even in long-distance constructions. This justifies my collapsing these two arguments of the reciprocal and simplifying formula (2.18b) into (2.19).

phor has both technical and conceptual problems; but I will defer discussion of them until section 2.2.1, where the claim that *each* raises out of the reciprocal is also questioned.

2.1.7.3 Multiple reciprocals: "Absorption"

Sentences containing multiple reciprocals with the same antecedent present a special problem for the Heim et al. (1991a) account: since distribution over the antecedent is supposed to be contributed by the reciprocal, the presence of a second reciprocal should result in having one distributor too many. But sentences like the following are in fact grammatical with the matrix subject as the antecedent of both reciprocals:

(2.27) a. John and Mary read each other's books in each other's languages.

b. John and Mary told each other that they love each other.

c. They gave each other pictures of each other.

The relevant reading of (a) says that John read Mary's book in Mary's language, and Mary read John's book in John's language. The relevant reading of (b) says that John told Mary that he loves her, and of (c) that John gave Mary a picture of Mary and vice versa. The *each*-raising account predicts that the *each* part of both reciprocals should raise to adjoin to the matrix subject, giving a configuration like (2.28). To account for such readings, Heim et al. allow for "absorption" of multiple coindexed instances of reciprocal *each* into a single distributor.[18]

(2.28) [John and Mary] told each other that they love each other.

Heim et al. (1991a) note that sentence (2.27b) is ambiguous between the "we" and "I" (independent and dependent) readings; the two readings can be generated according to the

[18]Absorption, it turns out, cannot apply to non-reciprocal "floated" *each;* see section 2.2.1 for the data and discussion.

process described in section (2.1.6). The higher reciprocal always distributes over *John and Mary,* while the lower one has two options, giving rise to the two readings: it can attach to the pronoun *they,* giving rise to the "we" reading, or it can raise to the matrix subject and absorb with the first reciprocal, giving rise to the "I" reading as shown in (2.28).

The absorption mechanism is rendered unnecessary in the revised analysis of Heim et al. (1991b) (presented in section 2.2.2), which abandons raising of *each* in favor of binding it in situ by an independently adjoined distributor. Sentences involving multiple reciprocals still raise problems for the revised account; they will be discussed in section 2.3.3.

2.2 Problems, and a revised proposal (Heim et al. 1991b)

The account of Heim et al. (1991a) made several strong claims and predictions that, if successful, would have made it extremely appealing. It attempted to derive the semantics of reciprocals compositionally from independently-motivated translations for the non-reciprocal uses of its component words *each* and *other;* it involved several instances of covert syntactic movement, which was expected to obey the usual syntactic constraints and have predictable effects, detectable through grammaticality judgements. Unfortunately, many of these predictions turned out to be too strong; the original (1991a) analysis of Heim et al. has a number of shortcomings, several of which are highlighted by Edwin Williams in his (1991) reply to their proposal. Heim et al. (1991b) responded with a reply of their own which repairs their analysis in response to Williams's criticisms, in large part by retracting the claim of covert movement by the reciprocal.

In the revised analysis the *each* part of the reciprocal, rather than raising, is bound in situ by an independently inserted covert distributor. (Recall that even in the version of Heim et al. 1991a, covert distributors can be freely inserted as necessary). They call this the "*each*-binding" version of their account, as contrasted to the original "*each*-raising" version. Their discussion is somewhat unclear on whether they actually embrace the revised proposal, and subsequent discussion in the literature has generally been based on the original analysis. But the revised proposal is easier to defend than the stronger original analysis, and is also better suited to discussion of the issues raised here; accordingly, discussion in subsequent chapters will be based on the revised analysis. As Heim et al. point out, it shares the essential features of the original, movement analysis—including the claim that "long-distance" reciprocals involve binding of the reciprocal by a non-local antecedent.

Sentence (2.29) would have the LF configuration (a) under the 1991a account; the 1991b version represents it as in (b).

(2.29) They like each other.

a. [They each$_i$]$_i$ like [e$_i$ other].

b. [They D$_i$]$_i$ like [each$_i$ other].

Several of Williams's criticisms turn out to involve the claim that *each* raises covertly from its surface position and assumes the role of a distributor adjoined to the reciprocal's antecedent NP. In the next section I discuss these arguments, and some of my own, as a way of motivating the switch to the revised Heim et al. (1991b) account. I single out these arguments here because the revised account, by eliminating movement of *each*, addresses them so thoroughly; I defer discussion of other problems, which persist in the revised account, to section 2.3.

2.2.1 On *each*-movement

Williams (1991) provides a multitude of arguments against the assertion by Heim et al. (1991a) that the semantics of non-reciprocal "floated" *each* are identical to those of distributed plurals and the *each* part of reciprocals. One of these arguments involves *absorption*, which Heim et al. (1991a) introduced as a way of accommodating multiple reciprocals with the same antecedent (see section 2.1.7.3). But while it is possible for two reciprocals with the same antecedent to absorb into a single distributor as in (2.30a), "floated" *each* cannot absorb with reciprocal *each.* This is unexpected, in view of the claim by Heim et al. (1991a) that reciprocal and non-reciprocal *each* have identical translations and undergo QR to the same destination.

(2.30) a. They gave each other pictures of each other.

b. * They each gave each other pictures of the other.

Williams also argues that the dependent/independent ambiguity seen in a reciprocal sentence like (2.31a) is "nothing more than the ambiguity latent in [the non-reciprocal] (2.31b)." Hence the account of Heim et al. (1991a) does not achieve the proper division of labor: in reciprocal sentences the source of distributivity (and binder of the pronoun under the dependent reading) is said to be the covertly raised *each,* while in non-reciprocal distributed sentences the same effect is achieved by an identically interpreted distributor inserted at LF. This is not a knock-down argument (Heim et al. respond that they sufficiently capture the similarity by giving identical translations to *D* and *each*), but it nevertheless detracts from the appeal of the Heim et al. (1991a) analysis.

(2.31) a. John and Mary think they like each other.

[[John and Mary] each$_i$] think they$_i$ like e$_i$ other.

b. John and Mary think they are sick.

[[John and Mary] D$_i$] think they$_i$ are sick.

Another line of argumentation against movement of *each,* curiously not followed by Williams, involves the status of the trace as the bearer of the reciprocal's Principle A requirement. Although it is clear that *something* in the reciprocal is responsible for its status as an anaphor, it is less certain that the culprit could be a movement trace. I find this part of the Heim et al. (1991a) proposal to have both conceptual and technical problems. To begin with, the status of e as an anaphor is presumably a lexical property, since it must be stipulated. But movement traces, at least the kind assumed to occur here (prior to the Copy Theory of movement), are generated as the result of movement and cannot be said to exist at a level compatible with lexical specification.

The Heim et al. analysis of "long-distance" reciprocals like example (2.26a), repeated below, lets the distributor raise non-locally and adjoin to a higher antecedent, resulting in the structure shown in (b). The reciprocal's Principle A requirement is satisfied by the presence of the coindexed pronoun *they;* but far from helping license the trace, movement of *each* past a coindexed pronoun as in (2.26b) would create the classic strong cross-over (SCO) configuration, and should be ungrammatical.[19] The configuration of the indexes in (2.26b) is parallel to that created by QR in the ungrammatical example (2.32).

(2.26) a. John and Mary think they like each other.

(Dependent reading: John thinks he likes Mary, and vice versa).

b. [[John and Mary]$_1$ each$_2$] think they$_2$ like [e_2 other].

(2.32) a. He loves every boy. ($\neq$ Every boy loves himself).

b. * [Every boy]$_i$ he$_i$ loves t$_i$. (SCO)

These objections lose their force in the revised analysis proposed by Heim et al. (1991b). In that analysis, *each* is not a quantifier but a variable; it gets bound by a covert distributor adjoined, independently, to the antecedent of the reciprocal. The role of the trace e is then taken over by the *each* part itself, and there is no obstacle to lexically specifying that the

[19]Strong cross-over follows if *each* raises through A′ movement. This is the most likely situation: there is no motivation for supposing that *each* raises via A movement, and it raises too far for it to be head movement.

latter is an anaphor.

Finally, consider another class of examples apparently overlooked by both Williams and Heim et al. The *each*-raising account implies that movement of *each,* and hence reciprocal construals, should be sensitive to movement islands. That is, it should be impossible for a reciprocal inside an island to take the dependent reciprocal ("long distance") interpretation. Moreover, since non-reciprocal sentences receive freely inserted distributors that do not involve movement, the dependent reading should be possible in the corresponding non-reciprocal sentence. In other words, the non-reciprocal (2.33a) should allow the dependent reading across the intervening factive island, but the presence of the reciprocal in (2.33b) should *inhibit* the dependent reading. This prediction is not borne out: both sentences are judged as grammatical under the dependent reading, contrary to what the *each*-raising account predicts.

(2.33) a. John and Mary wondered why they were fired.

b. John and Mary wondered why they were fired by each other's boss(es).

Under the revised, *each*-binding account, the dependent reading of sentence (2.33b) is represented as in (2.34): distributive operators are freely inserted in both sentences and movement is not involved. Hence there is no movement out of islands, and no crossover, and both sentences are correctly predicted to be grammatical.

(2.34) [[John and Mary] D_i] wonder why they_i were fired by each_i other's boss.

2.2.2 More on the "*each*-binding" account

Sentence (2.29), which would have the LF configuration (a) under the 1991a account, is represented as in (b) under the 1991b version.

(2.29) They like each other.

a. [They each_i]$_i$ like [e_i other].

b. [They D_i]$_i$ like [each_i other].

The switch from raising to binding the *each* part of the reciprocal solves several problems with the original account, including those raised in section 2.2.1: as we saw, it makes moot any syntactic objections associated with the ability of *each* to find its antecedent across movement islands, and with the fact that movement of *each* ought to generate a strong crossover violation. In addition, the revised account associates Principle A with the *each* part of the reciprocal itself, not with its movement trace, and so is conceptually cleaner.

The *each*-binding account also simplifies the treatment of sentences with multiple reciprocals: since each reciprocal does not introduce its own distributor, it is no longer necessary to stipulate distributor absorption. (Nor is it necessary to explain why non-reciprocal *each* cannot "absorb", as shown in section 2.2.1). Finally, it achieves the proper division of labor in accounting for the ambiguity between dependent and independent reciprocal readings: under the 1991b account, the readings are distinguished by the positions of independently-inserted distributors, just as in sentences without reciprocals. The independent and dependent readings of sentence (2.1b) are represented as (2.35a) and (2.35b), respectively:

(2.1b) John and Mary think they like each other.

a. [[John and Mary]$_1$ D] think that [[they_1 D_2]$_2$ like [each_2 other]$_3$] (independent)

b. [[John and Mary]$_1$ D_2]$_2$ think that [they_2 like [each_2 other]$_3$] (dependent)

Heim et al. (1991b) adopted some other modifications to their original account. Some of these are details irrelevant to my concerns; others involve issues that have not yet been raised, and are discussed more extensively elsewhere. (See section 2.3.1 for discussion of distributor position, and section 2.3.3 for discussion of chained multiple reciprocals). The remainder of this section touches on a couple of issues that are not immediately germane to the concerns of this dissertation.

2.2.2.1 The translation of distributors and "floated" *each*

The original proposal of Heim et al. (1991a) derived much of its appeal from their attempt to derive the semantics of reciprocals compositionally from those of *each* plus those of *other*, and to identify the semantics of reciprocal and non-reciprocal ("floated") *each* with those of implicit plural distributivity. But the second of these attempts at conceptual economy, like the first, is overly ambitious. We have already seen (in section 2.2.1) that although two reciprocals with the same antecedent are grammatical, non-reciprocal "floated" *each* cannot have the same antecedent as a reciprocal. The relevant examples are repeated here:

(2.30) a. They gave each other pictures of each other.

b. * They each gave each other pictures of the other.

Sentence (b) was predicted to be grammatical by the Heim et al. (1991a) analysis, since absorption should apply identically to raised reciprocal *each* and non-reciprocal "floated" *each*. In the revised analysis absorption has been abandoned, hence these facts are automatically accounted for. Note that the revised analysis continues to translate floated *each* as a quantifier, but treats reciprocal *each* as a bound variable (albeit one subject to Principle A, as mentioned). Hence the two types of *each* no longer receive the same translation.

This still leaves open the possibility that "floated" *each* is semantically identical with a covert distributor. However, Williams (1991) also shows that non-reciprocal "floated" *each* introduces stricter truth conditions than do reciprocals and implicitly distributed plurals; for example, sentence (2.35a) is false in a situation where there was a lot of hitting but there were some non-hitters, but sentences (b) and (c) are true. This is the phenomenon of *non-maximality*, described in section 1.4.[20]

(2.35) a. They were each hitting the others.

b. They were hitting each other.

c. They were hitting Bill.

Heim et al. (1991b) relax their semantics for distributors accordingly: they accept Williams's argument that plural distribution, and along with it reciprocal distribution, is not "down to individuals," and adopt a vague criterion of "sufficiently many relevant parts" of the distributor's complement. The *proper-atomic-part-of* operator, $\cdot\Pi$, is replaced by the *proper-part-of* operator Π . As suggested by the reference to "*relevant* parts," this relation is not intended to range over all possible proper parts of its right operand, but over some contextually appropriate collection of such parts. Heim et al. do not elaborate on this point, but the system of "covers" proposed by Schwarzschild (1996) provides collections of parts along the lines that they appear to have envisioned. The revised formula for distributors is given in (2.36). It differs from formula (2.12a) only in the substitution of the relation Π in place of $\cdot\Pi$.[21]

(2.36) $\mathrm{D}_{<\mathrm{e},<\mathrm{et},\mathrm{t}>>} \equiv \lambda N\, \lambda\varphi\, \forall x_j (x_j\ \Pi\ N)\, \varphi(x_j)$

The weaker relation Π is also used with the lower universal quantifier introduced by the reciprocal, which ranges over the objects of the reciprocal predicate. The reciprocal

[20]Such differences in the behavior of plurals and quantified NPs (but not reciprocals) are also discussed by Kroch (1979). His views are briefly presented in 5.3.4, along with Brisson's (1998) analysis of non-maximality by means of *ill-fitting covers.*

[21]Heim et al. (1991b) also define a Generalized Quantifier D in place of the universal quantifier $\forall$, whose truth conditions are such that its predicate argument must be true of "sufficiently many relevant parts" of the distributor's domain. However, they only use this quantifier in their reconstruction of Williams's (1991) system, not in their own system.

translation proposed by (Heim et al. 1991a) was given as (2.19). The revised version can be written as in (2.37). The two formulas only differ in the part relation used, and of course in whether they represent the entire reciprocal or just the remnant left after raising *each.*

(2.19) [e_j other(i) $]_k \Rightarrow \; \lambda P\, \lambda y\, \forall x_k (x_k \cdot\!\Pi\ X_i\ \&\ x_k \neq y)\, P(y, x_k)$

(2.37) [each$_j$ other(i) $]_k \Rightarrow \; \lambda P\, \lambda y\, \forall x_k (x_k\ \Pi\ X_i\ \&\ x_k \neq y)\, P(y, x_k)$

Since the quantification introduced by non-reciprocal *each* is actually "down to individuals," the latter retains the stronger interpretation in terms of $\cdot\!\Pi$. The change allows the interpretation of non-reciprocal *each* to be distinguished from that of reciprocals and distributors, but necessarily abandons one more claim to economy of the Heim et al. (1991a) account.

2.2.2.2 On quantifier-raising the NP-distributor complex

The analysis of Heim et al. (1991a) treats distributors as quantifiers, and the NP-distributor complex undergoes QR in order to be interpreted. The claim of QR, as Williams points out, implies that reciprocals should interact scopally with other quantifiers and should be sensitive to the usual restrictions on covert movement. But such expectations, as Williams puts it, are "uniformly disappointed."

Williams's argument is based on the contrast between the scope options of *each other* and those of standalone *each:* The latter has the properties of a true quantifier and can take widest scope, as shown by examples (2.38) and (2.39), in which *each* and *every* can be understood as having scope over *someone or other.* But sentence (2.40) does not have such a reading.

(2.38) a. Someone or other, at one time or another, has bothered each of our candidates.

b. Someone or other has said that each of the men likes the other.

(2.39) a. Someone or other has bothered every candidate.

b. Someone or other has said that every man likes the others.

(2.40) Someone or other has thought that they like each other.

$\neq$ Each of them is such that someone or other has thought that he liked the other.

As Williams (1991:p. 171) puts it, the (presumed) scope of *each* is "restricted to be only as wide as the highest instance of the variable it defines distribution and reciprocation for[.]" Williams's argument is based on the belief that QR is not clause-bound; as he notes, this assumption is also involved in generating the "long-distance" reciprocal configuration under the (Heim et al. 1991a) analysis. However, it is now accepted that QR is indeed clause-bound. On the one hand, this means that his examples cannot used to show what they were intended to show, that the *each* part of *each other* does not undergo QR. But on the other, it means that the Heim et al. analysis of dependent reciprocals cannot rely on quantifier raising across clauses.

As we saw, the revised "*each*-binding" analysis does not involve raising of *each* across clause boundaries, and therefore does not run afoul of QR locality requirements. In addition, in response to Williams's arguments Heim et al. (1991b) assume that: "antecedents of reciprocals do not undergo QR", but are obligatorily interpreted in situ. Quantifier-raising of the distributed NP was in any event combinatorially unnecessary, and could be abandoned without requiring any revisions to the semantics. Since they additionally represent the antecedents of reciprocals as ordinary distributed NPs, Heim et al. view their failure to raise as a consequence of the general property that "implicitly distributed plurals are not free to take scope." For example, if a distributor adjoined to the pronoun *they* in (2.41) could cause the pronoun-distributor complex to undergo QR, it would generate the reading "for each x, someone or other saw x." This reading is impossible, hence Heim et al. rule out QR for all implicitly distributed plurals, whether they are reciprocal antecedents or not.[22]

[22] Heim et al. actually give the following example, which would not be derivable via clause-bound QR even

(2.41) Someone or other saw them.

This conclusion, although originally based on the faulty assumption that it is possible to quantifier-raise across clause boundaries, is nevertheless consistent with the findings of later inquiries into this topic. In particular, Krifka (1992) and (independently) Dayal (Srivastav 1992, Dayal 1996) have shown conclusively that implicitly distributed NPs cannot undergo QR in wh-question constructions which allow QR of quantificational NPs.

The scope interaction of reciprocals with quantifiers is studied in more detail by Asudeh (1998:pp. 116–132), who confirms that reciprocals show no evidence of interacting scopally with other quantifiers. Unfortunately the data he considers are much too delicate to provide knock-down arguments about the status of reciprocals; but he succeeds in showing that reciprocals do not have the properties expected of ordinary quantifiers.

2.3 Additional problems

Although Heim et al.'s weaker (1991b) theory addresses several objections to the Heim et al. (1991a) account, other problems remain. The ones that concern us are related to the analysis of distributivity and to the long-distance analysis of dependent reciprocal sentences.

Section 2.3.1 presents Williams's (1991) arguments against the analysis of distributivity given by Heim et al. (1991a). But rather than adopt his conclusions (or the solution proposed by Heim et al. (1991b)), in section 2.3.1.3 I show that the issues he raises are adequately addressed if we replace the NP-adjoined distributors of Heim et al. (1991a) with equivalent operators adjoined to VP. From the point of view of accounting for reciprocal

if it applied to definite NPs. Fortunately, the same effects hold for single-clause examples such as (2.41).

(i) Someone or other has thought that they would win.

readings, this can be seen as a technical adjustment that leaves intact the essentials of the Heim et al. analysis.

I present two types of criticisms of the treatment of “long-distance” (dependent) reciprocals. The first type involves reciprocals whose intended range argument appears in a position so deeply embedded that a distributor adjoined to it could not bind the reciprocal. The relevant examples involve the interaction of two phenomena: pronouns dependent on distributed NPs that do not c-command them, and reciprocals whose antecedent is such a pronoun. They are discussed in section 2.3.2.

The second type of counterexample to the long distance analysis involves chained multiple reciprocals, i.e., cases where one reciprocal serves as the antecedent of another. In section 2.3.3 I show that the translation that Heim et al. (1991b) give for such examples cannot be derived in a principled way using their system.

The two types of problems relate to different elements of the system of Heim et al., but will be accounted for in a uniform way in the analysis of reciprocals I propose. This is shown in section 5.4.1.

2.3.1 On NP-adjoined distributors

Williams (1991) argues against the NP-adjoined distributors adopted by Heim et al. and in favor of his own alternative, which views distributivity as a relation between two elements. His arguments center on counterexamples to two consequences of the Heim et al. analysis: First, NP-adjunction of a distributor embeds the distributed-over subject into a position from which it does not c-command into the VP. Heim et al. (1991a) appeal to this feature of their analysis to explain the binding effects discussed in section 2.1.4.1, but Williams shows that the configuration they rule out is in fact possible. This is discussed in section 2.3.1.1. Second, NP-adjunction of distributors implies that the subjects of complex VPs must either have a uniformly collective or a uniformly distributive interpretation. Again, Williams shows that this prediction is incorrect. His arguments are presented in section 2.3.1.2.

While Williams's arguments bring out the inadequacies of the Heim et al. analysis, they do not demonstrate the correctness of his conclusion, that distributivity must be represented by means of some alternative system like his linking theory. In section 2.3.1.3 I show that the relevant examples are easily handled within the standard semantic framework by simply switching to VP-adjoined distributors.

2.3.1.1 C-command effects

Heim et al. (1991a) rule out a reading in which the reflexive pronoun *themselves* of sentence (2.42a) refers jointly to Mary and Sally. Because reciprocals are claimed to force a distributive reading, the syntactic structure (b) created by adjunction of the distributor should prevent the plural NP *Mary and Sally* from c-commanding (and hence, from binding) the reflexive. Since reflexive pronouns must be locally bound, the pronoun *themselves* cannot be assigned the index j.

(2.42) a. Mary and Sally introduced themselves to each other.

b. [[Mary and Sally]$_j$ each$_i$]$_i$ introduced themselves$_{i/*j}$ to e$_i$ other.

The relevant reading would mean that each of Mary and Sally, separately, introduced the two of them to the other. Indeed, this reading is absent. But as Williams (1991) points out, its absence should be attributed to pragmatic oddness; other examples, such as his (2.43a) or the more parallel example (2.43b) (suggested by Richard Larson and discussed by Heim et al. (1991b)), allow an anaphor to be given a plural interpretation.

(2.43) a. John and Mary gave each other pictures of themselves.

b. Simon and Garfunkel praised themselves to each other.

In order to account for the well-formedness of such examples, Heim et al. (1991b) appeal to the option of satisfying the Principle A requirement of the anaphor *themselves* at S-structure, prior to adjunction of the distributive operator to the subject. While this was a legitimate (if uninspiring) analysis in the syntactic framework of the time (Chomsky 1981, 1986), it is untenable if we adopt the Minimalist Program's (Chomsky 1993) requirement that Binding Theory conditions can only be checked at LF:[23] Since the interpretation of these sentences must involve distribution, an NP-adjoined distributor must be present at the interface with the interpretive component. This means that if the Binding Principles are checked at the interface, they will necessarily be checked after the distributors have been inserted, and the examples in (2.43) are predicted to be ungrammatical. At any rate, the revision considerably weakens the predictive force of their theory.

2.3.1.2 VP conjunction effects

A second argument against NP-adjoined distributors is supported by the following examples, also provided by Williams:

(2.44) a. They collided and criticized each other's driving.

b. They criticized each other's driving just after PRO colliding.

Since the verb *collide* of sentence (2.44a) requires a group antecedent, it is incompatible with a distributed subject; while the reciprocal predicate *criticize each other's driving* should force distributive interpretation on the subject. The structure posited by Heim et al. (1991a) for (2.44b) involves adjunction of *each* to the subject *they;* in such a structure, the non-distributed *they* is buried in the distributive constituent *they each*, and cannot be the

[23] The requirement that Binding Theory apply in a uniform way is a well-motivated sanity condition on syntactic analyses, and should be desirable regardless of one's position on the specifics of Minimalist syntax.

Within the Minimalist framework, enforcement of the Binding Principles exclusively in LF has been defended, inter alia, by Fox (1995).

antecedent of PRO since it does not c-command it:

(2.45) [They each_i] criticized t_i other's driving just after PRO colliding.

PRO, as Williams notes, cannot be made coreferent with *they* via discourse anaphora, the way an overt pronoun could. Heim et al. (1991b) suggest two possible accounts of sentences like (2.44b) (they do not discuss (2.44a) at all), neither of them very satisfactory. The first option they consider is to claim that the c-command requirement is satisfied prior to adjunction of the distributor, as in the case of reflexives. The second is to let the distributor adjoin VP-internally to the trace of a VP-internal subject, leaving the overt subject unembedded:

(2.46) They_1 [$_{VP}$ [e_1 each_2] criticized e_2 other's driving] just after PRO_1 colliding.

Heim et al. favor the second account on the grounds that it *differs* from their account of reflexive binding, and they find that these constructions do not pattern identically with those involving reflexives.[24] The weaknesses of the first alternative were already discussed in section 2.3.1.1. The second alternative is probably close in combinatory properties (though not in syntactic analysis) to the solution I discuss in the following section, namely, adjunction of distributors to VP rather than NP.

[24] Reflexives and conjoined VPs pattern differently with respect to floated *each,* which blocks plural reflexive binding but does not interfere with mixed group/distributed VP conjunction.

(i) The men_i each_j saw $\text{themselves}_{j/*i}$.

(ii) a. They each criticized the other's driving just after PRO colliding.

b. They each ran a catering service before opening a restaurant together.

Other differences between reciprocals and floated *each* were discussed in section 2.2.2.1, and more are given in section 5.1.1.

2.3.1.3 Adjunction to VP and the "relational" nature of distributivity

Williams (1991) describes (and criticizes) the Heim et al. view of distributors as one that treats distributivity as a property of NPs, rather than as a relation between two entities. However, the semantic translation of NP-adjoined distributors (repeated below) reveals them to be binary relations between NPs and VPs:

(2.12a) $\mathrm{D}_{<\mathrm{e},<\mathrm{et},\mathrm{t}>>} \equiv \lambda N\,\lambda\varphi\,\forall x_j(x_j \cdot\!\Pi\; N)\,\varphi(x_j)$

In fact, the semantic translation (and arity) of the distributors of Heim et al. is not responsible for the problems presented in this section; these can be shown to be due solely to the position of distributors, which causes them to combine with the rest of the sentence in the order ((NP D) VP). Both the VP conjunction problem and the c-command problem are easily addressed if we replace NP-adjoined distributors with VP-adjoined ones with the same semantics (but, necessarily, taking their arguments in the opposite order):[25]

(2.47) $\mathrm{D}_{et,et>} \equiv \lambda\varphi\,\lambda N\,\forall x_j(x_j \cdot\!\Pi\; N)\,\varphi(x_j)$

This VP-adjoined distributor will combine in the order (NP (D VP)), allowing for the conjunction of distributive and non-distributive VPs. Since its semantics are the same as those of the original distributor in (2.12a), it cannot be said to be any more (or less) "relational."

The presence of a VP-adjoined distributor does not stop the subject from c-commanding into the VP, allowing the embedded reflexive in (2.48a) (=(2.43b)) to be bound by *Simon*

[25]There is a fair-sized literature on the question of whether distributivity is a property of NPs or of VPs. The issue, and the relevant literature, is reviewed in detail by Lasersohn (1995). (He adopts the VP-adjunction alternative).

There seems to be some confusion regarding the origin of the argument from VP conjunction. Brisson (1998:p. 33) lists three different attributions in different sources, the earliest being Massey (1976). The relevant examples (but not the conclusion that distributivity must be a property of VPs) can be traced back even further. Lønning (1997) cites the following example, due to Hausser (1974):

(i) The horses gather and graze.

and Garfunkel. (It must be assumed that the intervening distributor does not destroy the locality of binding between the reflexive and its antecedent).[26] Example (2.44a) is given the translation shown in (2.48b), which allows the mixed group/distributive interpretation.

(2.48) a. Simon and Garfunkel [D praised themselves to each other].

b. They [$_{VP}$ [$_{VP}$ collided] and [D [$_{VP}$ criticized each other's driving]]]

The argument just presented uses the existence of conjoined VPs that admit a mixed group/distributive interpretation to argue for VP-adjoined distributors. The question then arises of whether there are examples involving conjoined NPs whose interpretation is similarly mixed; would such examples be incompatible with VP-adjoined distributors, supporting NP-adjoined distributors and leading us to a paradox? The answer is that such examples exist, but are compatible with VP-adjoined distributors. One such example (slightly modified here) is discussed by Schwarzschild (1996:p. 76): suppose that a number of computers were bought one at a time, and paid for in two installments each. A cartonful of diskettes was also bought and paid for in two installments. One might then truthfully say:

(2.49) The computers and the diskettes were paid for in two installments.

Here the conjoined NP *the computers* is interpreted distributively, but the NP *the diskettes* is interpreted collectively. Briefly, Schwarzschild's solution involves a VP-adjoined operator which distributes over the plural individual *the computers and the diskettes*, but not "down to individuals": the context allows the set of all diskettes is treated as a unit. I defer further discussion of this example to section 5.3, which is devoted to Schwarzschild's theory.

[26] Incidentally, floated *each* does block plural reference, as pointed out by Williams (1991:p. 166). This is true even when *each* is clearly adjoined to the VP, not the subject:

(i) They$_i$ probably each$_j$ saw themselves$_{j/*i}$.

VP-adjoined distributors, then, can account for the semantic relationships discussed in this section. These data were the basis of Williams's argument (back in the year 1991) that it was necessary to augment the traditional semantics of variable binding with a linking theory of the sort he proposed. But we see that the "relational" nature of distributivity can be captured in a compositional way within the Montagovian framework; the problem with the Heim et al. (1991a) account is due to the posited constituent structure, not to a lack of expressive power in the basic framework. (This insight is unlikely to be very surprising to adherents of Montague semantics).

2.3.2 Distributing without c-command

We now turn to a different type of challenge to the Heim et al. (1991a,b) analysis. I begin by presenting data that calls into question the analysis of dependent pronouns as variables bound by a distributive quantifier. Although I ultimately show that this issue can be resolved, such examples pose insurmountable problems for the part of their analysis that concerns reciprocals.

Heim et al. treat distributivity as quantification over the parts of a plural individual; the distributor carries its own index, indicating that the quantifier it introduces can bind variables in its scope. By thus allowing indices to multiply, the distributors introduced by Heim et al. can account for a multiplicity of readings in sentences containing plural pronouns. However, the variable binding mechanism implies that certain syntactic relations must hold, namely, that a distributor must c-command the variables it binds. It is generally agreed that this condition indeed holds (by both Heim et al. (1991a,b) and Williams (1991), among others), and indeed it is difficult to see how binding could work otherwise.

In (Dimitriadis 1999b), I presented examples that appear to call into question the claim that dependent pronouns are always bound by a c-commanding distributor. The analysis I propose there rescues this claim, but similar reciprocal constructions of this sort present additional difficulties for the scopal account of reciprocals. In chapter 4 I show that the evidence provided by such examples ultimately requires a rejection of Heim et al.'s "scopal" approach.

2.3.2.1 Non-reciprocal examples

Let us begin by reviewing some of the data considered by Heim et al. (1991a), and their conclusions about the possible readings of dependent pronouns. They claim five readings for example (2.50):

(2.50) John and Mary argue that they will win $100.

First of all, the main and the embedded predicates are, independently, two-ways ambiguous: John and Mary may be making an argument jointly or separately, and the argument may be that the two of them stand to win $100 jointly or separately. There are four combinations of collective and distributive modes for arguing and winning, which account for four independent readings and are represented as in (2.51a–d). Finally, there is the dependent reading, in which each of them argues a different proposition: John argues that he will win $100, and Mary argues that *she* will win $100.

(2.51) a. John and Mary_1 argue that they_1 will win $100.

b. John and Mary_1 argue that [they_1 D_2] will win $100.

c. [John and Mary_1 D_2] argue that they_1 will win $100.

d. [John and Mary_1 D_2] argue that [they_1 D_3] will win $100.

e. [John and Mary_1 D_2] argue that they_2 will win $100.

Next, Heim et al. turn to example (2.52), which they show to be only four-ways ambiguous; the dependent fifth reading is missing.

(2.52) The student that John and Mary taught argued that they would win \$100.

According to Heim et al., the dependent reading is missing because *John and Mary* occurs in a scope island, and thus its distributor does not c-command the pronoun *they* (and cannot be raised to a position where it does).

On the other hand, sentence (2.53a) is known to license the dependent reading. Heim et al. (1991a:p. 90) conclude that the possessive pronoun, along with its adjoined distributor, undergoes QR to adjoin to the containing NP. From there the possessive-distributor complex c-commands the dependent pronoun (and the reciprocal), as shown in (b).[27]

(2.53) a. Their coaches think they are faster than each other.

b. $[_{NP}$[their$_1$ D$_2$] [e coaches]] [think they$_2$ are-faster-than [each$_2$ other]]

The contrast between (2.52), which does not allow the dependent reading, and (2.53a), which does, is thus attributed to whether or not the intended antecedent of the dependent pronoun appears inside a scope island. But some reflection on what sentence (2.52) *could* mean, if it had the dependent reading, suggests that the problem is elsewhere: there is a single student, taught separately or jointly by John and Mary (depending on the presence or absence of an embedded distributor). This one student argued a single thing: that jointly or separately, John and Mary would win \$100.[28] This insight does not contradict our earlier analysis; in fact, it highlights the fact that a distributor cannot be applied to the complex NP that serves as the matrix subject, since it is singular.

Surprisingly, the dependent reading turns out to be contingent *only* on the plurality of the matrix subject, not on the c-command requirement: for many speakers, the missing

[27] As Heim et al. (1991b:fn. 3) note, this explanation becomes unavailable in their revised account, which does not allow QR of distributed-over NPs (or of reciprocal antecedents). They do not suggest an alternative treatment.

[28] There are constructions that allow a relative clause with a singular head to be interpreted as ranging over multiple individuals, but this is not one of them; see section 3.2 for details.

reading is immediately recovered if we substitute a plural number of students:[29]

(2.54) The students that John and Mary taught argued that they would win $100.

Here we have a reading in which John taught one or more students who argued that he would win $100, and Mary taught one or more different students who argued that she would. (There are additional ambiguities involving whether the students were taught individually or en masse, but they need not concern us).[30] Evidently the dependent reading does not require c-command between the dependent pronoun and its intended antecedent (or an adjoined distributor): the intended antecedent *John and Mary* is just as deeply embedded in example (2.54) as it was in (2.52). The essential element is the presence of the plural head noun *students,* and indeed, sentence (2.53a) loses the dependent reading if we substitute a singular subject, as in (2.55).[31] The only possible reading is the anomalous one, according to which a single coach believes contradictory things about her athletes.

(2.55) # Their coach thinks they are faster than each other.

When we try to represent the dependent reading of sentence (2.54), things get difficult. Informally, it seems that there should be a distributor over *John and Mary,* as shown in

[29]These examples, and the reciprocal variants discussed in the next section, were first presented in Dimitriadis (1999b).

[30]The system of Heim et al. does not include distribution over an object. For example, sentence (i) could say that John taught a class with five students in it, or that he taught five different students on possibly five different occasions. The second reading could be expressed in the spirit of their analysis by adjoining a distributor to the object and allowing it to QR over the VP, but a number of adjustments to the semantics would be required. See Brisson (1998:p. 116ff) for discussion of distribution over objects.

(i) John taught five students.

[31]Another example that does *not* allow the dependent reciprocal reading, this one from Williams (1986), is the following:

(i) People that know them say they like each other.

The head of the relative clause in this example is plural, but its indefiniteness blocks the dependent reading. Such factors are discussed in more detail in section 3.1.

(2.56), that can somehow bind the dependent pronoun. But a distributor adjoined to *John and Mary* is clearly too deeply embedded; even if some mechanism were assumed that could raise the NP-distributor complex out of the relative clause (a strong island), we would have no explanation why the plurality of the head noun is essential: an explanation along these lines would also predict that (2.52) should be grammatical.

(2.56) The students that [[John and Mary D_1] taught] argued that they_1 would win $100.

Perhaps there could be distribution over the complex NP *the students that John and Mary taught*, causing the matrix subject and the embedded pronoun *they* to somehow vary together. But the mechanism of Heim et al. cannot capture this relationship: in their system, dependent readings always involve pronouns bound by coreferring NPs. A distributor adjoined to the complex subject, as in (2.57), will range over students, while the dependent pronoun ought to range over John and Mary.

(2.57) [The students that John and Mary taught D_1] argued that they_1 would win $100.

We might try to insert a distributor at a position where it c-commands *they*, and co-index it with a distributor over *John and Mary* as follows:

(2.58) [The students that [John and Mary D_1] taught D_1] argued that they_1 would win $100.

While the above representation might suggest that *they* is dependent on the distributor adjoined to *John and Mary*, semantically the coindexation is meaningless. The embedded distributor (i.e., quantifier) will bind any variable in its scope with the index *1*, and outside its scope any coindexed variables will be treated just as if they carried a different index. Even assuming that some kind of absorption-like mechanism could achieve the intended effect, the resulting structure would coindex members of the set *John and Mary* with members of the set of their students, and could not express the difference between the students winning and John and Mary winning. Besides, assuming there are two students this structure would generate four different instances of arguing that someone would win $100, and we should only get two.

If we suppose that the matrix subject receives just the index *1* and not a distributor, we have perhaps better semantics but the wrong structure, since the only distributor is still too deeply embedded to bind either the matrix subject or the dependent pronoun *they.* The final option (and perhaps the best this system can do) is to adjoin the distributor to the matrix subject, and bind both *John and Mary* and *they* to this distributor:

(2.59) [The students that [John and Mary]$_1$ taught D$_1$] argued that they$_1$ would win $100.

This structure suggests that all the NPs involved should vary together; but even if we could provide a semantics for treating the indexed NP *John and Mary* as some kind of bound variable, we would still need a way to express the fact that *they* should range over *John and Mary,* not over the entire matrix subject. Such an analysis could be developed by employing choice functions (cf. Romero 1999), but will not be pursued here.

The problem here is not limited to the NP-adjoined distributor analysis of Heim et al.: the dependent pronoun cannot be c-commanded by its intended antecedent not because of an intervening distributor, but because the antecedent is embedded in a scope island. (In fact, the intended binder of the dependent pronoun is the distributor itself, not the distributed-over NP). The same problem arises with VP-adjoined distributors; for example, structure (2.59) would be as in (2.60) if VP-adjoined distributors were used. Clearly, this is no improvement since once again, the NP *John and Mary* does not c-command the pronoun *they,* and cannot raise to a position from where it does.

(2.60) [The students that John and Mary [D_1 taught]] [D_1 argued that they_1 would win $100.]

Having despaired of handling this construction with distributors, one might be tempted to conclude that the pronoun *they* is not bound after all, but is a referential pronoun referring to the plural individual *John* $\oplus$ *Mary*. In this case the question of binding it clearly does not arise; the fact that certain students argued that John would win and others argued that Mary would win is incidental, perhaps a pragmatic subentailment. In this way we would have given a *cumulative* analysis to this sentence. The problem is that such an analysis would not only be inconsistent with the theory of Heim et al. (who argue that the dependent pronoun in sentence (2.50) is in fact bound); it is contradicted by a body of evidence which I present in section 3.3, after a more thorough look at dependent pronouns in general. Until then, I will just take it for granted that the embedded pronoun does depend on (i.e., covary with) its intended antecedent.

In section 3.5 I propose that the dependent pronouns in such sentences should be treated as functional "paycheck" pronouns,[32] represented as proposed by Engdahl (1986). Specifically, the pronoun *they* in sentence (2.54) is interpreted as the function $F(x)$ that maps each student x to the person (either John or Mary) that taught x. This means that the argument variable ranges over students, not teachers, and therefore it can be bound by a distributor that ranges over the entire matrix subject:

(2.54) The students that John and Mary taught argued that they would win $100.

(2.61) [[The students John and Mary taught] D_1] argued that $F(x_1)$ will win $100.

In effect, translating the pronoun in this way makes (2.54) synonymous with:

[32]The term "paycheck pronoun" originates from the following type of canonical example.

(i) The man who put his paycheck in the bank was wiser than the man who fed it to his dog.

(2.62) The students that John and Mary taught argued that their teacher(s) would win $100.

2.3.2.2 Dependent reciprocals without c-command

The findings of the last section also apply to reciprocal sentences, which impose the same conditions on the "long distance reciprocal" (i.e., dependent) reading. Example (2.63a) has a non-contradictory, dependent reading, according to which John thinks he is taller and Mary thinks *she* is taller; example (2.63b) lacks the dependent reading, which is restored in the plural-headed example (c). The latter says that John's lawyer thinks that John will sue Mary, and Mary's lawyer thinks that Mary will sue John.

(2.63) a. John and Mary think they are taller than each other.

b. The lawyer who represents John and Mary thinks they will sue each other.

c. The lawyers who represent John and Mary think they will sue each other.

Such examples are difficult to process since they involve the double complexities of dependent reciprocation and anaphora without c-command. Some speakers do not accept dependent readings of the sort discussed in the last section at all, but for the most part, those who do also accept the reciprocal examples to which we now turn. Judgements are further complicated by the resemblance in truth conditions between the dependent and independent readings. In this section I will rely on a small set of examples presented out of context for ease of exposition; section 4.3 reviews a wider variety of examples in order to establish the existence of such readings.

Although the paycheck pronoun analysis sketched at the end of section 2.3.2.1 makes it possible to account for the non-reciprocal examples discussed there, reciprocal sentences present additional challenges: According to the scopal analysis, the dependent reading of sentence (2.63c) is generated through binding the reciprocal by the matrix distributor:

(2.64) [The lawyers that represent John and Mary D_1] think $F(x_1)$ will sue each other.

The problem is that the distributor D_1 ranges over lawyers, not clients: according to the scopal analysis, so should both the contrast and the range arguments of the reciprocal. Such a reading would say that John's lawyer thinks that John will sue Mary's lawyer (i.e., every lawyer other than his lawyer, or worse, every lawyer who is not John), and vice versa; this reading is impossible, and is rightly ruled out by the Principle A requirement that is part of the Heim et al. account. However, this system cannot generate the reading that does exist, in which the reciprocation is between clients, not lawyers.

The examples in (2.65) raise a similar problem, though without involving paycheck pronouns. Under its dependent reading, sentence (2.65b) says that John thinks his mother likes Mary's *mother,* and vice versa. In other words, the range of the reciprocal is the set of mothers. Again, the system of Heim et al. simply predicts that this reading is impossible. If it could somehow be relaxed to ignore the mismatch in indices, it might generate the LF shown in (c). This says that the range argument of the reciprocal should be the set consisting of John and Mary, not the set of mothers.

(2.65) a. The lawyers who represent John and Mary think that their clients will sue each other.

b. John and Mary think their mothers like each other.

c. [John and Mary D_i] think their_i mothers like each_i other.

In such examples the local antecedent of the reciprocal is a dependent pronoun that ranges over a set different from the matrix subject. Because the scopal analysis considers the matrix subject to be the true semantic antecedent of the reciprocal, it predicts that the matrix subject should be used as the reciprocal's range argument. But in fact the range argument always matches the local antecedent of the reciprocal, which in these examples is the embedded subject. I discuss these issues in greater detail in chapter 4, and I ultimately conclude that they force us to essentially reject the scopal analysis of dependent reciprocals.

2.3.3 Chained multiple reciprocals

The original, *each*-raising account of Heim et al. (1991a) introduced "absorption" of multiple distributors into one, in order to account for sentences containing multiple reciprocals with the same antecedent. Absorption was described in section 2.1.7.3, where it was used to account for the following sentences:

(2.27) a. John and Mary read each other's books in each other's languages.

b. John and Mary told each other that they love each other.

c. They gave each other pictures of each other.

The relevant reading of (a) says that John read Mary's book in Mary's language, and Mary read John's book in John's language. The relevant reading of (b) says that John told Mary that he loves her, and of (c) that John gave Mary a picture of Mary, and vice versa. The *each*-raising account predicts that the *each* part of both reciprocals should raise to adjoin to the matrix subject, giving a configuration like (2.66); the two copies of *each* then "absorb" into a single distributor in order to give the correct semantics.

(2.66) They gave each other pictures of each other.

The absorption mechanism is rendered unnecessary in the revised "*each*-binding" account of Heim et al. (1991b) which, instead of raising *each*, binds it in situ by an independently inserted distributor. In the above examples, the distributor binds the contrast argument of both reciprocals; no additional mechanism needs to be stipulated.

Heim et al. (1991a) only addressed the readings just discussed, but Williams (1991) points out that the sentences in (2.27) have additional readings not accounted for by Heim et al. For example, one reading of sentence (2.27c) says that John gave Mary a picture of John, and vice versa, and can be diagrammed as in (2.67).

(2.67) They gave each other pictures of each other.

These readings seem to require the higher reciprocal to function as the antecedent of the lower one, a configuration that Heim et al. (1991a) did not anticipate. I will refer to such readings as *chained reciprocal* constructions. The following example (also due to Williams) only allows the chained reciprocal reading, lacking a reading with both reciprocals linked to the matrix subject:

(2.68) John and Mary want each other to like each other.

a. John wants Mary to like John, and Mary wants John to like Mary.

b. * John wants Mary to like Mary, and Mary wants John to like John.

The reason for the ungrammaticality of reading (b), of course, is that reciprocals are anaphors, and the matrix subject is not close enough to the second reciprocal. (This can be seen by the ungrammaticality of *They want me to like each other*).[33]

In their reply to Williams, Heim et al. (1991b) attempt, unsuccessfully as I will show, to provide a way of accounting for chained reciprocal readings. In order to syntactically license and semantically derive this configuration, it is necessary for the second reciprocal to depend on the first, so that two applications of contraindexing cause the second reciprocal

[33]Williams claims that the Heim et al. (1991a) absorption mechanism overgenerates in the case of sentence (2.68), predicting a reading with both reciprocals linked to the matrix subject. Since Heim et al. assume that quantifier raising is not clause-bound in general, Williams claims that the theory of Heim et al. cannot rule out reading (2.68b). However, Heim et al. also stipulate that the reciprocal (specifically, the movement trace of the *each* part of the reciprocal) must be locally bound by an antecedent that bears the index of the reciprocal's distributor. This condition successfully rules out the illicit reading, as Heim et al. (1991b) point out.

to covary with the matrix subject.[34] Heim et al. (1991b) adopt this point of view, which corresponds to Williams's own linking configuration for this sentence, and represent these dependencies via the indexing scheme in (b):

(2.69) a. They want each other to like each other.

b. [They$_1$ D$_2$]$_2$ want [each$_2$ other]$_3$ to like [each$_3$ other]$_4$.

Heim et al. argued for independent reasons that the [*e other*] part of the reciprocal (or, in the non-movement (1991b) version, the entire reciprocal) is an R-expression that bears its own index; so it is not unreasonable that the first reciprocal could be the antecedent of the second. Heim et al. propose that "the implicit quantificational force of the *other*-NP be represented by a [distributor] D", so that a reciprocal is a suitable antecedent for another reciprocal. The indexing of (2.69) is then rewritten as (2.70a), for which they give the semantic translation (b):[35]

(2.70) a. [They$_1$ D$_2$]$_2$ want [[each$_2$ other] D$_3$]$_3$ to like [[each$_3$ other] D$_4$]$_4$

b. $\forall x_2(x_2 \ \Pi \ X_1)\forall x_3(x_3 \ \Pi \ X_1 \ \& \ x_3 \neq x_2)\forall x_4(x_4 \ \Pi \ X_1 \ \& \ x_4 \neq x_3)$

$[x_2$ wants x_3 to like $x_4]$

This translation correctly expresses the semantics of the intended reading. The problem is that the semantics of Heim et al. does not actually generate it: The variable x_4, introduced by the second reciprocal, is written as taking values among the parts of X_1; but the indexing scheme of (2.70a) does not license this substitution. Since the contrast argument x_3 of the second reciprocal is bound by the first reciprocal's distributor, D_3, their algorithm dictates

[34] In chained reciprocals whose matrix subject contains more than two members, the second reciprocal will range over more than the matrix subject.

[35] Heim et al. (1991b) actually illustrate these examples with the sentence *They gave each other pictures of each other,* under the reading saying that John gave Mary a picture of John (and vice versa). They consider this reading to be fully parallel, in all relevant respects, with the reading of (2.68) under discussion (and I agree).

that the range argument must be the sister of the distributor D_3: that is, the range argument of the first argument should be the translation of the reciprocal [*each*$_2$ *other*], not X_1. It is not immediately clear what individual entity it should translate into, but since *each*$_2$ is bound by the distributor D_2, the interpretation of [*each*$_2$ *other*] must depend on x_2. The most coherent interpretation is that [*each*$_2$ *other*] stands for the sum individual formed of all x such that $x \ \Pi \ X_1$ and $x \neq x_2$, which means that the second reciprocal is left to range over all values in X_1 that are neither equal to x_2 nor to x_3.[36]

It should be clear that this translation cannot generate the desired reading of (2.68), since what is desired is for the second reciprocal to take the same value as x_2. Moreover, when the matrix subject X_1 consists of just two atomic individuals, this translation would leave the second reciprocal with nothing to range over! Depending on how many individuals are involved, the resulting truth conditions resemble those of sentences (2.71a) or (b).

(2.71) a. Each of them wants the others to like each other.

b. * Each of them wants the other to like each other.

Could the problem be addressed by relaxing the rules for range-set assignment, allowing the lower reciprocal to be arbitrarily assigned the proper range argument, X_1? The behavior of reciprocals in general leaves no doubt that this is impossible; Heim et al. require that the range set be the sister of the distributor binding the reciprocal not for theory-internal reasons, but out of descriptive necessity. Sauerland (1995b) provides the following example (based on the one already discussed as (2.25), from Heim et al. (1991a)) to support the need for such a condition:

[36]Heim et al. do not provide revised translations for the reciprocal, beyond their statement that it would be enough that "the implicit quantificational force of the *other*-NP be represented by a [distributor] D" (p. 190); if we give D_3 the usual translation for distributors and work backwards from the intended semantics for reciprocals, we find that D_2 should somehow range over all parts of X_1 other than x_2, as I have stated.

(2.72) The women told the youngest three of them to give lectures to each other.

This sentence can only be understood as saying that the lectures were given to the other two of the three youngest women, not to all the women. The local binder of the reciprocal is the embedded subject PRO, so the only possible contrast set is the antecedent of PRO, *the youngest three (of the women).* An account that allows the range argument to be set "from the context" must also allow for the possibility that the context provides a range argument other than the sister of the reciprocal's binder, but in fact this never happens. Thus the reciprocal's range argument must be restricted to being the sister of the distributor that binds it, even if only by stipulation (as in the Heim et al. account).

In order to derive the correct translation of (2.68), then, it is necessary for the lower reciprocal to somehow have the same range set as its antecedent, the higher reciprocal; but compositional translation of the higher reciprocal causes it to be interpreted as a singular bound variable (or at best, as smaller than the matrix antecedent). This is just the "long-distance reciprocal" problem all over again: the lower reciprocal needs a plural contrast argument, but the only such antecedent is too far away. In the case of examples like *John and Mary think they like each other,* the scopal analysis claims that the reciprocal simply looks further for its antecedent; the contraindexation requirements of the lower reciprocal in (2.68) (under the desired reading) unambiguously indicate that it needs to depend on its local antecedent, i.e., on the higher reciprocal—not on the matrix subject.

Heim et al., as we have seen, do not treat these two types of examples as requiring similar analyses. Williams's proposal assigns them parallel structures, thereby expressing the similarity of the semantic relationships involved. Unfortunately, Williams's system of links sheds no light on how the links in question could be semantically interpreted. His proposal is briefly visited in the next chapter.

2.3.4 The determination of reciprocal range

In case the reader is still favorably inclined toward the idea of allowing the range argument of the reciprocal to be more freely determined, let us consider the implications of such a move for the scopal analysis of dependent reciprocals. The need for non-local movement (or binding) of the reciprocal, it turns out, is a consequence of the reciprocal's need for a plural range argument. If we were allowed to arbitrarily write in the correct range argument, the dependent reading of sentence (2.1b) could be expressed without long movement of the reciprocal: We could simply raise the reciprocal and adjoin it to the *dependent,* and hence singular, pronoun *they,* as in (2.73a). (In the *each*-binding variant of Heim et al. (1991b), we would adjoin to *they* a dummy distributor that binds the reciprocal's contrast argument).[37] Our compositional semantics would then give the translation (2.73b), which has the correct semantics as long as we are allowed to give to the open variable X the value John $\oplus$ Mary.

(2.1b) John and Mary think they like each other.

(2.73) a. [John and Mary D$_i$] think [they$_i$ each$_j$] like e$_j$ other.

b. $\forall x_i(x_i \,\dot{\Pi}\, J{\oplus}M)$ think$(x_i, {}^\wedge[(\forall x_j \,\dot{\Pi}\, x_i)\,(\forall x_k \,\dot{\Pi}\, X)\, x_j \neq x_k \Rightarrow$ like$(x_j, x_k)])$

Such an analysis has not been advocated by anyone, as far as I know. The reason is that the choice of translation for the range argument of the reciprocal (represented by the variable X) is completely arbitrary: such a mechanism would suggest that other choices

[37] Either variant requires that we allow distributors to "distribute" (trivially) over a singular individual. The account of Heim at al. actually forbids this: the operators $\dot{\Pi}$ and Π are defined as *proper-atomic-part-of* and *proper-part-of*, respectively, ruling out distribution over anything that lacks proper parts, i.e., over atomic individuals. But this is a technical property of their definition of distributors, specifically chosen to model the need for a plural antecedent for the reciprocal; it is not in itself motivation for non-local movement or binding of the reciprocal. Arguably, a theory that allows distributors to gratuitously distribute over singular variables is simpler than one that needlessly stipulates a distinction. For example, Kamp and Reyle (1993) make the uniformity of plural and singular processing rules one of their research goals, and explicitly allow operators defined for plurals to be applied, generally without visible effect, to singular antecedents as well. (See especially the discussion on pp. 394ff). Sternefeld's (1998) algebraic treatment of reciprocal and cumulative sentences defines freely insertible operators that, although he does not discuss the issue, are compatible with singular as well as plural operands.

of range argument are possible, given the appropriate pragmatic context. But in fact the reciprocal's range argument is strictly determined by its antecedent. In the account of Heim et al., this is reflected in the stipulation that the range argument must match the set over which the binder of the reciprocal ranges over. In a dependent reciprocal sentence, the local antecedent is a singular bound variable, and the closest antecedent for the range argument is the matrix subject. This provides the motivation for claiming that the reciprocal is indeed in a binding relationship with the matrix subject, i.e., for the scopal analysis of reciprocals.

In this way the restricted interpretation of the reciprocal's range argument is crucial to the case for non-local movement or binding of the reciprocal. A theory that is too flexible on setting the range argument does not need to bother with non-local movement at all: it would be enough to allow the reciprocal's dummy "distributor" to attach to a bound variable, and then magically set the range argument to the proper value. In other words, we have no need for the scopal analysis unless we adopt a highly constrained algorithm for determining the reciprocal's range argument. It follows that we cannot simply discard this algorithm in order to keep the scopal analysis from making the wrong predictions in the examples of the last two sections. The remainder of this dissertation largely revolves around the task of developing a sufficiently constrained account for the full range of dependent reciprocal constructions.

Chapter 3

Readings of dependent pronouns

In the dependent reading of a sentence like (3.1), the dependent pronoun is understood as taking values that covary with the parts of the matrix subject: John thinks that John is rich, and Mary thinks that Mary is rich.

(3.1) John and Mary think they are rich.

It is therefore natural to treat dependent pronouns as variables bound by a quantifier that distributes over their antecedent. This approach is explicit or implicit in most accounts of distributivity, including those of Heim et al. (1991a,b), Sauerland (1995a,b), Schwarzschild (1996), and Sternefeld (1998). But this point of view is called into question by examples, such as those presented in (Dimitriadis 1999b) and already reviewed in section 2.3.2, that allow the dependent construal without the need for c-command between the pronoun and its intended antecedent.

In this chapter, I focus on the properties of dependent pronouns in distributed non-reciprocal sentences. I will ultimately defend their analysis as bound variables, showing that we can account for apparent binding without c-command via a functional (paycheck) analysis of the dependent pronoun. I will demonstrate the function-like behavior of dependent pronouns by examining a variety of constructions that show systematically limited interpretive options, consistent with a functional interpretation of some sort.

A possible alternative is to interpret sentences containing dependent pronouns *cumulatively:* dependent pronouns are translated as referential plural pronouns, and it is left to the pragmatics to sort out the match between the parts of the dependent pronoun and the parts of its antecedent. Although such an analysis trivially explains the existence of dependent pronouns without c-command, in my view it is mistaken. In section 3.3 I present several arguments against the cumulative analysis. Among them are included the arguments of Heim et al. (1991a) against Higginbotham's (1985) cumulative-style analysis, which I extend to the non-c-command examples under consideration.

Because the older examples discussed here were introduced by Williams and Heim et al. in the context of studies of reciprocity, several of them involve reciprocals. In this chapter the presence of reciprocals is considered incidental: the central phenomenon is the relation between the dependent pronoun and its antecedent, and the dependent reading never disappears when we switch from a reciprocal to an equivalent non-reciprocal example. Reciprocals return to center stage in chapter 4, where the account of dependent pronouns developed in this chapter will be applied to dependent reciprocal sentences.

3.1 Dependent pronouns and binding

As discussed in section 2.1.4, Heim et al. (1991a) treat distributors as NP-adjoined operators that introduce quantification over the parts of the NP they adjoin to, and can bind a variable in their complement. The distributor D can be written as in (3.2). Sentence (3.1) is given the structure (3.3a), which leads to the translation in (b).

(3.2) $\mathrm{D}_{<\mathrm{e},<\mathrm{et},\mathrm{t}>>} \;\equiv\; \lambda N\, \lambda\varphi\, \forall x_j (x_j \cdot\!\Pi\, N)\, \varphi(x_j)$ (= (2.12a))

(3.3) a. [John and Mary_1 D_2] think [they_2 are rich.]

b. $\forall x_2(x_2 \cdot\Pi\ J{\oplus}M)$ think$(x_2,$ ^[rich(x_2)]$)$

This structure suggests that a distributor must c-command a pronoun that ranges over the parts of the distributed antecedent, and Heim et al. (1991a) argue that this is indeed the case. Sentence (3.4a) has several readings (Heim et al. count five), including the dependent reading, according to which John thinks he will win \$100 and Mary thinks *she* will win \$100.[1] Sentence (b) lacks the dependent reading, a fact that Heim et al. take as evidence that c-command is required between the pronoun and the distributor associated with its antecedent. These findings carry over to reciprocal sentences, which impose the same conditions on the dependent "long distance reciprocal" (i.e., dependent) reading. Thus example (3.5a) has a non-illogical, dependent reading, according to which John thinks he is taller and Mary thinks *she* is taller; but examples (3.5b,c) lack the dependent reading, and can only be understood as claiming something illogical.

(3.4) a. John and Mary think they will win \$100.

b. The student John and Mary taught argued that they will win \$100.

(3.5) a. John and Mary think they are taller than each other.

b. The guy who saw John and Mary thinks they are taller than each other.

c. People that know them say they are taller than each other.

Heim et al. attribute the contrast between (3.4a)/(3.5a), which allow the dependent reading, and (3.4b)/(3.5b), which do not, to whether or not the dependent pronoun and the reciprocal are c-commanded by a distributor adjoined to its intended antecedent. But this conclusion appears to be an artifact of the limited evidence studied. The problem with (3.5b), it turns out, is simply that there is a single guy, who necessarily argued a single, irrational thing: that John and Mary are taller than each other. In other words the matrix predicate has a

[1]The Heim et al. treatment of sentence (3.4a) was also the subject of section 2.3.2.1; a good part of that discussion is repeated here.

singular subject, and so its complement can only be asserted once; the dependent reading necessarily involves a plural subject.

This much is consistent with, and indeed follows from, the analysis of Heim et al. (1991a): since in their system dependent readings involve binding of the dependent pronoun by a distributor, and since they allow distribution over plural NPs only, it follows that the dependent reading is only possible when the dependent pronoun has a plural antecedent. But let us return to the other prediction of this analysis, the c-command requirement that Heim et al. appeal to in their own explanation of the ungrammaticality of (3.4b) and (3.5b). Surprisingly, the dependent reading does *not* always require that the dependent pronoun be c-commanded by its antecedent: For most speakers the missing reading of (3.4b) is immediately recovered if we substitute a plural number of students, as in (3.6a); similarly (if with some more difficulty), as we go from (3.5b) to (3.6b).

(3.6) a. The students John and Mary taught think they will win $100.

b. The guys who saw John and Mary think they are taller than each other.

Since Heim et al. consider dependent pronouns to be variables bound by a c-commanding distributor, these examples are entirely unexpected. Let us look more closely at the empirical conditions that determine the acceptability of the dependent reading. The dependent reading of sentence (3.6b) requires that John and Mary were each seen by a different guy (or guys), and that the guy who saw each one thinks that he or she is the taller of the two. The reading depends on our grasp of the match-up between the guys on one hand and John and Mary on the other, and consequently it is much easier to "get" such constructions when a natural one-to-one relationship between definite sets is involved. The following examples, which assume pairings between coaches and runners and between lawyers and clients, respectively, are much easier to process under the dependent reading.

(3.7) a. The coaches that trained them think they will win.

b. The lawyers that represent John and Mary think they will refuse to settle.

In example (3.7a) the subject of the embedded clause, the dependent pronoun *they,* ranges over runners; the only antecedent referring to the set of runners is the pronoun *them,* which is embedded in the relative clause modifying *coaches* and cannot be raised to a position that c-commands the dependent pronoun.

The reader may verify that it is also possible in the above examples to imagine multiple coaches for each runner, or multiple lawyers for each client, as long as no lawyer or coach is associated with multiple sides. Sentence (3.7a) can also be understood in a context in which each coach has trained several runners, as long as each coach expects all of his or her runners to win, and no runner was trained by more than one of the coaches under consideration. In short, what is required is a pairing between, e.g., coaches and runners; there may be multiple individuals in each group, as long as there is no overlap between the groups.

The analysis assumed so far cannot express the dependent reading of the above sentences. According to Heim et al., the distributive reading of sentence (3.4a) is expressed by translating an NP-adjoined distributor as a quantifier that ranges over the parts of the subject NP, and binds a variable representing the embedded pronoun:

(3.8) [John and Mary D_i] think they$_i$ will win $100.

Such an analysis cannot be applied to a sentence like (3.7a): a distributor adjoined over the matrix subject as in (3.9a) would range over coaches, not runners, while a distributor adjoined to the pronoun *them* (the intended antecedent of *they*), as in (3.9b), would be too deeply embedded to c-command the pronoun *they:*

(3.7a) The coaches that trained them think they will win.

(3.9) a. [The coaches that [e trained them] D_i] think $they_i$ will win.

b. [The coaches that [e trained [them D_i]]] think $they_i$ will win.

Since distributors have quantificational force, Heim et al. (1991a) claimed that they undergo QR, along with their adjoined NP; this was in fact the explanation given to examples like (3.10), which was known to allow the dependent reading:

(3.10) a. Their coaches think they will win.

b. $[_{NP}$ [Their $D_i]_i$ [t_i coaches]] think $they_i$ will win.

Since Heim et al. believed that dependent pronouns cannot find their antecedent inside a relative clause, this account hinged on the distinction between possessive NPs, which can be raised out of the NP they modify, and relative clauses, which constitute a strong island and are therefore expected to block QR. This means that the dependent reading of sentences like (3.7a) cannot be explained by raising an embedded distributor out of the relative clause: First, because such movement is highly implausible given that relative clauses are strong islands for extraction; and second, because if such raising were possible, it should also be possible out of relative clauses with *singular* head nouns, as in sentence (3.4b). But such sentences do not in fact allow the dependent reading; even if raising out of a relative clause were syntactically plausible, then, such raising could not explain why the dependent reading is contingent on the plurality of the head noun, i.e., it could not explain the contrast between the plural-headed (3.6a,b), which allow the dependent reading, and the singular-headed (3.4b)/(3.5b), which do not. (See section 2.3.2 for more discussion).

3.1.1 Towards a solution

If dependent pronouns are bound variables, our semantics provide no straightforward way to represent dependent readings without c-command. Our semantic framework requires that all forms of semantic covariation must involve a variable in the scope of some quantifier, which is tantamount to requiring that at *some* point in the derivation, the variable must appear in a semantic argument of a constituent containing the quantifier.[2]

There are several possible approaches to accounting for the examples considered here. First, we could argue that dependent pronouns are not really bound, assuming a *cumulative* analysis instead, and attribute the correspondence between parts of the pronoun and its antecedent to the pragmatics. Second, we could provide a way for the distributor to bind out of the relative clause, along the lines of the "functional relative clauses" studied, among others, by Sharvit (1999). Third, we can give the dependent pronoun a functional analysis, as I propose in section 3.5; this allows the bound variable in its translation to be bound by something that does c-command it. Finally, we could explore the potential of some novel mechanism for expressing dependencies between the pronoun and its antecedent. I reject the first two of these strategies in sections 3.3 and 3.4; in section 3.5 I argue in favor of the third, in the form of a proposal that gives dependent pronouns of this sort a paycheck-type translation along the lines proposed by Engdahl (1986).

3.1.2 More types of dependent pronouns

In order to place the above examples in context, we must consider a somewhat wider range of constructions. Sentence (3.11b), an example from Williams (1986:p. 281), was also given as evidence that the dependent reading requires c-command.

(3.11) People that know them say they like each other.

I have shown that c-command is not in fact necessary, and that the absence of the dependent reading of (3.4b) is due to the fact that its subject is singular; but in this case the subject

[2]We will leave aside the issues raised by Hintikka's (1999) conceptual separation of binding scope from c-command domain.

is plural, so another explanation must account for the absence of a dependent reading for (3.11). In view of my earlier remarks on the interpretation of sentence (3.7a), it appears that the dependent reading is unavailable because the subject, being indefinite, cannot set up a unique mapping between the referent of *them* and a unique set of people who know them. (In other words, it does not have a unique *witness set*; see Szabolcsi and Zwarts (1993)).

We have seen that the plurality of the subject is necessary to the dependent reading. In the examples seen so far, complex NP subjects with singular heads were necessarily interpreted as singular. But this is not true in general: Example (3.12) has a dependent reading, under which Bill and Dave saw different persons. Similarly, the dependent reading is possible in example (3.13a):

(3.12) Bill and Dave saw someone they knew.

(3.13) a. John and Mary found the student who argued that they would win $100.

b. John and Mary found the students who argued that they would win $100.

Here the dependent reading is *not* possible if we take (3.13a) to describe a situation involving a single student; as before, such a student could only have argued a single thing, that John and Mary would win $100. But this sentence is also compatible with a situation in which John found a certain student who argued that John would win $100, and Mary found a *different* student who argued that Mary would win it. The same situation can also be described using sentence (3.13b).[3] On the other hand, we saw that sentence (3.4b) cannot be read as involving more than one student. The difference is that the complex NP in (3.13a) contains the dependent pronoun *they*, which is bound by a c-commanding distributor. Hence the denotation of the complex NP *the student who argued that they would win $100* depends on the value of the bound variable represented by *they*, giving us more than

[3]The dependent reading of (3.13b) can be used to describe any situation where one or more students argued that John would win $100, and one or more *different* students argued that Mary would win it (and, of course, John and Mary found these students).

one student as this variable ranges over *John and Mary.*

These brief remarks on the role of the head noun have probably raised more questions than they have answered. Fortunately, these issues have been examined in detail by Kamp and Reyle (1993:ch. 4), to whose analysis of plurals we now turn.

3.2 Kamp and Reyle (1993): Plurals in DRT

In their detailed study of plurals, Kamp and Reyle (1993) arrive at a number of general principles governing their interpretation, and a framework that implements them. Their analysis is within the framework of Discourse Representation Theory (DRT); a close look at DRT is beyond the scope of this work, as is, regrettably, a comparison of the relative merits of DRT and the standard semantics as frameworks for the analysis of plurals. I will limit my presentation to a very brief introduction to relevant aspects of DRT in the next section, followed by a description of Kamp and Reyle's treatment of plurals and a rough adaptation of it into the semantic framework assumed here. For the purposes of this discussion, I will simply assume that their findings can be expressed equally well in either framework.

3.2.1 A quick introduction to DRT

Discourse Representation Theory (DRT) provides a dynamic framework that expresses the semantics of embedded clauses as recursively embedded Discourse Representation Structures (DRSs). DRSs are built incrementally: as each part of a sentence or sequence of sentences is interpreted, its translation is added to the DRS being built. The precise rules for the order of interpretation need not concern us here, but elements higher in a tree are generally processed before the elements they c-command; hence the translation of a bound pronoun can refer to the translation of its binder. A new sentence does not in general begin a new DRS; its content is simply added to the current DRS. Embedded DRSs are introduced by particular sentence elements, such as quantifiers and negation.

Each DRS includes a set of *discourse referents* (written on the top line of the DRS) followed by a set of *DRS conditions*, which are the propositions representing the propositional content of the discourse or sentence fragment being interpreted. The discourse referents are formally open variables: a DRS is evaluated as true with respect to some model $\mathcal{M}$ if its discourse referents can be associated with entities in $\mathcal{M}$ that satisfy all the DRS conditions in the body of the DRS, including conditions that contain embedded DRSs.

A DRS condition may contain variables that are identified with any discourse referent that is *accessible* to that DRS condition, that is, introduced by either the current DRS or one of the DRSs that properly enclose it. For example, the sequence of sentences (3.14a) is given the DRS in (b). The first sentence introduces the discourse referents x and y, and the three DRS conditions that describe them. The subject of the second sentence introduces the discourse referent z, which is immediately identified with the referent x. The negation in the VP *doesn't own a Porsche* then introduces an embedded DRS, which must be evaluated as false if the DRS condition that contains it is to be evaluated as true. Its last DRS condition, *z owns u*, involves the variable z which is defined in the enclosing DRS. Because the NP *a Porsche* is indefinite, the discourse referent u that it introduces is placed in the current DRS, which in this case is the embedded DRS introduced by the negation operator; a definite appearing in the same position would have been added to the list of referents of the top-level DRS. Since the indefinite appeared inside an embedded DRS, it is not available for reference by subsequent clauses; for example, the pronoun in the continuation (c) can be translated as referring to the book y, but not to the Porsche u. (The translation of sentence (c) is added to the top-level DRS, since (c) is outside the scope of the negation operator).

(3.14) a. Jones likes Ulysses. He doesn't own a Porsche.

b.

$$\begin{array}{|c|} \hline \\ x\ y\ z \\ \\ \text{Jones(x)} \\ \text{Ulysses(y)} \\ \text{x likes y} \\ \text{z = x} \\ \neg \begin{array}{|c|} \hline u \\ \\ \text{Porsche(u)} \\ \text{z owns u} \\ \hline \end{array} \\ \hline \end{array}$$

c. It is very expensive.

While indefinites add discourse referents to the current DRS, definite NPs add theirs to the top-level DRS and are thereby accessible to subsequent pronouns. For example, the pronoun *it* in the following fragment can be understood as referring to *War and Peace.*

(3.15) John does not like *War and Peace. It* is very long.

Quantifiers and conditionals also introduce DRS conditions that contain embedded DRSs. It is well-known that pronouns in the *then*-clause of a conditional can refer to referents introduced in its *if*-clause, as in the donkey sentence (3.16a). DRT expresses this property directly, stating that the *then*-clause of a conditional *extends* the DRS established by its *if*-clause, and therefore has access to the variables introduced there. (On the other hand, the DRS of the antecedent does not have access to the variables of the consequent).[4] Accord-

[4]There are some counterexamples to the generalization that a pronoun in the antecedent cannot refer to entities introduced in the consequent. In his "remaining problems and unresolved issues" chapter, von Fintel (1994:p. 178, fn. 25) mentions the following examples, which he attributes to lecture notes by Irene Heim:

(i) Most letters are answered if they are shorter than 5 pages.
(ii) Few people like New York if they didn't grow up there.

ingly, (3.16a) corresponds to the well-formed DRS in (b).

(3.16) a. If Jones owns a book on semantics, he uses it.

b.

$$
\boxed{
\begin{array}{c}
x \\[1em]
\text{Jones(x)} \\[1em]
\boxed{\begin{array}{c} y \\[0.5em] \text{book on semantics(y)} \\[0.5em] \text{x owns y} \end{array}}
\;\Rightarrow\;
\boxed{\begin{array}{c} z\ w \\[0.5em] \text{z = x} \\[0.5em] \text{w=y} \\[0.5em] \text{z uses w} \end{array}}
\end{array}
}
$$

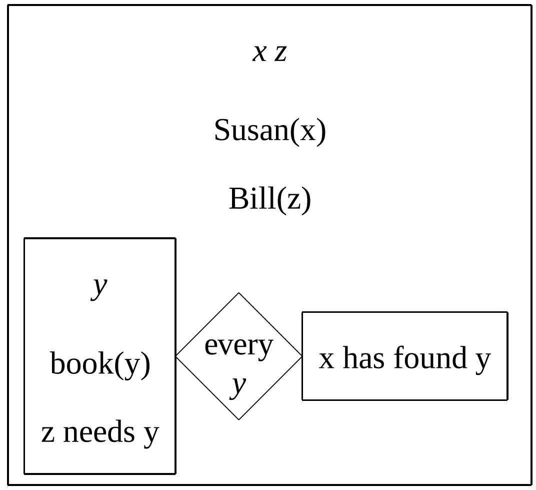
The truth conditions of a conditional DRS are those that apply to material implication: the conditional DRS is true if any choice of values for the discourse referents of its left-hand DRS that makes that part true can be extended to a choice of values that makes its right-hand DRS true as well.

Quantifiers similarly establish a “duplex condition” consisting, in addition to the quantifier and its variable, of an embedded DRS containing the conditions for the restrictor, and another containing the conditions being asserted (nuclear scope). Such DRSs represent generalized quantifiers, and have the truth conditions appropriate to each such quantifier. Again, the nuclear scope extends the DRS of the restrictor and therefore has access to its variables.

(3.17) a. Susan has found every book which Bill needs.

Such examples are rare, but they are a puzzle for most treatments of conditionals, including that of Kamp and Reyle.

b.



3.2.2 Plurals and dependent pronouns

We now come to plurals. Kamp and Reyle assume the plural ontology of Link (1983), which posits a lattice of atomic (singular) or sum (plural) "individuals," ordered by the part-of relationship (written in DRT using set notation). Singular and plural individuals are then in principle of the same type, but are distinguished by the DRS conditions *at(x)* and *non-at(X)* which are true of atoms and non-atoms, respectively. By convention, these conditions are normally suppressed in the DRSs Kamp and Reyle write; plural referents are indicated by using capital letters to name them, while singular referents are written with lowercase letters. (There are also referents that are unspecified for number, written with lowercase Greek letters).

A collectively interpreted plural sentence is translated just like a singular one; its subject just happens to be a non-atomic individual. Distributivity is represented by introducing quantification over the parts of the subject. Consider the following readings of sentence (3.18):

(3.18) The lawyers hired a secretary they liked.

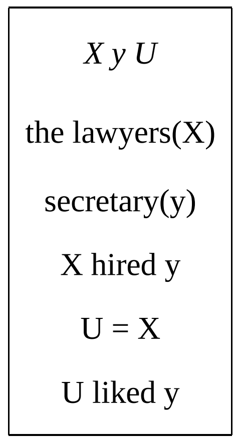
a. $\exists X$ lawyers(X) & X hired a secretary X liked.

b. $\exists X$ lawyers(X) & $(\forall x \in X)$ x hired a secretary x liked.

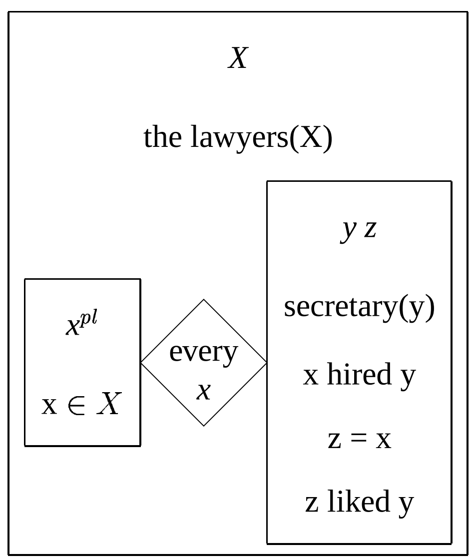
c. % $\exists X$ lawyers(X) & $(\forall x \in X)$ x hired a secretary X liked.

The collective reading (3.18a) is represented as in (3.19a). It says that the lawyers, as a group, hired a secretary whom they like (again, as a group). The distributive reading is expressed by introducing universal quantification over the parts of the subject; the embedded pronoun *they* can be identified with either x or X, as shown in (3.18b) and (c), respectively. The first of these readings says that each lawyer hired a secretary that he or she liked. This reading is, of course, what we have been calling the dependent reading; it is shown in (3.19b). (The superscript *pl* is explained on page 82).

(3.19) a.



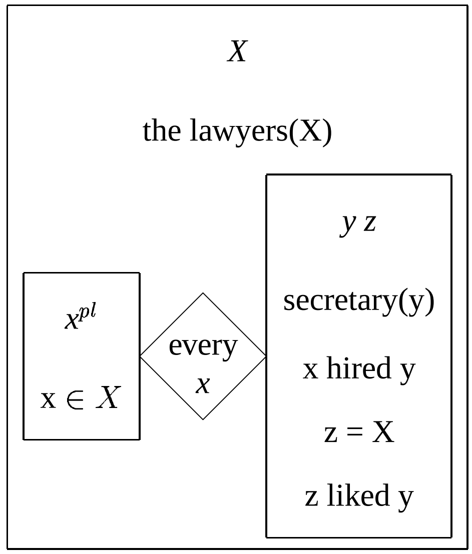
b.



Reading (3.18c) would express a (very questionable) distributed but independent reading:

a situation under which each lawyer hired a secretary that all the lawyers liked. Kamp and Reyle (pp. 328, 353) voice reservations about whether this reading actually exists, noting that many speakers find it unacceptable. But ultimately they decide to allow their system to generate it, and represent it as follows:

(3.20)

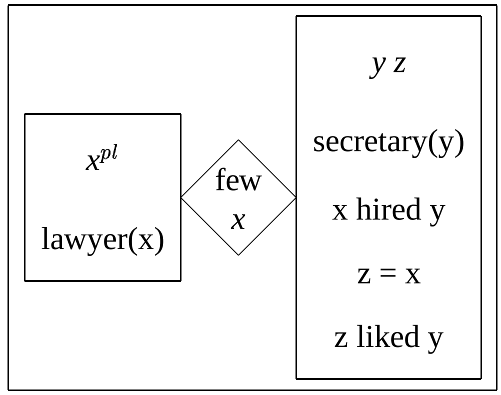


Note that both the distributed DRSs given involve the hiring of multiple secretaries; this is a simple consequence of the fact that discourse referents for indefinites are introduced into the current DRS, in this case, into the right hand side of the duplex condition introduced by the distributor.

The presumed existence of reading (3.18c) leads Kamp and Reyle to conclude that non-quantificational NPs introduce a discourse referent even when they are distributively interpreted, i.e., when they are otherwise processed like quantificational NPs. Explicitly quantified sentences clearly lack this type of reading. For example, sentence (3.21a) only has the distributed, dependent reading; neither the collective reading nor the equivalent of (3.18c), which would involve reference to a plural discourse referent, are possible. Accordingly, Kamp and Reyle conclude that quantified NPs do not introduce a plural referent for their subject. (But see the discussion of *abstraction* on p. 83).

(3.21) a. Few lawyers hired a secretary they liked.

b.



Example (3.21a) does allow distributed readings in which *they* refers to some plural third party, perhaps the set of all lawyers or (given the proper context) the set of all their malpractice insurance underwriters.

3.2.3 Plural-marked referents

In the dependent readings of sentences (3.18) and (3.21a), the pronoun *they* has been identified with atomic antecedents even though it is morphologically plural. This is clearly correct since English requires plural, not singular, pronouns in these sentences, but creates a mismatch between the syntax and the semantics. Kamp and Reyle go to some trouble to keep track of when a morphologically plural pronoun may be used. They point out that such a pronoun is appropriate just in case it is bound by a quantifier that ranges over a morphologically plural NP; thus sentences (3.18) and (3.21a), but not (3.22), have a reading in which *they* ranges over parts of the subject:

(3.18) The lawyers hired a secretary they liked.

(3.21a) Few lawyers hired a secretary they liked.

(3.22) Every lawyer hired a secretary they liked.

In order to make use of this generalization, Kamp and Reyle mark with the superscript *pl* any variable bound by a quantifier that ranges over a set introduced by a morphologically plural NP. Their anaphora rules then state that the antecedent of a plural pronoun must be either non-atomic or marked *pl*.

We saw that the pronoun in (3.21a) cannot refer to the set of only the lawyers that were happy with the secretary they hired; but this set can be referred to once a sentence is finished, as in the continuation shown in (3.23a). To represent this kind of anaphora, Kamp and Reyle allow for the creation of discourse referents comprised of all qualifying members of a quantified-over set via *abstraction*, shown in (3.23b). This is just Link's (1983) sum operation over a predicate, written as in (c) in the standard notation.[5] The variable X is associated with the sum of all individuals x with the stipulated properties. The referent thus defined can then be used as the antecedent of the pronoun *they* in the second sentence of (3.23a).

(3.23) a. Few lawyers hired a secretary they liked. *They* had discussed the applicants beforehand.

b. $\mathrm{X} = \Sigma x \boxed{\begin{array}{c} x\ y\ z \\ \\ \text{lawyer(x)} \\ \text{secretary(y)} \\ \text{x hired y} \\ \text{z = x} \\ \text{z liked y} \end{array}}$

c. X = $\sigma x\ (\exists y$ lawyer(x) & secretary(y) & x hired y & x liked $y)$.

[5] Link defines $\sigma x\ Px$ as the supremum of all individuals in the extension of *P. Because *P is by definition closed for sums (and because the lattice of individuals has certain convenient properties), this is the same as the sum of all individuals in the extension of P.

3.2.4 Dependent plural NPs

Kamp and Reyle use the term *dependent plurals* for bare-plural NPs that are semantically neutral between singular and plural (bound) interpretations. (Their use of the term should not be confused with what I have been calling *dependent pronouns:* plural pronouns bound by a distributor rather than by a quantificational NP). For example, sentence (3.24) is true in a situation where some students bought one book and some bought more than one, and the book or books that each student bought will keep that student busy for a month.

(3.24) The students bought books that will keep them busy for a month.

Plural pronouns can be bound by c-commanding distributors or quantifiers, as long as the antecedent NP is plural. (*Few lawyers,* but not *every boy*). But unlike pronouns, dependent NPs are subject to a locality constraint: they are only licensed if there is an already processed (usually: c-commanding) quantified or distributively interpreted plural NP *in the same clause.* Sentence (3.25a) only has an odd reading, involving multiple automatic transmissions in the same car. But sentence (3.25b) has a dependent plural reading, under which each woman bought one or more cars that had one or more automatic transmissions. Plural pronouns are not subject to such constraints, as demonstrated by the fact that sentence (c) allows the dependent reading.

(3.25) a. The women bought a car which had automatic transmissions.
b. The women bought cars which had automatic transmissions.
c. The women bought a car which they liked.

Kamp and Reyle conclude that introduction of a dependent plural NP discourse referent must be licensed by the existence, in the same DRS, of a morphologically plural discourse referent. In order to keep track of which NPs can license dependent plurals, atomic variable introduced by quantifiers and distributors over morphologically plural NPs are marked with the superscript *pl*. Such referents are only introduced by quantifiers and distributors over plural NPs, so this criterion gives us the proper licensing condition. Dependent plural referents are represented as variables that are neither identified as atomic nor as non-atomic; they are written using lower-case Greek letters. Even though they may range over singularities, such "neutral" referents are marked with the feature *pl*, and can therefore license other dependent plurals. For example, to treat *automatic transmissions* as a dependent plural in examples (3.25a,b), the subject of the relative clause must be represented by a variable marked *pl;* this will only happen if the relative clause is headed by the plural *cars* (which is itself interpreted as a dependent plural, and thereby marked *pl*, after being licensed by the NP *the women*).

(3.26) The women bought cars^{pl} which e^{pl} had automatic transmissions^{pl}.

Dependent pronouns, Kamp and Reyle point out, sometimes also need to be licensed in a similar way. This occurs when a dependent pronoun is not directly bound by a c-commanding antecedent; their analysis involves two elements, a licensing mechanism for the plural and a paycheck translation for the pronoun.

Kamp and Reyle discuss examples (3.27a) and (b) under a situation in which each child found a teacher who opened his or her present.

(3.27) Every director gave a present to a child from the orphanage.

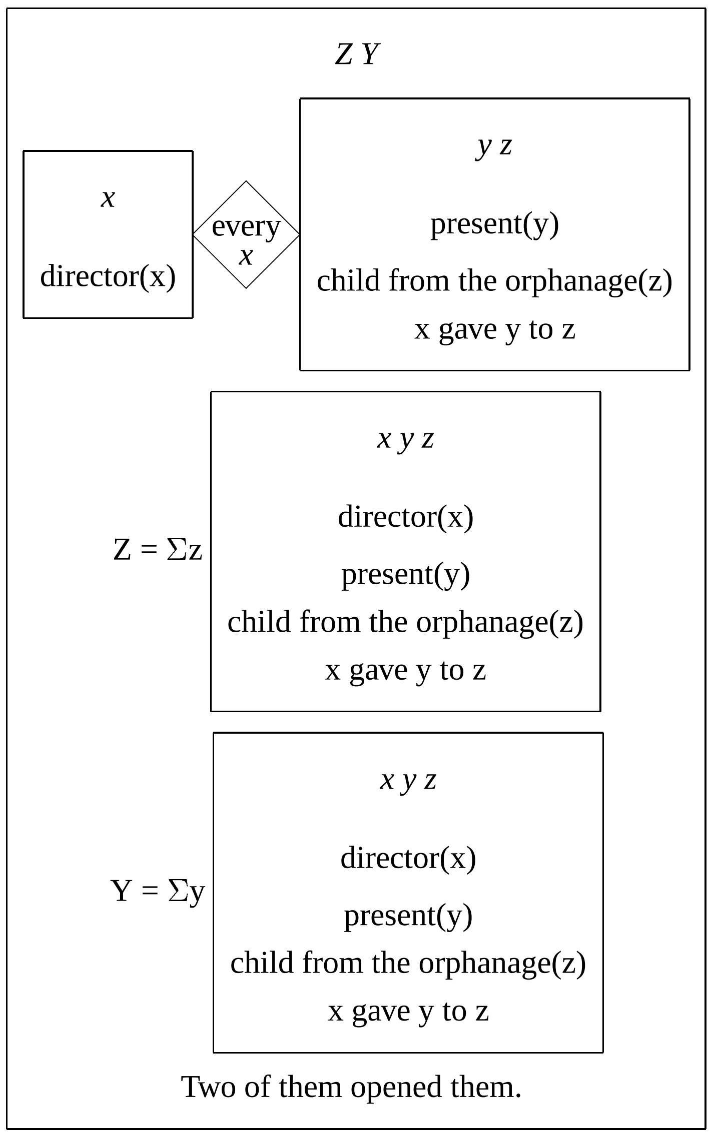
a. Two of them found a teacher who opened them.

b. Two of them found teachers who opened them.

Continuation (3.27a) does not have a reading in which each of the two children found a different teacher who opened his or her present (or presents). This reading is possible with continuation (3.27b), showing that the plural head *teachers* is needed to license the dependent reading of the pronoun here. Consider first the translation of the pronouns in the simpler sequence (3.28). The subject pronoun *them* refers to the set of children from the orphanage, which can be constructed by abstraction over the DRS of the previous sentence; similarly, we can construct a referent for the set of children, resulting in the following partially interpreted DRS.

(3.28) Every director gave a present to a child from the orphanage. Two of them opened them.

(3.29)



As the second sentence is interpreted, the lower pronoun needs to be construed as dependent on the higher one. Kamp and Reyle achieve this by making use of the following principle:

(3.30) When a set is introduced via Abstraction over some complex condition δ, then the information contained in the constituent DRSs of δ is available as information concerning the members of that set. This means that when we distribute over such a set, the DRS occurring on the right-hand side of the Abstraction equation may be "copied" into the left-hand DRS of the duplex condition which the distribution operator introduces. (p. 379)

Using this principle, the DRS shown in (3.29) is extended by the following duplex condition (the entire DRS is not shown in the interests of space).

(3.31)

u^{pl} x y

$u \in U$

director(x)

present(y)

child from the orphanage(u)

x gave y to u

every u

u opened them

The stage is now set to assign to the pronoun *them* the antecedent y, which for each child u is the present that some director gave to u. In effect, the second sentence of (3.28) is interpreted as saying "Two of the children who were given a present by a director opened the present they were given."

This would be the end of the story for Kamp and Reyle if it were not for the technical issue of licensing the plural on the object pronoun. According to their system, a dependent plural can only be licensed by an antecedent that is marked *pl*. In order to license the plural in a properly constrained way—recall that example (3.27a) lacks the dependent reading—they define an additional type of plural diacritic, *pl(x)* (where *x* is the variable that licensed the superscript). Since I will not attempt to replicate this part of their analysis, the details need not concern us here.

3.3 Are dependent pronouns real?

In section 3.1 we discussed sentences involving dependent pronouns that are not c-commanded by their intended antecedents. The obvious analysis of dependent pronouns, as quantifier-variable constructions, necessarily involves c-command and therefore cannot account for such examples. In looking for an alternative analysis, the simplest option would be to deny that a binding-like relationship is involved at all. If the dependent pronoun is actually interpreted not as a bound variable but as a plural, referential pronoun, then there is no c-command requirement and no problem to explain. This section is devoted to considering, and rejecting, this line of analysis.

3.3.1 The cumulative alternative

We accept sentence (3.32a) as having a sense in which it is true in a situation where each man kissed only one baby, his own.[6] One way to derive this reading is to give sentence (3.32a) a *cumulative* interpretation (cf. Scha 1984), which requires that every man kissed at least one baby and that every baby was kissed by at least one man, but nothing more.[7]

[6]The reading that involves this situation is described as *codistribution* by Sauerland (1995b). I will reserve his term for effects of this sort between two definite NPs; as I show in this section, such effects should not be confused with the relationship between an NP containing a dependent pronoun and its antecedent.

[7]The truth conditions associated with the codistributed representation always entail those of the weaker cumulative construal; in other words, every situation (model) that satisfies the codistributive truth conditions also satisfies the cumulative truth conditions.

Clearly these conditions are satisfied if every man kissed his own baby (or babies); if a speaker chooses to infer from (3.32a) this stronger meaning, this is between her and her pragmatics. Why not, then, apply the same analysis to sentence (3.32b), and even to (c) and (d)?

(3.32) a. The men kissed the babies.

b. The men kissed their babies.

c. The men who carried the babies kissed them.

d. The men urged their babies to play with each other.

This would constitute what I will refer to as the *cumulative analysis* of dependent pronouns. Despite its initial plausibility, I will show that it is wrong: Although the cumulative analysis is appropriate for some constructions, it is generally acknowledged that others, particularly those involving pronouns, involve a pairing of the members of one NP with those of another that is more structured than the cumulative analysis can account for; this must be considered a distinct reading. (This view is held by Heim et al. (1991a) and Schwarzschild (1996), among others).

We begin by considering the readings of sentence (3.1), repeated here:

(3.1) John and Mary think they are sick.

Sentences of this type were first discussed by Higginbotham (1981, 1985). His analysis is in the spirit of what I have called the cumulative analysis: The pronoun is coindexed with its antecedent, but no connection is asserted between elements of a decomposition of the embedded subject (into *John* and *Mary*) and a decomposition of the matrix subject. Instead, the following rule licenses the partial readings:

(3.33) Suppose that pronominal X is linked to antecedent Y in phrase marker Σ. [...]

Include some values of Y among the values of X.

(Higginbotham 1985:p. 573)

Heim et al. (1991a), who adopt a bound-variable analysis for the dependent reading, show convincingly that Higginbotham's rule is too vague. It would allow (3.1) to describe any one of the following states of affairs:

(3.1) John and Mary think they are sick.

(3.34) a. John thinks John and Mary are sick, and Mary thinks the same.

b. John thinks John is sick, and Mary thinks Mary is sick.

c. * John thinks Mary is sick, and Mary thinks John is sick.

d. * John thinks John is sick, and Mary thinks John and Mary are sick.

e. * John thinks Mary is sick, and Mary thinks John and Mary are sick.

f. * John thinks Mary is sick, and Mary thinks Mary is sick.

g. (Etc.)

Heim et al. (1991a) note that only the first two of these readings are possible: The "fixed" reading (a), and the bound-like (dependent) reading (b). The "crossed" reading (c) is impossible, as are the mixed readings (d) and (e).[8] Reading (f) is permitted by Higginbotham's rule, but it is ruled out by any analysis that explicitly appeals to cumulativity as defined above: the sum of all referents for the interpretation of *they* in (f) is just *Mary,* which is not equal to the entire intended antecedent, *John and Mary.* But the unavailability of readings (c)–(e) cannot be predicted by the cumulative analysis.

It must be acknowledged at this point that it is not completely impossible to accept (3.1) as a description of one of the states of affairs (c) through (e), given some goodwill and a bit of practice with such examples: After all, we have no reason to *rule out* a cumulative

[8] I am using my own terminology to describe these readings, not that of Heim et al.

reading for (3.1), which would say that John and Mary, between them, hold beliefs about the group of people consisting of John and Mary. However, it should be plain that the status of such readings is very different from the status of (a) and (b). At any rate the difference in acceptability between the dependent reading (b) and the crossed reading (c) cannot be predicted by any true cumulative analysis.

3.3.2 Context restrictors and strongest meaning

A defender of the cumulative analysis would need to explain why, out of all the logically possible readings of the above examples, only the fixed and the dependent readings are actually available. To this end, one might appeal to the relative pragmatic plausibility of various cumulative pairings of the parts of the relevant NPs. For example, Sauerland (1995a) discusses sentence (3.35) in a context such as a folk dance, where the men and the women are arranged in pairs. In such a situation, it is natural to take the sentence as saying that each woman faces her partner.

(3.35) The women face the men.

Sauerland proposes a very powerful mechanism (for which he credits Irene Heim) that can generate the desired truth conditions as an elaboration of a cumulative analysis. This is accomplished by modifying the truth conditions of the relevant predicate by means of *contextual restrictors*, context-supplied propositions that can be adjoined into suitable positions at LF. (As with codistribution, any sentence under the dependent-pronoun reading satisfies the truth conditions for the cumulative reading; thus it is only necessary to strengthen the truth conditions in order to derive the truth conditions of the dependent reading from those of the cumulative reading).

For example, sentence (3.35) is translated as (3.36). Note that these context restrictors result in truth conditions consistent with a bound-variable translation, even though the sentence is in fact given a cumulative translation in the style of Sternefeld (1998).

(3.36) [the women] [the men] * (κ_{12} & $\lambda 2\, \lambda 1\, [t_1 \text{face} t_2]$),

where $\kappa_{12}(x, y) = 1$ iff x and y are a couple.

Similarly, (3.37a) could be translated as shown in (b).

(3.37) a. The voters thought their candidates would win.

b. [the voters] [their candidates] * (κ_{12} & $\lambda 2\, \lambda 1\, [t_1$ thought t_2 would win$]$),

where $\kappa_{12}(x, y) = 1$ iff x voted for y (or, y is x's candidate).

We also need some way to select the proper set of contextual restrictors out of all logically possible choices. This can be accomplished by means of some pragmatic principle that, other things being equal, chooses meaningful statements over (relative) platitudes. Sauerland appeals to the *strongest meaning hypothesis* of Dalrymple et al. (1994), which asserts that of the acceptable possible interpretations for a sentence, we tend to choose the logically strongest one that is not contradicted by information already available to us. It is then possible to account for the preferred reading of (3.35) by claiming that a reading in which each woman faced a man to whom she bore a specific relationship (dance partner) is stronger than the reading in which each woman faced some random man.

Before going on I should make clear that Sauerland has *not* advocated applying the codistributivity mechanism to pronouns, or to NPs containing a pronoun (such as example (3.37)). To my knowledge, nothing I argue in this section contradicts a position that he actually holds or held. But since pronouns can be interpreted as plural referential expressions, and give rise to cumulative readings, the codistributivity mechanism is at least in principle available to them, and one could plausibly (at least at first blush) appeal to it as a way of evading the problems raised by examples like (3.1).

(3.1) John and Mary think they will win.

But recall that sentence (3.1) does not have a *crossed* reading, which would say that John thinks Mary will win, and vice versa. Sauerland's context restrictors, by themselves, cannot predict the lack of the crossed reading: it is just as easy to insert a context restrictor that makes the pronoun identical with the matrix subject as it is to insert one that forces the pronoun to be distinct from the subject. Neither can the Strongest Meaning Hypothesis, since matching John with Mary and Mary with John is logically neither stronger nor weaker than matching John with himself and Mary with herself. For purposes of discussion, let us assume that there is in addition a way of ranking restrictors according to some sort of *a priori* scale of preference. We might might try to account for the dependent reading of (3.1) by claiming that a meaning in which John's belief is about himself is preferred to one in which it is about Mary.

In this case the condition κ would be simply $x = y$:

(3.38) [$(J \oplus M)_i$] [they$_i$] * (κ_{12} & $\lambda 2\, \lambda 1\, [t_1$ thinks t_2 will win$])$,
where $\kappa_{12}(x, y) = 1$ iff $x = y$.

Let us evaluate this approach in an example with richer context. Given the background that Street and Weinberg ran against each other in an election that can only have one winner, consider the possible readings of the following sentences:

(3.39) a. The people who voted for Street and Weinberg thought that their candidates would win the election.
b. The people who voted for Street and Weinberg thought that they would win the election.

Example (3.39a) is strongly biased toward the plausible, dependent reading, according to which each voter expected the candidate they voted for to win the election. It might also allow the (unrealistic) fixed reading, in which each voter expected *both* candidates to win.[9] However, once again the crossed reading (according to which each candidate's supporters thought that *the other* candidate would win) is impossible. Like example (3.39a), sentence (3.39b) favors the dependent reading, which says that every voter expected the candidate they voted for to win the election; It also allows the (unrealistic) fixed reading. We could try to explain the prominence of the dependent reading by claiming that it is natural for voters to expect the candidate they voted for to win, and so we prefer this "stronger" reading to the fixed or the crossed reading.

But if it is plausible for voters to expect the candidate they voted for to win, example (3.40a) should have a "natural" reading which says that those voters who voted for Street expected *Weinberg* to lose, and vice versa (that is, the crossed reading should be the natural one). But such a reading is impossible. Perhaps voters can simply be assumed to have expectations, optimistic *or* pessimistic, about whoever they voted for? Then we have no explanation for the fact that (3.40b) also lacks the crossed reading.

(3.40) a. The voters who voted for Street and Weinberg thought that they would lose.

b. The voters who voted against Street and Weinberg thought that they would win.

Although this game could be continued for a while, it should be clear by now that the dependent reading of these sentences relies on interpreting the pronoun as having not just any salient relation to its antecedent, but one that is mentioned in (or is otherwise derived from) the preceding discourse. This suggests that the Strongest Meaning Hypothesis as I have represented it here is not applicable, but it still leaves open the possibility that Sauerland's analysis, coupled with a more appropriate way of choosing contextual restrictors,

[9] This reading is not unrealistic in the state of Florida.

could allow us to treat non-c-commanded dependent pronouns as cumulatively interpreted plurals.

3.3.3 The importance of being a pronoun

The cumulative analysis of dependent pronouns is based on treating pronouns on a par with full NPs. In this section, I show that their behavior is in fact markedly different with respect to the availability of the dependent reading.

Sentence (3.41a) is ambiguous between the (pure) cumulative and the codistributed reading, according to which each father coached his own son(s). The codistributed reading appears to be truly pragmatically induced, and is subject to partial cancellation, as shown by the fact that we can continue (3.41a) with (3.41b). Variant (3.42) is very different: the dependent reading is very prominent and does not allow cancellation, so that we cannot easily continue (3.42a) with (3.42b).

(3.41) a. The fathers coached the sons in Little League baseball.

b. Most fathers coached their own sons, but Bill and Jake coached each other's sons.

(or: For reasons of fairness, nobody coached his own son.)

(3.42) a. The fathers coached their sons in Little League baseball.

b. ??Most fathers coached their own sons, but . . .

It could be argued (and I have no reason to disagree) that sentence (3.41a) is ambiguous between a pure cumulative and a codistributed reading, and that continuation (b) cancels the codistributed reading, limiting us to the cumulative. If so, a defender of the cumulative analysis of dependent pronouns should conclude that (3.42a) only has the codistributed reading, lacking the plain cumulative reading; thus if the codistributed reading is canceled by continuation (b), no acceptable reading would remain. But why would the presence of a pronoun (which, by assumption, is *not* being treated as bound) suppress the cumulative reading? Under this account there is no explanation for the contrast, which is trivially explained if we accept that the pronoun is bound by a distributor.

We find the same behavior if we let the embedded NP consist of just a pronoun. Consider the following sentences, once again given the background that Street and Weinberg ran against each other in an election that can only have one winner.[10]

(3.39b) The people who voted for Street and Weinberg thought that they would win the election.

(3.43) The people who voted for Street and Weinberg thought that Street and Weinberg would win the election.

We have seen that sentence (3.39b) favors the dependent reading, which says that every voter expected the candidate they voted for to win the election. The unrealistic reading, which ascribes to every voter the belief that *both* candidates would win, is decidedly less prominent. But sentence (3.43b) only allows the fixed reading, contrary to what a cumulative analysis would predict: Since the pronoun in (3.39b) is assumed to take *Street and Weinberg* as its antecedent, the two sentences should have identical readings since the codistribution mechanism should apply identically to both.

Higginbotham's rule fares somewhat better here, since it is specifically restricted to pronouns; but again, it fails to predict the non-existence of a crossed reading for (3.39b), under which some or all voters expected the candidate they did *not* vote for to win the election.

Sentence (3.39b) is of particular interest because the dependent pronoun is not c-com-

[10] I am grateful to Anthony Kroch for bringing contrasts of this sort to my attention.

manded by its antecedent (which is trapped in a relative clause, a scope island).[11] Since this pronoun cannot be straightforwardly interpreted as a bound variable, a cumulative analysis would be particularly welcome—had it been tenable.

Consider also what a true cumulative reading would mean here: It would merely say that each of Street and Weinberg's supporters expected one of the two of them to win the election, but nothing more specific; there might be some optimistic and some pessimistic supporters in both camps, as long as *someone* expected each one of them to win. Supposing that Street and Weinberg had been the only candidates in that election, (3.39a) should be paraphrasable as

(3.44) The people who voted for Street and Weinberg thought that someone/someone's candidate would win the election.

It should be clear that sentences (3.39a) and (b) say a lot more than that.

3.3.4 The fixed reading

So far we have considered only interpretations of (3.1) in which the pronoun refers to one or more of John and Mary. In addition to the readings given in (3.34), let us now consider the following possibilities:

(3.1) John and Mary think they are sick.

(3.45) a. John and Mary think that [the Spice Girls are sick].

b. John and Mary think that [Mary and Margaret are sick].

c. John and Mary think that [John, Mary, Bill and Margaret are sick].

d. * John thinks Bill is sick, and Mary thinks Margaret is sick.

[11] In order to draw the contrast between (3.39b) and (3.43) it was in fact necessary to switch to such examples, since sentences like *John and Mary think that John and Mary are sick* incur a Principle C violation.

e. * John thinks John is sick, and Mary thinks the Spice Girls are sick.

Interpretations (a–c) are easily available, provided only that the prior context has established the desired referent for *they* as a possible pronominal antecedent. For example, the following context firmly establishes interpretation (a):

(3.46) The Spice Girls haven't toured recently.
John and Mary think *they* are sick.

The well-formed interpretations have in common the property that John and Mary believe the same proposition. As Heim et al. (1991a) point out, such readings can always be represented by employing a pronoun that is coreferential with, rather than bound by, its antecedent. (If the antecedent of the pronoun does not c-command it, the pronoun *must* be treated as referential rather than bound). Interpretations (d) and (e), on the other hand, require John and Mary to believe different propositions, and are impossible or at least much harder to get: The only well-formed reading in which John and Mary believe different propositions is the dependent reading (3.34b) discussed earlier, in which John's belief is about himself and Mary's belief is about *her*self. In other words, the only good non-fixed reading is the one expressed by the bound-variable interpretation for the pronoun; and conversely, a referential (not bound) pronoun must necessarily be understood according to a fixed reading. Note also that this effect is independent of whether John, Mary or both are properly included in the antecedent of *they*. Hence I will refer to any reading where all elements of the subject believe the same proposition as a *fixed* (or independent) reading, regardless of whether or not the subject of the embedded clause matches the subject of the matrix clause.

Let us now consider what the range of available readings tells us about their likely interpretation: A dependent reading is possible only when the antecedent of the pronoun appears in the same sentence (even if it does not c-command the pronoun). All cases in which the referent of the pronoun does not appear in the sentence allow only fixed readings.[12] Such a restriction makes no sense under the cumulative analysis: cumulative readings should be possible between a referential pronoun and a higher NP, just like they are possible between two referential NPs. Their absence constitutes strong evidence that dependent pronouns are indeed bound expressions of some sort, not cumulatively interpreted plurals. This conclusion accounts for the range of possible interpretations, although we still need an explanation of how such binding is possible without c-command.

3.3.5 Final comments on the cumulative alternative

So far I have considered the following phenomena that cannot be explained by the cumulative analysis: (a) the lack of the crossed reading, (b) the limited availability of "salient" relations that might restrict a cumulative interpretation by means of Sauerland's restrictors, (c) the importance of having a pronoun, not just any NP, in the dependent position, and (d) the fact that the dependent reading is only possible under conditions much more restrictive than those for the fixed reading, essentially being possible only when the antecedent of the pronoun appears somewhere in the same sentence.

Let us consider, briefly, one more argument. There are constructions whose semantics is simply incompatible with the assignment, in the cumulative style, of plural reference to the pronoun. First, consider the following variation of example (3.39a):

[12] An anonymous SALT 9 reviewer brought up the following type of exception to this generalization:

(i) Q: What did the men tell you about their children?
A: They said that they are doing well in school.

Here each man said that his child (or children) is doing well in school. Here the embedded pronoun refers to each man's own child; no antecedent referring to children is present in the same sentence, but an expression (*their children*) mapping men to their children is available in the immediately preceding discourse. In section 3.5 I show that we should analyze the embedded pronoun as a function taking each man to his children. These data do not contradict the argument being made here, since the pronoun is interpreted as a function operating on parts of an antecedent (*the men*) that does appear in the same sentence.

(3.47) The people who voted for Street and Weinberg thought that *their candidate* would win the election.

Because the head of the dependent NP is a singular noun, this example cannot even be interpreted unless we assign a bound interpretation to the pronoun *their.* If this pronoun referred to the set of all voters, then the NP *their candidate* should refer to some unique candidate that all voters voted for. But by assumption there is no such person; instead there are two candidates, each with his supporters. Hence the pronoun must be a bound variable of some sort, ranging over parts of the set of voters.

Schwarzschild (1996:p. 114) provides some additional examples that are incompatible with a cumulative translation of the pronoun. The pronoun in (3.48a) must be interpreted as a bound variable if this sentence is used in a situation where grades were given one at a time, and each student left the room immediately after receiving his or her grade. Another example is given in (b): The pronoun *their* must be interpreted as a singular bound variable, if the soldiers were from different home towns and received money from different friends.

(3.48) a. The students left the room immediately after receiving *their* grades.

b. Those soldiers received money from a friend in their home town.

It should be clear that the dependent reading of the sentences we have considered relies on interpreting the embedded pronoun as if it is a variable bound by a higher quantifier ranging over the members of its antecedent NP; and that this mechanism is specific to pronouns, since a full NP in place of the pronoun (as in example (3.43)) cannot receive such an interpretation.

I used the hedge “as if it is” in the previous paragraph because the structural configuration of (3.39b) prohibits binding of the pronoun by its intended antecedent. In the next section, I argue against an analysis that would let a definite NP bind outside the relative clause in the style of the "functional relative clauses" of Sharvit (1999). With this possibility eliminated, in section 3.5 I conclude that dependent pronouns should be treated as "paycheck pronouns" containing a function-denoting variable, in the style of Engdahl's (1986) adaptation of Cooper (1979).

3.4 Functional relative clauses?

In sentence (3.49) there is no antecedent that could bind the pronoun *them* as a bound variable; the intended antecedent is buried in the relative clause. Since the dependent reading is nevertheless available and we have ruled out the cumulative option, our conclusion must be that either the intended antecedent of the pronoun is somehow able to bind outside the relative clause, or the pronoun is not directly bound by the NP *Street and Weinberg*, but by something else. This section considers (and rejects) the first possibility, binding of the pronoun from inside the relative clause.

(3.49) The voters who support Street and Weinberg hope they will win.

The form of the relative constructions we have discussed resembles the "functional relative clauses" studied by Sharvit (1999). These are relative clauses containing a quantifier that appears to bind a pronoun outside the relative clause, as in the Hebrew example (3.50a). English does not freely allow this type of binding in non-identity sentences, but does allow it when the head of the relative clause contains an anaphor, as in (3.50b). (Sharvit takes the paucity of such readings to be an idiosyncratic property of English).

(3.50) a. ha-iSa Se kol gever xibek cavta oto.
the-woman that every man hugged pinched him
'For every man x, the woman that x hugged pinched x.'

b. The picture of himself which every student hated annoyed his friends.

In Sharvit's analysis of such sentences, the referential index of a quantificational NP can in effect escape the relative clause through absorption into the relative clause operator; the headed relative clause is translated into a functional expression which is applied (via a generalized quantifier) to a two-place complement to generate one of its arguments from the other. The resulting translation of sentence (3.50a) is shown in condensed form in (3.51a). Here A is the unique *witness set* of the generalized quantifier *every man*, and g is the unique function with domain A that maps elements of A to women they hugged; its full definition is given in (3.51b).

(3.51) a. $\forall x \in A\, \text{pinch}'(g(x), x)$

b. $\iota g\, [\text{Dom}(g) = A\ \&\ \forall y \in A\, [\text{woman}'(g(y))\ \&\ \text{hug}'(y, g(y))]]$

I give only this brief description of Sharvit's analysis because the functional relative clauses she discusses have grammaticality conditions very different from those of the dependent constructions with relative clauses. The first difference is that English, as already mentioned, does not easily allow functional relative clauses in non-identity sentences; while all our examples of dependent readings without c-command involved non-identity sentences. A second difference is that functional relative clauses may have singular heads (as with example (3.50b)), while as we saw in section 3.1, the dependent reading is only possible if a relative-clause subject has a plural head noun; thus example (3.52a) does not allow the dependent reading, but the plural-headed (b) does.

(3.52) a. The coach that trained them thinks they will win.

b. The coaches that trained them think they will win.

Finally, the functional reading of quantificational relative clauses like (3.50b) is sensitive to the syntactic position of the quantifier (possibly due to a weak crossover effect): the embedded quantifier must c-command the gap in the relative clause for the functional reading to be available, hence the following examples (also from Sharvit (1999)) are ungrammatical.

(3.53) a. * ha-iSa Se ϕ mexabeket kol gever covetet oto.
the-woman that is-hugging every man is-pinching him

b. * The picture of himself that ϕ depicted every student amused his friends.

Dependent pronouns, on the other hand, can take their antecedent from any position inside a relative clause; in example (3.49), repeated here, the gap is in subject position and is therefore not c-commanded by the intended antecedent of the pronoun, *Street and Weinberg.*

(3.49) The voters who support Street and Weinberg hope they will win.

Because of these differences between the quantificational relative clauses studied by Sharvit (1999) and the ones we have been concerned with, we cannot extend Sharvit's analysis to relative clauses with definite embedded NPs; even if we allowed definite NPs to undergo the same kind of operations that Sharvit identifies for quantifiers, the resulting theory would not be able to predict the distributional differences between the readings involving functional relative clauses with embedded quantifiers and those with embedded definites.[13]

[13] One possibility along such lines would be to claim that the object in the relative clause of (3.49) can raise to a position where it c-commands the gap, perhaps aided by the presence of an adjoined distributing quantifier. This would at least generate the configuration that, according to Sharvit, is necessary for the existence of a functional relative clause.

Such an analysis, which would treat a distributively interpreted object as quantificational, would be very close to that proposed by Pritchett (1990) in connection with functional questions. Taking the existence of

3.5 The functional analysis of dependent pronouns

In the previous sections I argued against the cumulative analysis of dependent pronouns, as well as the possibility that an antecedent can bind a dependent pronoun from inside a relative clause via some "functional relative clause" mechanism. Let us now turn to an analysis based on Cooper's (1979) treatment of "donkey" pronouns.

Cooper was able to account for many instances of apparent pronominal binding without c-command, via the following translation schema:

(3.54) Pronouns may be associated with any one of the translations of the form
$\lambda P \, \exists x \, \forall y[[^{\vee}\Pi](y) \leftrightarrow y = x] \;\&\; P(x)$,
where Π is a property-denoting expression containing only free variables and parentheses.

To handle the pronoun *it* in (3.55), we can let $\Pi = S(u)$, where S is a free variable of type <e, et> and u is a free variable over individuals (intended to be bound, via quantifying-in, by the quantifier ranging over *every man who owns a donkey).* The context may then supply a value for S such that $S(x)$ is the property of being x's donkey.

(3.55) Every man who owns a donkey beats *it.*

Similarly, in the "paycheck" sentence (3.56) $S(x)$ would be the property of being x's paycheck.

(3.56) The man who put his paycheck in the bank was wiser than the man who fed *it* to his dog.

pair-list readings to indicate that an NP has scope over the wh-word, Pritchett concluded that referential as well as quantificational NPs raise at LF. However, his argument was persuasively refuted by Krifka (1992) and, independently, by Dayal (Srivastav 1992, Dayal 1996), who noted important differences between quantificational and referential NPs in (purportedly) functional questions. Therefore such a raising analysis is untenable.

Engdahl (1986) recasts Cooper's formula into functional form, dispensing with the Russellian assertion of uniqueness included in Cooper's translation. Her translation is as follows:

(3.57) $\lambda P\, P(\Phi)$,

where Φ is a variable ranging over function expressions; e.g., Φ might be $W(x)$, the function giving x's donkey.

The function Φ may be of any arity (may accept any number of arguments), including 0 (in which case it just denotes an individual).

We can adopt the same analysis for dependent pronouns: In a sentence like (3.6a), repeated below, the pronoun *they* is not bound by *John and Mary* but denotes the expression $\lambda P\, P(S(u))$, where S is a function that maps every student taught by John or Mary to the person in the set {John, Mary} who taught him or her. (Note that the dependent reading of (3.6a) presupposes that John and Mary taught disjoint sets of students; if there is overlap, our intuitions about the meaning of (3.6a) get confused). Sentence (3.6a) then translates as (3.58), which says roughly that each of the students taught by John and Mary thinks that the person that taught them will win $100.

(3.6a) The students John and Mary taught think they will win $100.

(3.58) $\forall x \in \{y : y$ a student that John or Mary taught$\}$ think$(x,$ ^[win-100(S(x))$])

($S(x)$ = x's teacher)

Ordinary cases of dependent pronouns, in which the intended antecedent does c-command the pronoun, can be translated as the identity function or simply as bound variables.

There is a special case of dependent pronouns which we should mention here.[14] These are examples where the dependent pronoun is interpreted as a function supplied by the

[14]The complications associated with this example were brought up earlier, in fn. 12 of this chapter.

preceding context, not by the current sentence. In (3.59a), the embedded pronoun *they* must involve the function mapping each man to his children. In the translation shown in (b), Φ is the function mapping men to their children (an open variable that will be supplied by the context), and u is a variable that is eventually bound by the distributor over the parts of the matrix subject, *they*.

(3.59) a. Q: What did the men tell you about their children?

A: They said that *they* are doing well in school.

b. they = $\Phi(u)$ = u's child/children

As a result the dependent embedded pronoun ranges over individuals (children) that are not mentioned elsewhere in the *same* sentence. Our analysis explains how this reading is derived: the bound variable u in the translation of the dependent pronoun ranges over men, not children, and is bound in the ordinary way by the distributor over the matrix subject. Children are only involved as the output of the dependent pronoun's reference function.

A note may also be in order on the source of the reference function that the context supplies. As we saw in section 3.3.2, pragmatics alone is not enough to license any value for the reference function. The reader may be reminded of the following well-known contrast, discussed by Heim (1982:p. 21) and attributed to Barbara Partee:

(3.60) I dropped ten marbles and found all of them, except one. It is probably under the sofa.

(3.61) I dropped ten marbles and found only nine of them. ??It is probably under the sofa.

The example shows that mere salience is not sufficient to license pronominal anaphora: explicit mention is generally required. Likewise, in the examples repeated below the pronoun is understood to refer to whichever politician a group of voters voted for, or didn't vote for, etc., according to whatever relationship has actually been mentioned.

(3.40) a. The voters who voted for Street and Weinberg thought that they would lose.

b. The voters who voted against Street and Weinberg thought that they would win.

3.6 Split dependent plurals

The analysis of dependent pronouns in terms of functions is supported by the fact that dependent pronouns can have split antecedents, giving what I will call a *split dependent* reading. Ordinary plural pronouns like the one in (3.62) can take split antecedents; the same type of split reading can co-occur with the dependent plural interpretation, so that (3.63a) has reading (3.63b):

(3.62) Tom told Mary that they$_{t,m}$ should leave.

(3.63) a. Manny, Moe and Jack (each) told me that *we* are neighbors.

b. Manny told me that Manny and I are neighbors, *and* . . .

Moe told me that Moe and I are neighbors, *and* . . .

Jack told me that Jack and I are neighbors.

Sentence (3.63a) has any number of other readings: it could mean that Manny, Moe and Jack told me that all four of us are neighbors, or that Manny and I only are neighbors, or that I and some third person (or persons) are neighbors, etc. But reading (3.63b) is the only one in which Manny, Moe and Jack communicated to me *different propositions*. (This is the same pattern discussed in section 3.3.4).

Being neighbors requires a plural antecedent, but being rich does not. Accordingly, sentence (3.64) allows two different dependent readings, the *split dependent* reading (a) and the *singular dependent* reading (b):

(3.64) John and Mary told Harry that *they* are rich.

a. John told Harry that John and Harry are rich, *and* . . .

Mary told Harry that Mary and Harry are rich.

b. John told Harry that John is rich, *and* . . .

Mary told Harry that Mary is rich.

Once again, there is also a multitude of fixed readings, which I group together: Perhaps John and Mary told Harry that *John and Mary* are rich, or that all three of them are rich, or that the Rockefellers are rich, etc. Whether they involve third parties or just the participants of this sentence, all these other readings have the property that John and Mary said the same thing.

(3.65) a. John and Mary told Harry that the Rockefellers are rich.

b. John and Mary told Harry that Bill and Donald are rich.

c. John and Mary told Harry that John and Mary are rich.

d. John and Mary told Harry that John, Mary and Harry are rich.

(etc.)

Interpretations (3.64a,b) are the only possible dependent readings of sentence (3.64); as we have established, there is no "crossed" reading where John told Harry that Mary is rich, and Mary told Harry that *John* is rich (as in (3.66a)). There are also no readings mixed between split and singular dependence, as in (3.66b), or between fixed and dependent readings, as in (3.66c). In other words, the interpretation of *they* is determined only once per construal, proving that we are dealing with genuine ambiguity, not vagueness.

(3.66) a. * John told Harry that Mary is rich, *and* . . .

Mary told Harry that John is rich.

b. * John told Harry that John and Harry are rich, *and* . . . (split +

Mary told Harry that Mary is rich. singular)

c. * John told Harry that John is rich, *and* . . . (dependent +
Mary told Harry that the Rockefellers are rich. "fixed")

We now have the following typology of licit readings: "fixed" readings that could refer to anything, as long as all speakers state the same proposition; a "singular dependent" reading, in which the dependent pronoun is identified with each speaker separately; and a "split dependent" reading, in which the dependent pronoun refers to one speaker plus some other, fixed argument of the sentence.

Only one reading of the pronoun can be used in any given reading of a sentence. If we take (3.64) to say that John told Harry that John is rich, Mary cannot have said that Harry and Mary are rich, or that some other persons are rich; she must have said that *Mary* is rich. Similarly, if *they* is understood to refer to a fixed group as in (3.65), the group may consist of any set of individuals at all, but must be the same for *every* speaker. In other words, the interpretation of *they* is determined only once per construal, proving that we are dealing with genuine ambiguity, not vagueness.

Since ordinary split anaphora has been described in terms of assigning multiple indices to the pronoun (see Higginbotham 1981), we might consider treating split dependent pronouns in the same way, assigning them one referential and one bound index. However, it appears that the fixed readings enjoy much greater freedom for antecedent selection than does the fixed part of the split dependent reading. As we have seen, a fixed-reading pronoun can easily be understood as referring to individuals mentioned earlier, as in the construals given in (3.65) above. But for some reason, the fixed part of split dependent pronouns appears to be restricted to individuals in the current sentence, as in all the examples we have seen so far (3.63b, 3.64a). Even in the presence of suitable prior context, it seems difficult, if not impossible, to include a discourse-supplied entity (but see below):

(3.67) Jane is hard to get along with. John and Mary said that they disagreed over trivial things. =

?? John said that he and Jane disagreed, and Mary said that she and Jane disagreed.

In any case it seems safe to say that such readings, if possible, are not nearly as easy to obtain as non-dependent reference to a third party. I will not venture a general explanation of this effect, beyond noting that it supports the accepted division between bound and referential pronouns (as spelled out, for example, by Reinhart (1983)); and that it shows that dependent pronouns should not be translated as including a referentially interpreted index. Instead, I represent the dependent readings as functions from individuals to individuals. A split dependent pronoun is neither referential nor a bound variable, but its meaning can be expressed as a function that takes any individual x to the plural individual consisting of x plus some other, fixed individual. A singular dependent pronoun is represented simply as the identity function. (Note that the split dependent function is an extension of the identity function). Representing ordinary dependent pronouns as the identity function gives us the technical advantage of distinguishing them from pronouns bound by an ordinary quantifier.

While it seems to make no difference whether ordinary (c-commanded) dependent pronouns are represented as variables or as the identity function, the fixed readings should be represented (or at least representable) as referential or bound variables as appropriate, *not* as the constant function; otherwise they would be expected to obey the same restrictions that the fixed part of dependent pronouns obeys.

Finally, note that there is at least one way that a split dependent pronoun can pick out an individual from outside the sentence. An anonymous reviewer for SALT 9 suggested the following as a counterexample to the generalization that the antecedents of split-dependent pronouns must come from the same sentence:

(3.68) Q: What did the women tell you about themselves and their husbands?

A: They told me that they like each other.

The embedded pronoun *they* in the answer refers to each woman and her husband. In view of the account being developed here, this pronoun must be interpreted in terms of a function mapping each woman to the sum of herself and her husband (or possibly as the sum of each woman with the output of a function mapping women to husbands, as in (3.69b)):

(3.69) a. they = $\Phi(x) = x \oplus$ husband-of(x)

b. they = $x \oplus \Omega(x) = x \oplus$ husband-of(x)

In this case, then, both halves of the split dependent pronoun vary, and both are actually bound by the distributor ranging over women. The only part of its translation that is not drawn from the current sentence is the mapping function itself, which is drawn from the immediate context.

Because the non-identity part of the function is not referential but another function, our generalization about the fixed part of split dependent pronouns is not violated. We do need to add this type to our inventory of dependent pronoun functions, which we can now summarize as follows:

(3.70) **A typology of reference modes for dependent pronouns**

1. Fixed:

   they = $<$any *fixed* group$>$

2. Dependent:

   (a) Singular (identity map): they = $x \mapsto x$ (or bound variable)

   (b) Functional: they = $x \mapsto \Phi(x)$

3. Mixed:

   (a) Split-dependent: they = $x \mapsto (x \oplus \text{Harry})$

   (b) Split-dependent (functional): they = $x \mapsto (x \oplus \Phi(x))$

Perhaps, given a properly complicated context, the elementary fixed and dependent reference types can be combined in other ways; e.g., it might be possible to obtain the sum of two non-identity functions, or the sum of a non-identity function and a fixed referent. I will not delve into the question of whether such modes are actually instantiated.

Chapter 4

A non-scopal analysis of reciprocals

In the previous chapter, I showed that pronouns which appear to depend on a distributed antecedent that does not c-command them can be analyzed as functional (paycheck) pronouns, allowing them a translation in which there is no binding without c-command. But although the functional translation I proposed is sufficient for non-reciprocal sentences containing dependent pronouns, examples in which such constructions are the antecedents of reciprocals present problems of their own. In section 2.3.2 I showed that sentences such as (4.1a) cannot be handled by simply giving wide scope to the reciprocal (or to the distributor that binds it) as proposed by Heim et al. (1991a,b): there is no c-commanding antecedent that matches the intended antecedent of the reciprocal. The problem extends to parallel constructions such as (4.1b) and (c), in which the paycheck pronoun is expanded into a full NP.

(4.1) a. The lawyers who represent John and Mary think they should sue each other.

b. The lawyers who represent John and Mary think their clients should sue each other.

c. These lawyers think their clients should sue each other.

We now turn to these matters, beginning with a review of the problem as it applies to reciprocals.

4.1 A short review of the scopal analysis

Our canonical example of an ordinary dependent reciprocal is given in (4.2). This example may have the independent reading given in (a), according to which John and Mary think the same thing: "we like each other"; or it may have the dependent reading given in (b), according to which John and Mary think different things: John thinks "I like Mary," and Mary thinks "I like John."

(4.2) John and Mary think they like each other.

a. John and Mary think "We like each other."

b. John thinks "I like Mary," and Mary thinks "I like John."

The independent reading is straightforward: the pronoun *they* refers to *John and Mary,* and the entire reciprocal clause stays in the scope of the belief operator, as shown in (4.3a). For the dependent reading, the pronoun is translated as a variable bound by the matrix distributor and is therefore not a suitable antecedent for the reciprocal, which needs a plural antecedent to reciprocate over. In this discussion it is convenient to assume the analysis of Heim et al. 1991b, rather than the stronger but more problem-prone account in Heim et al. 1991a. Heim et al. (1991b) generate the dependent reading (4.2b) by letting the reciprocal be bound by a distributor adjoined to the matrix, not the embedded, subject, as shown by the indexing scheme in (4.3b).

(4.3) a. [[John and Mary]$_1$ D] think that [[they$_1$ D$_2$]$_2$ like [each$_2$ other]$_3$]

= John thinks "we like each other", and Mary thinks the same.

b. [John and Mary$_1$ D$_2$] think [that they$_2$ like [each$_2$ other]$_3$]

= John thinks "I like Mary", and Mary thinks "I like John".

The matrix distributor in (4.3b) binds the reciprocal's contrast argument, which must be the same as the local antecedent of the reciprocal, the bound variable *they*$_2$; it also determines the set of entities over which the object of the reciprocal predicate will range, the so-called *range argument* of the reciprocal. The system of Heim et al. (1991b), requires this to be the sister of the distributor that binds the reciprocal, i.e., the NP *John and Mary.*

Covert distributors are freely inserted at LF as necessary. The distributor introduces universal quantification over the parts of the set it adjoins to. Heim et al. represent it as an NP-adjoined operator whose translation can be written as follows:

(4.4) a. $\mathrm{D}_{<\mathrm{e},<\mathrm{et},\mathrm{t}>>} \equiv \lambda N \, \lambda\varphi \, \forall x_j (x_j \; \Pi \; N) \, \varphi(x_j)$

b. $[\mathrm{NP}_i \; \mathrm{D}_j]\varphi \Rightarrow \forall x_j (x_j \; \Pi \; \mathrm{NP}'_i) \, \varphi'(x_j)$

The definition of Π guarantees that the distributor cannot be applied to a singular NP.[1]

Heim et al. (1991b) translate the reciprocal as an operator that raises to adjoin to VP; it introduces a second universal quantifier, which binds a variable occupying the object position in the reciprocal predicate. Its translation can be written as follows:

(4.5) [each$_j$ other(i)]$_k$ $\Rightarrow \lambda P \, \lambda y \, \forall x_k (x_k \; \Pi \; X_i \; \& \; x_k \neq y) \, P(y, x_k)$

The independent and dependent readings are translated as in (4.6a) and (b), respectively. In each formula the lowest universal quantifier was contributed by the reciprocal.

(4.6) a. $\forall x (x \; \Pi \; J{\oplus}M_1) \, \mathrm{think}(x,$

$\hat{}[\forall x_2 (x_2 \; \Pi \; J{\oplus}M_1) \, \forall x_3 (x_3 \; \Pi \; X_1) \, x_2 \neq x_3 \Rightarrow \mathrm{like}(x_2, x_3)])$ (Independent)

[1] Heim et al. (1991b) define Π as the *proper-part-of* relation. Since atomic individuals do not have proper parts, they cannot serve as the antecedents of reciprocals.

b. $\forall x_2(x_2 \Pi\ J{\oplus}M_1)$ think$(x_2, {}^\wedge[\forall x_3(x_3 \Pi\ X_1)\, x_2 \neq x_3 \Rightarrow$ like$(x_2, x_3)])$

(Dependent)

4.2 Dependent anaphora without identity

Now consider a sentence like (4.7), under the dependent reading.

(4.7) The lawyers that represent them say they will sue each other.

A correct translation should state that each lawyer, x, says that x's client will sue the other clients (or, the other lawyers' clients).

As defined, the framework of Heim et al. (1991b) cannot generate this reading. To express the fact that reciprocals are subject to Binding Principle A, Heim et al. stipulate that the contrast argument of the reciprocal must be locally A-bound; in effect, that the long-distance binder must be coindexed with the local antecedent. This requirement leads to the prediction that the sentences in (4.1) should lack a dependent reading.

Let us assume that this condition can be suitably relaxed without overgeneralizing. Even with the paycheck analysis of the dependent pronoun, the framework of Heim et al. (1991b) cannot generate the correct reading. If we were to allow a long-distance reciprocal in this sentence, it might receive the following interpretation:

(4.8) $(\forall x_2 \cdot\Pi$ lawyers$_1)$ say$(x_2, {}^\wedge[\forall x_3(x_3 \cdot\Pi\ X_1\ \&\ W(x_2) \neq x_3)$ sue$(W(x_2), x_3)])$

(W is the function mapping lawyers to their clients).

The range argument of the reciprocal, the free variable X, is required to be coindexed with the set over which the long-distance binder (the matrix distributor) quantifies. Because in this case this is the set *lawyers*$_1$, (4.8) says, incorrectly, that each lawyer expects his or her client to sue the other *lawyers.* The correct translation can be generated if we stipulate that

X should be interpreted as the set of clients instead of being coindexed with *lawyers*$_1$. But given the scopal analysis adopted by Heim et al., there is no systematic way to require this: the scopal analysis asserts that the true semantic antecedent of a dependent reciprocal is the matrix subject, not the embedded dependent pronoun. To allow pragmatic adjustments to come into play here would imply that the range of the reciprocal could be potentially any set, depending on the context, when in fact it is rigidly determined: it can only be the set of entities that the reciprocal's *contrast argument*, $W(x_2)$, ranges over.

This problem is not limited to Heim et al.'s account; as we will see in chapter 5, its equivalent is also found in other scopal treatments of reciprocals, including those of Sternefeld (1998) and Schwarzschild (1996). Sternefeld, for example, captures the semantics of weak reciprocity via membership in the cumulation of the reciprocal predicate. (In contrast, the account of Heim et al. relies on direct quantification and encodes the semantics of strong reciprocity). In dependent reciprocal sentences the entire reciprocal raises to the matrix clause, so that the dependent reading of (4.7) would be as follows:

(4.9) $(\exists X)(X = \text{lawyers}_1 \ \& <X, X> \in {}^{**}\lambda x\, \lambda y\, [x \neq y \ \& \ \text{say}(x, {}^\wedge[\text{sue}(W(x), y)])])$

This is the same meaning as given by the Heim et al. (1991a) analysis: it claims that each lawyer x said that his or her client, $W(x)$, will sue one or more of the other *lawyers*, y. The correct semantics would require replacing the last occurrence of y with $W(y)$.

Thus the problem of finding the right range for the reciprocal is not specific to the analysis of Heim et al. (1991b), or to their particular assumptions about distributivity or type of reciprocal relation. The core of the problem is that the range of the dependent reciprocal depends on its local antecedent, but the scopal analysis cannot properly take its contribution into account. Indeed, it could be argued that the essence of the scopal analysis lies precisely in excluding the local antecedent: the core claim of the scopal analysis is that in a dependent ("long distance") reciprocal sentence, the reciprocal enters into a binding relationship with the matrix subject *instead* of with the local subject.

As it stands, then, the scopal approach to reciprocals cannot account for dependent readings in which the dependent pronoun corresponds to a function other than the identity: it predicts, wrongly, that the range and contrast arguments will match the long-distance binder, not the dependent pronoun. Having adopted a functional analysis for the dependent expressions, we could address the issue of reciprocals by internally applying the pronoun's reference function to the (non-local) range and contrast arguments of the reciprocal, and keeping the long-distance binding relationships as they are. This move would add another unbound variable to the translation of the reciprocal (recall that the range argument is also a free variable), this one based on the local binder. But once we have given ourselves access to the pronoun's reference function, a simpler alternative is possible: we can drop all reference to the long-distance binder, and let the range argument of the reciprocal be the range of the reference function. For concreteness, I base the discussion in this chapter on the analysis of Heim et al. (1991b); as we will see in chapter 5, similar adjustments can be made to more sophisticated scopal treatments.

4.3 The grammaticality of dependent reciprocals

The preceding critique of the scopal analysis stands or falls on the well-formedness of examples like (4.7), which involve a dependent reciprocal whose local (dependent) antecedent ranges over a different set of individuals than *its* antecedent. Given the complexity of the examples there is inevitably significant variation across speakers, and even for the same speaker on different occasions. Until this point I have simply presented examples and described them as acceptable. In this section, we review a range of dependent reciprocal constructions in more detail.

(4.7) The lawyers that represent them say they will sue each other.

To begin with, it should be taken for granted that dependent pronouns can have bound-like rather than cumulative readings, even in the absence of c-command between the pronoun and its intended antecedent. This position was defended in some detail in section 3.3. For example, I demonstrated that the possessive pronoun in sentence (4.10a) should be treated as a singular bound variable, not as a plural, cumulatively interpreted referential pronoun; and that the pronoun in (4.10b) should be translated as a paycheck pronoun containing a bound variable, and again interpreted as having singular (but varying) denotation. Second, it should be clear that so-called long-distance reciprocals do exist, at least in those cases where the dependent pronoun ranges over the parts of its antecedent, as in (4.11).

(4.10) a. The people who voted for Street and Weinberg thought that their candidates would win the election. (= (3.39a))

b. The people who voted for Street and Weinberg thought that they would win the election. (= (3.39b))

(4.11) a. The children told me they like each other.

b. John and Mary think they will defeat each other.

Example (4.11a) could be truthfully uttered in a context where each child has confided in me that he or she likes one or more of the other children. Example (b) might be used if John and Mary are about to face each other in a tennis match that each of them expects to win.[2] The remaining question, then, is whether dependent reciprocal readings are possible in those cases where the local antecedent of the reciprocal is not c-commanded by its intended

[2]The existence of dependent reciprocal readings is essential to the scopal analysis of reciprocals. If the embedded pronoun in (i) could be analyzed as a cumulatively interpreted plural, the reciprocal should simply take it as its antecedent instead of being bound by the distributor over the matrix subject.

(i) John and Mary think they like each other.

The sole motivation for positing non-local binding of the reciprocal in the scopal theories of Heim et al. (1991a,b) or others is the existence of dependent reciprocal readings, and (as discussed in section 2.3.4) the need to find the reciprocal's range argument in a principled way. Thus if we adopt the cumulative analysis, there is no longer any motivation for the scopal analysis of reciprocals.

antecedent. One example that has been widely cited in the literature (but does not *provably* involve lack of c-command) is the following:

(4.12) Their coaches think they are faster than each other. (= (2.53a))

The dependent, non-illogical reading of this sentence says that each coach thinks that the runner he or she trained is faster than the other runners. This dependent reading is somewhat harder to get than that of sentence (4.11b), but most speakers accept it and the consensus in the literature is that it is possible. But speakers that accept sentence (4.12) usually find example (4.13) equally acceptable under its dependent reading, even though the intended antecedent of the embedded pronoun occurs in a strong island.

(4.13) The coaches that trained them think they are faster than each other.

Similarly, if John and Mary have had a falling out and each has retained a lawyer, sentence (4.14) is acceptable under its dependent reading; it says that John's lawyer thinks that John will sue Mary, and that Mary's lawyer thinks that Mary will sue John.

(4.14) The lawyers who represent John and Mary think they will sue each other.
(= (2.63c))

Such examples are admittedly difficult to "get" for many speakers, and judgements vary over time with others. But there is overall a preponderance of affirmative judgements that establishes the well-formedness of this type of construction.

Dependent readings are generally compatible with paycheck pronouns even when the pronoun ranges over individuals that are not mentioned elsewhere in the *current* sentence. Suppose that John and Mary's mothers are neighbors. Recently John went to visit his mother and from the things she said, got the impression that she had been spying on Mary's mother. Meanwhile Mary went to visit *her* mother, and concluded that she had been spying on John's mother. In this context, example (4.15) is generally found acceptable under its dependent reading:

(4.15) John and Mary are worried about their mothers. They think *they* are spying on each other.

It is also possible for the antecedent of the reciprocal to be an NP containing a dependent pronoun. The following variations of the examples we have seen, considered in the same contexts, allow the dependent reading:

(4.16) a. The people who voted for Street and Weinberg expected their candidates to defeat each other.

b. John and Mary think their mothers are spying on each other.

c. These lawyers expect their clients to sue each other.

A number of factors affect the availability of dependent readings. Examples whose fixed reading involves illogical or contradictory beliefs are sometimes hard to get under the dependent reading, presumably because of interference between the two readings. (The unacceptability of the fixed reading makes such examples all the more valuable). On the other hand, a continuation that is only compatible with the bound-like interpretation usually helps bring out the dependent reading, as with example (4.17). (Capitalization indicates emphasis). The presence of overt complementizers might have a weak adverse effect on the dependent reading, while bridge verbs and Exceptional Case Marking constructions as in (4.16a) tend to favor it.

(4.17) John and Mary are certain that THEY like each other, but they wonder what the other one thinks.

These examples (and countless variations) are to various extents easy or difficult to get, but I found all of them to be acceptable to multiple speakers for whom the corresponding crossed readings were impossible, indicating that they did not confuse the dependent reading with the fixed, cumulative one. Overall I could not find grounds for contrasting the behavior of reciprocals with, e.g., paycheck pronoun antecedents as in (4.14), to the behavior of those with full NP antecedents, as in (4.16c). Accordingly I have treated all of the constructions considered as compatible with the dependent reciprocal reading. Although it is likely that there are factors not considered here which affect the relative prominence of these readings, it is hoped that such effects are independent of the fundamental mechanism of dependent reciprocal interpretation.

The examples reviewed in this section all pose problems for the scopal analysis of reciprocals, which claims that the range of a dependent reciprocal should be its putative non-local antecedent; as we have seen, the range of dependent reciprocals is always identical with the set that the *local* antecedent ranges over—even when this set never appears as a plural referent in the sentence. We now turn to the question of deriving the proper semantics for such constructions.

4.4 Dependent pronouns as restricted functions

We have seen that the interpretation of dependent reciprocals involves reciprocation over the same set that the local antecedent of the reciprocal, canonically a dependent pronoun, ranges over. I propose to derive the correct semantics by appealing to the range of the function represented by the dependent pronoun; in order to do this, it is of course necessary that the range of a reference function should be somehow retrievable. In order to accomplish this I amend the functional representation of pronouns to use *restricted* reference functions, i.e., functions that are only defined on some limited domain. We may further make the domain explicit by writing it into the definition of the reference function:

(4.18) $r = \lambda x\, \iota z\, (x \;\Pi\; \textsc{Ant}\ \&\ z = W(x))$

Here ANT (for *antecedent*) is an open variable, the plural individual that is coextensive with the domain of r. The function $W(x)$ is some (unrestricted) reference function of the type considered until now, that is, the version used by Engdahl (1986). For example, the split dependent pronoun in (4.19a) would correspond to the function given in (b):

(4.19) a. John and Mary told Harry that *they* are neighbors.

b. $r = \lambda x\, \iota z\, (x \;\Pi\; \text{John} \oplus \text{Mary}\ \&\ z = x \oplus \text{Harry})$

The domain and range of a restricted function of this sort can be retrieved by application of Link's (1983) maximality operator σ , defined as follows:

(4.20) $\sigma x\, Px = \iota x\, ({*}Px\ \&\ \forall y({*}Py \rightarrow y \;\Pi\; x))$

Recall that $*P$ is the closure of P under the sum operation.[3] The operator σ selects the maximal individual for which $*P$ is defined; since $*P$ is by definition closed for sums, the value of σ is also the sum of all individuals x for which $P(x)$ is true. The domain and range of a restricted function r can then be retrieved by use of the following formulas:

(4.21) a. $\sigma y\, (\exists z\, r(y) = z)$ = DS(r) = The maximal y in the domain of r

b. $\sigma z\, (\exists y\, r(y) = z)$ = RS(r) = The maximal z in the range of r

Expression (4.21a) simply recovers the domain restrictor ANT. It denotes a (possibly) plural individual d with the property that any individual in the domain of r is a part of d; this is not, strictly speaking, the domain of the function r (which is a *set* of possibly

[3] A short review of the semantics of plurals can be found in section 1.3.

overlapping individuals), but is sufficient for our purposes. I will refer to formula (4.21a) as the *domain sum* of r, and abbreviate it $\mathrm{DS}(r)$.

Similarly, I will refer to expression (4.21b) as the *range sum* of r, and abbreviate it $RS(r)$. As long as we restrict ourselves to quantification over atomic individuals, we can use domain and range sums instead of the domain or range without loss of information. But the (true) range of a split-dependent pronoun like the one represented by function (4.19b) consists of non-atomic individuals. Atomic individuals (e.g., *Mary* or *Harry*) are never returned by r, so they are not in the range; but they are part of the range sum. If it is desirable to only obtain individuals that are actually generated by r, we must either use the true range or obtain our individuals by applying r to the parts of the domain:

(4.22) a. $\forall x(x \in \text{Range}(r)) \ldots$

b. $\forall x(\exists w \in DS(r) \;\&\; x = r(w)) \ldots$

Provided the *domain* of r is properly described by quantifying over atomic individuals only, either of the above expressions will let x take only the values *John*$\oplus$*Harry* and *Mary*$\oplus$*Harry*. Note that if this condition holds we can write a predicate that is true only of elements in the (true) range of r:

(4.23) $\text{Range}(r) = \lambda x\, (\exists z \in DS(r) : x = r(z))$

Since the Heim et al. treatment of reciprocals on which I am basing this discussion involves distribution over atomic individuals only, all of these alternatives give us exactly the same results. Even when we consider examples involving predicates that do not distribute down to individuals, it is difficult to choose between these alternative ways of supplying the reciprocal with the range of the pronoun function. In section 4.6 I will discuss some examples which suggest that the reciprocation should only be applied to individuals that are actually in the range. This suggests that one of the forms in (4.22) should probably be used. In chapter 5, however, I adapt the present treatment of dependent reciprocals to the framework of Schwarzschild (1996); his treatment of reciprocals independently requires the values substituted into the object of the reciprocal predicate to be drawn from a *cover*, a contextually salient collection of individuals that includes some individuals, atomic or plural, and excludes others. (See section 5.3.2). This condition ensures that only appropriate individuals will be included, and renders redundant the use of the more powerful formalisms shown in (4.22).

Since the case for the stronger versions is not clear, I will somewhat arbitrarily adopt the range sum operator RS. This choice should be understood as provisional, subject to revision if an independent argument could be provided for or against any of the possible alternatives.

To make the restricted function available to the reciprocal, we also need to modify the representation of the pronoun in LF. In Engdahl's (1986) formulation, the pronoun is a free variable that may have a functional translation such as $S(u)$, where the argument u is a variable bound by a higher quantifier. Let us express this with the bipartite structure [R u], where R is a free variable interpreted as a restricted reference function. After Functional Application, the value of the pair is $R(u)$. (I assume that this can freely type-raise to $\lambda P\, P(R(u))$ as necessary).

4.5 A nonscopal semantics for reciprocals

We are finally ready to define a reciprocal operator that finds its range from the reference function corresponding to its antecedent. We begin with the translation of Heim et al. (1991b), who treat the reciprocal as a VP operator that contributes a universal quantifier with scope over the VP. Their translation for the reciprocal was given earlier as (4.5), and is repeated here. Given our definition of dependent pronouns, we can rewrite this as (4.24), eliminating non-local binding of the reciprocal without loss of coverage.

(4.5) [each$_j$ other(i)]$_k \Rightarrow\ \lambda P\, \lambda y\, \forall x_k (x_k\ \Pi\ X_i\ \&\ x_k \neq y)\, P(y, x_k)$

(4.24) [each$_j$ other(i)]$_k \Rightarrow\ \lambda P\, \lambda y\, \forall x_k (x_k \in RS(r)\ \&\ x_k \neq y))\, P(y, x_k)$

The range of the reciprocal is computed from the function r, supplied as a free variable that is constrained to match the antecedent's restricted reference function. In the system of Heim et al., the contrast argument of the reciprocal is provided through binding by a distributor, and the range is a free variable that must be coindexed with the sister of the contrast argument's binder (Heim et al. 1991a:fn. 3). In the proposed bipartite representation, $\lambda x\, r(x)$ appears in the right place—the same position that the domain of a distributor appears under the independent reading. Compare the constituent structure of the independent reading in (a) with that of the dependent reading in (b):

(4.25) a. [[John and Mary]$_1$ D$_4$]$_4$ think [they$_1$ D$_2$]$_2$ like [each$_2$ other]$_3$

= John and Mary think that [John and Mary like each other].

b. [[John and Mary]$_1$ D$_2$]$_2$ think [$\lambda x\, r(x)_1\ u_2$]$_2$ like [each$_2$ other]$_3$

= John thinks John likes Mary, Mary thinks Mary likes John.

In the dependent reading (b), the argument y of the reciprocal predicate (which serves as the subject of the reciprocal predicate and also as the contrast argument of the reciprocal) is provided by the dependent pronoun *they*$_2$. This has been translated as [[$\lambda x\, r(x)$] [u_2]], i.e., evaluates to $r(u_2)$ after its components are combined via Functional Application. The variable u_2 is bound by the matrix distributor D$_2$, and (4.25b) translates as in (4.26). In this example r is the restricted identity function, so the range of r contains just John and Mary and $r(x_2)$ is just x_2.

(4.26) $\forall x_2(x_2\ \Pi\ j \oplus m)$ think$(x_2, {}^\wedge[\forall x_3(x_3 \in RS(r)\ \&\ x_3 \neq r(x_2))$ like$(r(x_2), x_3)]\ )$

In the dependent reading of example (4.27), r maps lawyers to their clients; its domain is the set of John and Mary's lawyers, and $r(x_2)$ is x_2's client. The reciprocal thus pairs each client to all clients disjoint from him or her, as it should.

(4.27) The lawyers that represent John and Mary advised them to sue each other.

In this way the proposed translation can account for dependent reciprocal sentences, even when the remote binder is not coreferential with the local antecedent. But what about simple reciprocal sentences, like *John and Mary like each other?* I will assume that a referential NP can always type-lift to the restricted identity function on the subparts of that NP. Of course if it denotes an atomic individual, the range of the resulting function will have only one element and no reciprocation will be possible.

As I have repeatedly stated, the representation of distribution as universal quantification over atomic parts is an oversimplification; the account I propose shares in this. One aspect that should be pointed out in particular is the distinctness condition $x_k \neq y$. It is perfectly adequate when only atomic individuals are being compared, since such individuals can only be either identical or disjoint. But plural individuals may have some proper part in common, or one may be part of the other, without the two being identical. What should the reciprocal's distinctness condition do in such cases? Are pairs of overlapping individuals allowed to be paired by the reciprocal, and what kinds of overlap are permitted? I will consider these questions in section 4.8; for the moment I will simply retain the simplistic $x_k \neq y$, but it should be read as a placeholder for a distinctness condition which will be made precise in section 4.8.

4.5.1 The combinatorics of the proposed analysis

The function r that the reciprocal uses is not supplied to the reciprocal as an argument, but remains an open variable. Consequently one could imagine passing to the reciprocal a function that has a larger domain than its antecedent. For example, suppose that the range argument of the reciprocal in example (4.2), under the dependent reading, was based on the identity function on the set $\{$John, Mary, Kate, Amanda$\}$. The resulting translation would say that John thinks he likes Mary, Kate and Amanda, and Mary thinks she likes John, Kate and Amanda. This reading, of course, is impossible.

(4.2) John and Mary think they like each other.

In its essence, this is a weakness that this account shares with that of Heim et al. Recall that they need to stipulate that the range argument of the reciprocal is coindexed with the sister of the reciprocal's binder, the NP to which the binder-distributor is adjoined. Correspondingly, I ensure the proper choice of reference function by requiring r to be the *antecedent's* reference function; one can imagine a rule that coindexes the free variable representing the reference function in the reciprocal and in its antecedent, forcing variable interpretation to assign them the same value.

The reference function can then be found via the same structural description, *sister of the binder's head*, that Heim et al. used to find the reciprocal's range argument. (An illustration of the parallelism was given by example (4.25)). For this reason, I consider the proposed analysis to be neither better nor worse than that of Heim et al. with respect to specifying the range argument. The difference, however, is that Heim et al. apply this structural description to the matrix binder in the case of dependent reciprocals, while the present analysis always applies it to the local antecedent.

Consequently the present analysis can be said to be "nonscopal" in the same sense that Heim et al. describe their own analysis as "scopal." As we have seen, by adopting the less "scopal" analysis we can handle a wider range of dependent reciprocal examples.

The analysis developed here is wholly consistent with my representation of the Heim et al. system with respect to binding of the reciprocal's *contrast* argument: both systems identify the contrast argument with the subject of the reciprocal predicate. This is actually inconsistent with the intent of the Heim et al. analysis: as discussed in section 2.1.5, Heim et al. intended the contrast argument to be determined by the binder of the reciprocal, that is, by the matrix distributor in the case of dependent reciprocals. But as we have seen, the contrast argument should in fact always be the local antecedent.

My analysis follows Heim et al. in one important respect: the independent and dependent readings given in (4.25) differ with respect to the location of the distributor that provides quantification over the subject position of the reciprocal predicate. In the independent reading (a) the distributor is adjoined to the embedded subject, while in the dependent reading the quantification is provided by the matrix distributor. In the present, "nonscopal" account, this asymmetry can be considered incidental to the translation of the reciprocal, since both its range and its contrast argument depend on the local antecedent in both cases.[4]

4.5.2 Larger dependent NPs

The dependent reading of example (4.28) is not readily handled by the analysis presented thus far.

(4.28) John and Bill want their clients to sue each other.

[4]This point of view reflects the approach of Williams (1991), who argued that the difference between the readings given in (4.25) was the result of asymmetries in the interpretation of the *pronoun*, not of the reciprocal. Williams's argument greatly influenced my own approach to such examples. Unfortunately, as we will see in section 5.1.4, Williams does not provide a workable framework for a nonscopal analysis of reciprocals.

The dependent reading of this example says that John wants his client to sue Bill's client, and vice versa. Here the NP *their clients* plays the role of a function mapping John and Bill to their clients; the problem lies in showing how the reciprocal can be making use of this NP, which is not in the form of a paycheck pronoun. The details depend on the analysis of possessives one adopts. The one I sketch here is along the lines proposed by Jacobson (1999b).

For *Y* a relational noun like *client*, let us take an NP of the form *X's Y* to be translated as $\iota z\, Y'(X', z)$. For simplicity we will ignore the question of exactly how to derive this compositionally (it is easy to simply assign the appropriate translations to the genitive morpheme *'s*, and to the possessive pronoun).

In the NP *their clients*, the possessor *their* is an ordinary dependent pronoun, i.e., it takes values identical to those of the distributor that binds it. The structural description we have used until now to determine the reciprocal range will lead us to the pronoun: it is the sister of the head of the reciprocal's *local* antecedent, the NP *their clients.* Compare the analysis of example (4.25b) with the structure assigned to sentence (4.28):

(4.25b) [[John and Mary]$_1$ D$_2$]$_2$ think [$\lambda x\, r(x)_1\ u_2$]$_2$ like [each$_2$ other]$_3$

= John thinks John likes Mary, Mary thinks Mary likes John.

(4.29) [[John and Bill]$_1$ D$_2$]$_2$ want [their$_1$ clients$_3$]$_3$ to sue [each$_3$ other]$_4$

= John wants John's client to sue Mary's client, Mary wants Mary's client to sue John's client.

This means that a purely structural criterion for the range of the reciprocal would select the dependent pronoun *their* as the source of the reciprocal's range argument. We could treat the pronoun as the identity paycheck function, rather than as a simple variable, but it is the *wrong* function: the pronoun by itself expresses the identity function, but what we need is the function taking lawyers to clients! It is necessary, then, to include the relational noun *clients* in our derivation of the reference function. Perhaps then we should take the noun *clients* as our reference function, as usual with the smallest possible domain (the set of John and Mary's lawyers)? To begin with, it is not clear how this could be accomplished by an algorithm that also works for the ordinary dependent cases. Worse, this approach would imply that common nouns such as *client* have domains that depend on the context of use. I find this more than a little problematic: pronouns inherently depend on an antecedent for their interpretation; if they represent a function, it is reasonable for the antecedent to partially determine the properties of the function, including its domain. But common nouns are not anaphoric in the same way; to link the interpretation of *client* with that of *John and Mary* seems arbitrary.[5]

A purely structural criterion, then, will not identify the proper range argument in this case. The alternative is to use the domain of the pronoun and somehow combine it with the function encoded by the relational noun. However, it is not clear how this should be done compositionally. I do not present a treatment here, because the problem has an elegant solution in the framework of Variable Free Semantics developed by Jacobson (1999a,b). An analysis along those lines is presented in chapter 6; until then, let us simply assume that the presence of an NP containing a dependent pronoun allows us to deduce a suitable reference function: this function performs the same mapping between individuals as the NP containing the dependent pronoun, and has the same domain as the embedded pronoun.[6]

[5]Perhaps an indirect connection is more plausible: the noun's domain could be said to be determined by the domain of its specifier, the possessive pronoun *their.* But even this involves a degree of context sensitivity that seems unwarranted for common nouns.

[6]One way of achieving this is to compute the reference function by lambda abstraction over the antecedent of the reciprocal: if the translation of the dependent antecedent is φ, the reference function can be written as $\lambda u\,\varphi$, where u is the name of the open variable that saturates the argument of the dependent pronoun embedded in φ. Since this variable should remain open in the antecedent itself, this operation would need to be performed "off-line," not as part of the compositional derivation of the sentence, to compute the value of the reference function.

4.6 Split-dependent reciprocals

Split dependent pronouns, which we saw in section 3.6, pick out a series of plural individuals. Accordingly, we expect them to function as reciprocal antecedents in two different ways: either by reciprocation over the plural denotation of the pronoun itself, or as the local antecedent in a dependent reciprocal construction analogous to those we have been discussing. But as we will see in this section, we find that only the first type is actually possible.

Sentence (4.30a) has a reading given in (b)/(c), in which each man urges mutual support between himself and Mary. (The pronoun represents the function $\lambda x\, x \oplus \text{Mary}$). Although this should be classified as a dependent reading, it differs in an important way from the dependent reciprocal readings we have been considering. Its analysis under the system of Heim et al. would not involve a wide scope reciprocal, but would be as given in (c). Here we have two distributors, one within the scope of the other. The embedded distributor D_j ranges over whatever each value of *they* refers to; the reciprocal must be bound by the embedded distributor, mapping, for example, Tom to Mary and Mary to Tom when *they* refers to Tom $\oplus$ Mary.

(4.30) a. Tom, Dick and Harry told Mary that *they* should support each other.

b. Tom told Mary that they$_{t,m}$ should support each other$_{m,t}$.
Dick told Mary that they$_{d,m}$ should support each other$_{m,d}$.
Harry told Mary that they$_{h,m}$ should support each other$_{m,h}$.

c. [(T,D&H) D$_i$] told Mary that [they$_i$ D$_j$] should support [each$_j$ other]

This reading, then, does not involve a "long-distance," or dependent, reciprocal: The distributor adjoined to the local antecedent of the reciprocal controls its range set, under the Heim et al. analysis as well as my own.

Sentence (4.30a) can also have the singular (non-split) dependent reading, shown in (4.31). In this reading the pronoun *they* refers in turn to Tom, Dick and Harry alone, and reciprocation is among them only; Mary is involved in neither supporting nor in being supported. As usual, sentence (4.30a) also has any number of "fixed" readings, under which Tom, Dick and Harry have stated the same proposition about mutual support by some group of people: for example, that the Rockefellers should support each other.

(4.31) Tom told Mary that Tom should support Dick (or Harry), *and* ...
Dick told Mary that Dick should support Harry, *and* ...
Harry told Mary that Harry should support Tom.

(4.32) Tom Dick and Harry each told Mary that:
a. ... the Rockefellers should support each other.
b. ... Tom, Dick and Harry should support each other.
c. ... Mary, Ann and Harry should support each other. (Etc.)

Leaving aside the fixed readings, as usual, we find that reading (4.30b) is the *only* way the reciprocal can be interpreted when the pronoun has the split-dependent reading. If the reciprocal could be bound by the matrix distributor, it would give the following structure:

(4.33) a. [(T,D&H) D_i] told Mary that they_i should support [each_i other].
b. $\forall x_i \in \{T, D, H\} \, \forall x_j (x_j \in \{T, D, H\} \;\&\; x_j \neq x_i)$ [x_i told Mary that $x_i \oplus$ Mary should support x_j]

This structure does not correspond to an actual reading: it would say, among other things, that Tom told Mary that Tom and Mary should support Dick.

Since we have seen dependent readings with other dependent pronouns that were translated via non-identity functions such as λx *teacher-of*(x), we might expect a dependent reading under which the reciprocal matches each pair of the form $x \oplus$ *Mary* to some other pair or pairs of the same form:

(4.34) Tom told Mary that Tom and Mary should support Dick and Mary (also, Harry and Mary), etc.

This reading is not possible, either.

The analysis I have proposed does not predict the absence of these readings.[7] What rules them out, then? One possibility is that the reciprocal must be construed with the closest available distributor, which in this case is the embedded distributor. Alternately, the reason may be that the sets over which *they* ranges are not disjoint, but have *Mary* in common: *Tom* $\oplus$ *Mary, Dick* $\oplus$ *Mary, Harry* $\oplus$ *Mary.* If the reciprocal operator is in fact required to match individuals that have no part in common, reading (4.34) would be ruled out, since no licit reciprocation is possible.

In the next section I will present some more constructions involving non-atomic local antecedents for the reciprocal. As we will see, their behavior tips the scale towards the first of the two alternatives.

[7]The Heim et al. (1991a,b) account, which requires identity between the local and the remote antecedent of the reciprocal, would correctly rule out the long-distance readings of (4.30a) simply because the split-dependent pronoun does not represent the identity function on its binder. But as we have seen, this restriction incorrectly rules out many grammatical long-distance reciprocals.

The 1991a version of their theory would also rule out the long-distance readings because they would generate relations involving plural entities: *Tom* $\oplus$ *Mary, Dick* $\oplus$ *Mary,* and so on. Their framework, via the definition of $\dot{\Pi}$, requires that the reciprocal variables should be singular. But this is not a deep property of their system; indeed, it is absent from the revised, 1991b version, which uses Π in the place of $\dot{\Pi}$ (*proper-part-of* instead of *proper-atomic-part-of*). At any rate, this condition is too strong: Sentence (i) is readily interpreted as saying that the Smiths do not like Tom, and vice versa. (This reading is readily generated by Schwarzschild's (1996) treatment of distributivity by means of *covers*, discussed in section 5.3).

(i) Tom and the Smiths don't like each other.

4.7 Reciprocals and collective action

Let us examine the hypothesis that a dependent reciprocal must always depend on the nearest available distributor. Fortunately it is possible to test this without resorting to reciprocation between two non-atomic individuals. Imagine that John and Mary have each written a book. In addition, they are co-owners of a small publishing house, and they both wish to publish their two books in their own press. However, their business is so precarious that they can only afford to publish one of their two books this year. Since they are good friends and also altruistic, each of them feels that their press should publish the other's book. My consultants consistently find the following sentence an acceptable description of this state of affairs:

(4.35) John and Mary think that they$_{\mathrm{J+M}}$ should publish each other.

This example is a bit surprising, since the local antecedent does not appear to be a dependent pronoun. Since we have seen that the long-distance reading (i.e., what we have been calling the dependent reading) requires a dependent local antecedent, we must conclude that the antecedent is actually a dependent pronoun. Perhaps it represents the constant function that maps everything to $J \oplus M$, or perhaps the split-antecedent function $\lambda x\, x \oplus J \oplus M$ (recall that according to the semantics of plurals we assume, $J \oplus J \oplus M = J \oplus M$).

Now consider another test case: Suppose that Al and Bill have both been nominated as candidates for the presidency of their social club. They both feel that they should pool their votes in order to give one of them a bigger chance of winning the election; since they are good friends and also altruistic, each of them wants the other to be the president. We might try to describe this situation with the following sentence:

(4.36) * Al and Bill think that they$_{\mathrm{A+B}}$ should vote for each other.

My consultants did *not* in general consider this a possible way to describe this situation. It only has an irrelevant non-dependent reading, which says that Al thinks "we should vote for each other", and so does Bill. However, sentence (4.36) was judged acceptable in a situation where Al and Bill together cast a *single* ballot. The crucial factor, then, is whether the embedded subject is interpreted distributively or collectively. The context that allows a collective construal for (4.36) provides it with an interpretation parallel to that of (4.35). The situations mentioned correspond to the following assignment of distributors and indices:

(4.37) a. John and Mary D_i think that they$_{\mathrm{J+M}}$ should publish each$_i$ other.
(John thinks "we should publish Mary")

b. * Al and Bill D_i think that they$_{\mathrm{A+B}}$ D_j should vote for each$_i$ other.
(Al thinks "we should vote for Bill")

c. Al and Bill D_i think that they$_{\mathrm{A+B}}$ D_j should vote for each$_j$ other.
(Al thinks "we should vote for each other")

d. Al and Bill D_i think that they$_{\mathrm{A+B}}$ should vote for each$_i$ other.
(Al thinks "we should cast our joint vote for Bill")

We conclude that the reciprocal must always be associated with the distributor closest to it, even if the latter modifies a dependent pronoun bound by another distributor. Note that the embedded pronouns in (4.37b,c) and (d) have the same binder, the distributor over the matrix subject. But in (b)/(c), the lower distributor cannot be skipped over by the reciprocal, and forces the independent ("narrow scope") reading.

This condition is sufficient to explain the absence of reading (4.34) for sentence (4.30a) (repeated below): *support* is interpreted distributively, and the distributor over the embedded subject blocks the reciprocal from ranging over the members of the matrix subject.

(4.30a) Tom, Dick and Harry told Mary that they should support each other.

(4.34) $\neq$ Tom told Mary that Tom and Mary should support Dick and Mary (also, Harry and Mary), etc.

Let us try to confirm this analysis by considering a split-dependent example in which that the values of the dependent pronoun have no fixed part:[8]

(4.38) Q: What did the women tell you about themselves and their husbands?

A: They told me that *they* are richer than each other.

$f(x) = \lambda x\, x \oplus \text{husband-of-}(x)$

The reading of interest is the one in which each woman said that she and her husband are richer than the other couple in question. Since there is no overlap in the members of the couples whose riches are being compared, the well-formedness of this example should depend on the arrangement of distributors only: The closest-distributor condition predicts that a distributor over the members of the couple would block the reciprocal from ranging over couples. The legitimate readings of this example are far from clear, but several of my consultants accept it. To the extent that it exists, however, it seems to treat being rich as a collective property of the husband and wife teams; thus there is only one distributor, and this example is consistent with the closest-distributor condition.[9]

[8] This type of split-dependent pronoun was brought to my attention by an anonymous reviewer for SALT 9.

[9] The following example, which was deemed acceptable under its dependent reading by some of my consultants, appears to violate the closest-distributor condition.

(i) John and Mary think their parents should visit each other.

I do not have an explanation for this inconsistency at present.

4.7.1 Reciprocation without distributors

It is an old chestnut in studies of reciprocity that the reciprocal forces a distributive interpretation on its antecedent. This was never intended to be a profound discovery: reciprocal sentences describe a web of relationships between various actors and patients (or subjects and objects, or whatever), so without distribution, where would the variety come from? Oblivious to the force of this argument, however, most of my non-linguist English language consultants insist that sentences like (4.39) are acceptable under their collective action reading:

(4.39) Bill and Peter, together, carried the piano across each other's lawns.

Although such examples have for the most part remained below the radar of the linguistic literature, they are not completely unknown. Moltmann (1992:fn. 3) mentions the following example, which she attributes to Higginbotham:

(4.40) John and Mary divided each other's belongings (among themselves).

Moltmann suggests that this example should be analyzed either by distributing over the predicate's event argument rather than over the subject, or by allowing the reciprocal to relate the entire antecedent to its parts instead of relating parts to parts. Moltmann also gives the following example:

(4.41) a. John and Mary gave themselves books about each other.

b. Hans und Maria schenkten sich Bücher übereinander.
Hans and Maria gave them (refl.) books about each other

Example (a) has a reading under which John and Mary, as a group, gave to John and Mary as a group a book about John, and they also gave to John and Mary as a group a book about Mary. Moltmann notes that the German version of this example, (4.41b), lacks this reading.[10] This raises the question, which I have entirely avoided in this dissertation, of the extent to which the fine points of reciprocal interpretation are subject to cross-linguistic variation.

How should such examples be interpreted? Note that the translation of an ordinary, distributed reciprocal sentence involves two universal quantifiers, one distributing over the subject and one ranging over the reciprocal's range argument. The meaning of the present examples, on the other hand, involves a non-distributed antecedent plus, once again, quantification over the reciprocal's range. In other words, the correct translation should involve a single universal quantifier.

This means that the "each-raising" account of Heim et al. (1991a) could not do the job: it introduces *two* universal quantifiers, one for *each* and one for the remainder, and claims that the *each* part of the reciprocal must necessarily raise to become a distributor over the reciprocal's antecedent. The proposed reciprocal translation (4.24), however, which is based on the (Heim et al. 1991b) account, assumes that distribution over the subject is not due to the reciprocal but to an independently inserted distributor. Therefore the translation of the reciprocal, according to either my account or that of (Heim et al. 1991b), only contains *one* universal quantifier; and that is the desired number for the translation of sentence (4.39). In other words, to properly interpret sentence (4.39) we need to simply translate the *each other* according to formula (4.24), repeated below. The reciprocal VP will be translated as shown in abbreviation in (4.42).

(4.24) [each$_j$ other(i)]$_k \Rightarrow\ \lambda P\, \lambda y\, \forall x_k (x_k \in RS(r)\ \&\ x_k \neq y))\, P(y, x_k)$

(4.42) $\lambda y\, \forall x_k (x_k \in RS(r)\ \&\ x_k \neq y)$ carry-piano-across-lawn-of$'(y, x_k)$

[10] Both languages also allow the ordinary dependent reading, which says that Mary gave herself a book about John and John gave himself a book about Mary. Moltmann reports that German also allows a reading according to which John gave Mary a book about Mary, and vice versa. She attributes this reading to the ability of *sich* to the dual status of *sich* as a reciprocal as well as a reflexive.

Combining the reciprocal VP with the entire subject, without the mediation of distribution, results in the expression in (4.43). This is more or less right, saying that Bill and Peter, together, carried the piano across the lawn of every individual given by r that is distinct, in the appropriate sense, from *Bill* $\oplus$ *Peter.* This will be correct provided that r is the identity function over parts of the individual *Bill* $\oplus$ *Peter,* and that the distinctness condition treats *Bill* as distinct from *Bill* $\oplus$ *Peter.*

(4.43) $\forall x_k(x_k \in RS(r)\ \&\ x_k \neq B{\oplus}P)$ carry-piano-across-lawn-of$'(B{\oplus}P, x_k)$

Given that it is after all possible for a reciprocal to have a non-distributive antecedent, do we need to reconsider our analysis of example (4.35)? In the previous section I analyzed this as involving dependent reciprocal (i.e., "long-distance") binding, as in (4.37a):

(4.37a) John and Mary D$_i$ think that they$_{\mathrm{J+M}}$ should publish each$_i$ other.
(John thinks "we should publish Mary")

Since we now know that reciprocals can take non-distributed NPs as their antecedents, we must consider the possibility that the pronoun *they* is not a dependent pronoun after all, but is simply a non-distributed antecedent of the reciprocal of the same type as the examples seen in this section. But such an analysis would not give the correct semantics: it would ascribe to John the belief "we (together) should publish each other." The reading described by (4.37a) is a characteristically dependent reading, since John's belief is only about publishing Mary and Mary's belief is only about publishing John. Thus we must continue to treat such examples as dependent reciprocals.

The examples considered in this section are unusual in two respects. First, as discussed they involve reciprocation without distribution over the subject. And second, consequently, they relate a collectively interpreted plural subject to its own proper parts. For example, *John* $\oplus$ *Mary* is related to *John,* and *John* $\oplus$ *Mary* to *Mary.* The second point means that the proper distinctness condition for reciprocals cannot require complete disjointness. We will consider the implications of this in the next section.

4.8 The distinctness condition

Let us finally turn to the proper definition of the distinctness condition for the reciprocal, which was simplistically given in formula (4.24) as $x \neq y$. This question has not received much attention in the literature, since the better-studied properties of the reciprocal do not seem to depend crucially on it. An exception is Beck (1999), who studies in detail the position and presuppositional status, though not the actual content, of the distinctness condition. In this matter I will follow tradition and give no more attention than is necessary to this question.

Since Heim et al. (1991a) restricted their attention to distribution and reciprocation down to atomic individuals, it was sufficient to write the distinctness condition as a simple inequality, $x \neq z$. Here x represents the contrast argument and (typically) subject of the reciprocal predicate, while z is part of the range argument and is used as the object. (Since the variable z is supplied by the reciprocal by distribution over the range argument, let us refer to z for the purposes of this discussion as the *derived argument*). Although atomic individuals are either the same or different, two non-atomic individuals may be non-identical but have parts in common. If applied to plural individuals, the simple test of equality used by Heim et al. would pronounce a set of one hundred individuals to be distinct, in the sense relevant to reciprocals, from the set of ninety-nine of these same individuals.

Sauerland (1995b) employs the stronger condition given in (4.44). Here x is again the contrast argument and subject, y the range argument, and z is the derived argument. Sauerland's condition is motivated by the behavior of non-pronominal *other,* in examples such as

(4.45). The interpretation of this sentence suggests that the NP *the other student* must refer to the unique student that is not part of the group of two students already mentioned.

(4.44) $[\![\text{other}]\!](x)(y)(z)$ = 1 iff z is part of y and z is not part of x.

(4.45) Two of the three students lived in Cambridge. The other student lived in Somerville.

Apparently Sauerland was only concerned with atomic values for z. Note that his condition is asymmetric: it states that the derived argument z is not part of the contrast argument x, but not that x is not part of z. Moreover, even if neither x nor z is part of the other, it allows for the possibility that they could nevertheless both be non-atomic individuals with some part in common. If we allow z to take non-atomic values, it turns out that at least for non-reciprocal *other,* Sauerland's formulation is too weak. In (Dimitriadis 1999b) I used the following distinctness condition, which states that x and z must be completely disjoint:

(4.46) The contrast argument c is distinct from a derived argument z if and only if $c \wedge z = \mathbf{0}$.

Here "$\wedge$" is the meet operation on the semilattice of individuals, corresponding to intersection in the union model, and **0** is the lowest (empty) semilattice element; i.e., c and z can have no part in common. I chose this stronger condition because in an example like (4.47), the denotation of the NP *the others* clearly must be disjoint from the two boys already mentioned.

(4.47) Two of the boys brought food. *The others* brought wine.

This is probably the strongest conceivable distinctness condition between individuals, and it seems to be the correct one for non-reciprocal *other.* But is it the correct distinctness condition for reciprocals? Although Heim et al. (1991a) aspired to reduce the analysis of reciprocals to a union of the properties of non-reciprocal *each* and "pronominal" *other,* I showed in chapter 2 that the fine points of reciprocal interpretation cannot be predicted in this way. The data discussed in the last section, as the reader is probably aware, are incompatible with both Sauerland's (1995b) distinctness condition and with my own (4.46). Let us consider again sentence (4.35) under the intended reading:

(4.35) John and Mary think they should publish each other.

= John thinks they$_{\mathrm{J+M}}$ should publish Mary, and Mary thinks they$_{\mathrm{J+M}}$ should publish John.

In this example the antecedent of the reciprocal is *John* $\oplus$ *Mary*. According to the analysis of reciprocals I have presented, this is also the reciprocal's contrast argument. Therefore the individuals being related by the reciprocal (*John* $\oplus$ *Mary* with *John*, and *John* $\oplus$ *Mary* with *Mary*) are not disjoint, contradicting the disjointness condition (4.46) assumed in (Dimitriadis 1999b). Moreover, in this example the derived argument *is* a part of the contrast argument: *Mary* $\sqsubseteq$ *John* $\oplus$ *Mary*. This contradicts Sauerland's version of the distinctness condition.

Since the strict equality condition used by Heim et al. is clearly too weak, we can maintain the remainder of the proposed analysis of reciprocals by adopting the following distinctness condition:

(4.48) The contrast argument c is distinct from a derived argument z if and only if $c \not\sqsubseteq z$.

This is simply the inversion of Sauerland's condition: It says that the contrast argument c should not be part of the derived argument z, while Sauerland's condition required that z should not be part of c. While it is consistent with the limited range of examples we have examined, it is not the only conceivable condition that is consistent with them;[11] but given the

[11]For example, the disjunctive condition "x and z are disjoint *or* z is a part of x" is also consistent with the examples under discussion. Unlike formula (4.48), this distinctness condition would not allow the reciprocal

unreliability of judgements about examples involving reciprocation between non-atomic individuals, I will not attempt to evaluate the merits of formula (4.48) against those of more complex alternatives.

4.9 Conclusions

This chapter has defended an analysis of dependent ("long-distance") reciprocals that can account for the problematic examples, involving paycheck-style dependent pronouns, discussed in chapter 2. This involved the claim that the reference functions of dependent pronouns have domains, and that the translation of long-distance reciprocals directly references such domains. In the resulting theory, the range of the reciprocal always depends on its local antecedent only, which is also the contrast argument; in dependent readings, an open variable in the antecedent is itself bound by a non-local distributor, but the range argument is still determined by the local antecedent. This state of affairs somewhat reflects Williams's (1991) approach to reciprocals, according to which the dependent and independent reciprocal readings only differ in whether or not the embedded antecedent is interpreted as bound pronoun.

In comparison, the account of Heim et al. (1991a,b) involves explicit dependence of the reciprocal on the long-distance binder, which binds the contrast argument as well as determine the range argument; the local antecedent is only involved in satisfying syntactic binding conditions. As we have seen, it is not possible to account for the full range of dependent reciprocal examples without involving the local antecedent in the computation of the range argument.

The treatment of reciprocals provided in this chapter is far from complete. For concreteness, I have based my discussion on the analysis of Heim et al. (1991b), and have not

predicate to be applied to two partially overlapping individuals, neither of which is a subset of the other. For example, it would treat $John \oplus Mary$ as sufficiently distinct from $John \oplus Bill$.

addressed any shortcomings of their account that are orthogonal to the issue of reciprocal scope. In particular, the resulting semantics for reciprocals follow Heim et al. in improperly requiring strong reciprocity (that is, universal quantification over both arguments of the reciprocal predicate). More recent treatments, such as those of Sternefeld (1998) and Schwarzschild (1996), are better at capturing the nature of reciprocation; but while their treatments of dependent reciprocals differ, they share with the Heim et al. analysis the inability to handle the full range of dependent reciprocal examples: The range of the reciprocal is incorrectly predicted to match the long-distance binder, not the local antecedent. Chapter 5 examines these matters in some detail. In section 5.4, I show how the account proposed here can be straightforwardly combined with Schwarzschild's (1996) treatment of distribution and reciprocals in order to address many of these issues.

The mechanism I have outlined here expresses, in a limited way, the dual function of dependent reciprocal antecedents as singular bound variables and as the set-denoting range argument of the reciprocal. The proposed analysis relies on a free variable whose value is determined by syntactic structure (namely, the range of the reciprocal is based on a function-valued free variable that must be coindexed with the functional part of the reciprocal's local antecedent). Although this structural requirement is identical with the requirement adopted by Heim et al., it is still not ideal. Such free variables are harder to constrain that the strict Principle A behavior of reciprocals would have us wish for; they carry the unrealistic expectation that the usual interpretation could be overridden, if only the right pragmatic context could be found. As we have seen, this way of utilizing function domains is also ill-suited to dealing with examples such as (4.28), in which a dependent pronoun is embedded in a larger NP that constitutes the antecedent of the reciprocal. In this case, the antecedent of the reciprocal does not contain its reference function in an accessible form.

(4.28) John and Bill want their clients to sue each other.

So far I have accounted for these examples by urging the reader to assume the existence of a reference function corresponding to such NPs. But a more direct approach is possible in the framework of Jacobson's (1999a) *variable-free semantics,* in which all pronouns are represented as functions rather than as assignment-dependent variables. In this system it is straightforward for the reciprocal to directly access the restricted function represented by the dependent pronoun, eliminating the need for an unbound variable. I will develop the technical details of this approach to reciprocal interpretation in chapter 6.

Chapter 5

Other treatments of reciprocals

Although in the preceding chapters I restricted my discussion to the scopal analysis of reciprocals proposed by Heim et al. (1991a,b), their approach to dependent reciprocals is reflected in most other analyses of reciprocals that I have come across; in consequence, these analyses share with the Heim et al. analysis their inability to take into account the role of the local antecedent in the interpretation of dependent reciprocal sentences like the following:

(5.1) a. Their coaches think they will defeat each other.

b. The lawyers that represent them$_i$ say they$_i$ will sue each other.

c. John and Mary think their mothers like each other.

In this chapter, I will review briefly three more treatments of reciprocals. Section 5.1 presents the reciprocal analysis of Williams (1991), who is a rare exception in defending what he calls a non-scopal analysis of reciprocals: one that consistently treats the local antecedent of the reciprocal as its semantic antecedent, even in dependent reciprocal constructions. Unfortunately, Williams does not make any progress on the central problem of how a compositional semantics could derive the reciprocal's *plural* range argument from its singular, bound-variable antecedent.

Sections 5.2 and 5.3 review two scopal analyses of reciprocals, those of Sternefeld (1998) and of Schwarzschild (1996). Like those of Heim et al. and Williams (1991), each of these treatments is also a theory of plurals. Sternefeld's analysis is superficially very different from the Heim et al. account, but it nevertheless shares the essential elements of the scopal analysis. Schwarzschild's account is of additional interest because I believe that it provides a good basis for addressing various issues that have been beyond the scope of the Heim et al. treatment (and largely, of my own). In section 5.4, I show how the approach I developed in chapter 4 can be combined with his analysis.

5.1 The non-scopal view: Williams (1991)

Perhaps because Williams's (1991) account has the disadvantage of not having an obvious translation into Montague semantics, it has not inspired other treatments of reciprocals along the same lines. In this section I present a version of Williams's viewpoint and his analysis, but do not attempt to provide a formal semantics for his system. Instead, I will highlight the insights that I find particularly compelling; they have been incorporated into the account presented in chapter 4, which is more directly derived from the Heim et al. (1991b) analysis.

Williams's very short paper covers a lot of ground with few examples, and in filling in the details of his analysis one has to extrapolate considerably. My elucidation of Williams's linking framework turns out to be markedly different from the one worked out by Heim et al. in their (1991b) response. Since I do not intend to resurrect Williams's linking framework, but merely to highlight his insightful generalizations, I will only give a brief account of its detailed mechanics and will not discuss the criticisms directed at it by Heim et al. (1991b).

Williams (1986, 1991) maintains that the ambiguity involved in what Heim at al. call the "puzzle of scope" (example (5.2)) is "nothing more than the ambiguity latent in (5.3)":

(5.2) John and Mary think they like each other.

a. John thinks John likes Mary and Mary thinks Mary likes John. (dependent)

b. John and Mary think that John and Mary like each other. (independent)

(5.3) They think they are sick.

a. John thinks John is sick and Mary thinks Mary is sick. (dependent)

b. John and Mary think that John and Mary are sick. (independent)

Williams calls the (a) readings *distributed,* the (b) *group* readings; they are what I called *dependent* and *independent,* respectively.

The independent reading of the embedded pronoun in (5.2) is only compatible with what the scopal account of Heim et al. characterizes as "narrow" apparent scope for the reciprocal, while the dependent reading is only compatible with "wide" scope: there are only two readings rather than the four that might be expected if reciprocals can be assigned scope.[1] Heim et al. can rule out the missing readings, so this is not exactly a problem for their theory; but given the independent need for multiple construals of the embedded pronoun, it certainly lends appeal to Williams's thesis: that the ambiguity in (5.2) is wholly due to the possible construals of the pronoun *they,* not to any properties of the reciprocal. Williams's intent is that the embedded pronoun should depend on the matrix subject, and the reciprocal should depend on the embedded pronoun only; the two readings of the sentence should then be derived according to which kind of relationship (identity or distribution) there is between the dependent pronoun and its antecedent.

[1]The possible readings are discussed in section 2.1.6. A narrow-scope reciprocal must distribute over an independent embedded pronoun (a dependent embedded pronoun would be singular, giving the reciprocal nothing to distribute over), and a wide-scope reciprocal is only grammatical when combined with a dependent pronoun. (The dependent pronoun satisfies the reciprocal's Principle A requirement).

5.1.1 Weak and strict distribution, and floated *each*

Another empirical motivation for Williams's solution comes from the observation that non-reciprocal "floated" *each* introduces stricter truth conditions than do reciprocals and implicitly distributed plurals; for example sentence (5.4a) is false in a situation where there was a lot of hitting but there were some non-hitters, but sentences (b) and (c) are true.

(5.4) a. They were each hitting the others.

b. They were hitting each other.

c. They were hitting Bill.

Williams concludes that there are two kinds of distribution-like relationships, which he calls *weak* and *strict* distributivity; the former corresponds to the semantics of reciprocals and implicitly plural NPs, the latter to the semantics of non-reciprocal *each.* Weak distributivity is not "enforced down to the level of individuals," while strict distributivity is.[2]

The difference in the semantics of reciprocals and non-reciprocal *each* correlates with their behavior in sentences containing another reciprocal: while it is possible for two reciprocals to have the same antecedent, "floated" *each* cannot share an antecedent with a reciprocal.

(5.5) a. They gave each other pictures of each other.

b. * They each gave each other pictures of the other.

Floated *each*, but not reciprocals, also blocks plural interpretation of reflexives. Example (5.6a) is compatible with the "we" reading (along with "I" and "you" readings); it says

[2]Some of the issues involved in defining distributivity were reviewed in section 1.4. Williams's definitions are not entirely compatible with the definition of weak distributivity given there: that definition makes no provision for exceptions, i.e., individuals who did not participate in the stated action at all, while Williams includes exceptions in the scope of his proposal.

that each of the pictures exchanged depicted John and Mary together. But in (5.6b), which contains floated *each*, the reflexive can only have the "I" or "you" readings.

(5.6) a. John and Mary gave each other pictures of themselves.

b. John and Mary each gave the other pictures of themselves.

Finally, Williams notes that the same elements pattern together with respect to the syntactic plurality of the object NP in the following sentences:

(5.7)	a.		They each have a new nose.	(strict)
	b.	*	They each have new noses.	
(5.8)	a.	*	They have a new nose.	(weak)
	b.		They have new noses.	
(5.9)	a.	*	They gave each other a new nose.	(weak)
	b.		They gave each other new noses.	

Not everyone agrees with Williams's judgements. Many of the English speakers I consulted accepted the plural as well as the singular version of example (5.7); but it is still true that this example, which involves "floated" *each*, patterns differently from the other two, which involve implicit or reciprocal distribution and require plural NPs.

Williams concludes that a strictly distributive relation must be between a higher plural NP and a singular lower NP, creating a singular variable that is, for example, not suitable as an antecedent of a reciprocal; while a weakly distributive relation is between two plural NPs, either of which can serve as a reciprocal antecedent.

The above contrasts, Williams points out, hold even when the antecedent refers to just two individuals, even though in such cases the two types of distributivity have the same truth conditions. It follows that the syntactic contribution of floated *each* differs grammatically from that of the *each* of *each other*, and from that of implicit distributors.

As we saw in section 2.2.2.2, the “strongly distributive” non-reciprocal *each* is also able to interact scopally with other quantifiers, while distributed NPs are not. We can conclude that the contrast is between true quantifiers, which undergo QR and range over the atomic elements of their complement, and distributed NPs, which do neither. Williams does not make this generalization, since he does not include quantificational NPs in his classification. Instead, he attempts to capture the difference within an approach to distributivity that is intended to be irreducibly relational.

5.1.2 The linking theory

In order to express the relationships discussed above, Williams (1991) proposes that the indices commonly used to express syntactic and referential relationships should be replaced by directed links that encode a variety of relationships between their endpoints. His proposal, which follows along the lines of earlier work by Higginbotham (1983, 1985) and Williams (1986), is summarized as follows:

(5.10) a. There are two kinds of links, or binding relations: distributing and non-distributing.
b. A nondistributing link links a singular to a singular or a plural to a plural.
c. A (strictly) distributing link links a plural to a singular.

A plural pronoun can be linked to its antecedent using either kind of link, illustrated as follows:

(5.11) They think they are smart.
d/n

A distributing link imposes strict distributivity, while a non-distributing link allows collective or weakly distributive interpretations. *Each other* links to its antecedent nondistributively as a plural, whereas *each* imposes a distributive link between the subject and the VP. A distributive link cannot be headed by a singular NP, for obvious reasons. Quantification and strong distributivity involve distributive links, while reciprocals, dependent plural readings, and weakly distributed plurals involve non-distributive links.

The association of weak distributivity with non-distributing links merits particular attention: it is clear from Williams's discussion that weak distributivity is a relation between two plurals, and therefore should be encoded with a *non*-distributing link. The choice of names here is unfortunate, and potentially confusing: one would expect a link described as *nondistributing* to be incompatible with any kind of distributivity.

Williams suggests the following rules for deriving the truth conditions of sentences containing bound (dependent) pronouns and reciprocals. The rules are presented before the linking theory itself, so they do not refer to linking type.

(5.12) **Rules for reciprocal construal**

a. The antecedent of the reciprocal must be plural and must receive at least a (weakly) distributed interpretation.

b. For any x substituted for the antecedent position, substitute only y's not equal to x but from the same distributed set for the position occupied by the reciprocal.

(5.13) **Rule for bound pronouns**

For any x substituted for the antecedent position, substitute x for the position occupied by the bound pronoun.

Rule (5.13) is intended to include distributive cases; as distribution ranges over the parts of the antecedent, a bound pronoun takes on the same values, as for example in the dependent reading of the sentence

(5.2) John and Mary think they like each other.

Note the mention of "the same distributed set [as x]" in the second rule for the reciprocal. This functions, of course, as the range argument of the reciprocal. In the ("long-distance" reciprocal) cases where the antecedent is a dependent pronoun, the "set" in question is the set of all values over which the pronoun will range, i.e., the plural antecedent of the dependent pronoun. In a compositional semantics, this set is particularly difficult to recover from a bound pronoun. The standard translation of a bound pronoun (according, e.g., to the system of Heim and Kratzer (1998) which I assume) is as a variable. If *they* in sentence (5.2) carries the index i, its translation would be $[\![i]\!]^g$. The interpretation of this is whichever individual is assigned to the variable i by the variable assignment g. Further up in the semantic combinatorics, an expression containing the open variable i has this variable bound by a quantifier that causes it to range over some set.[3] Until that point, the translation of the pronoun is a single (assignment-dependent) individual—and the set of values over which it will range is not retrievable from it!

This is exactly the problem that the scopal analysis of reciprocals succeeds in solving: since the reciprocal cannot be interpreted without reference to the plural antecedent of its dependent pronoun antecedent, the scopal analysis raises it, or binds it, sufficiently high in the structure to provide it with the plural antecedent it needs. To be considered a viable alternative, Williams's proposal must provide a way to interpret the reciprocal without recourse to a non-local antecedent. Unfortunately, as I will argue, the linking theory he sketches can express the desired relationships, but brings us no closer to an understanding of how they might be implemented in the semantics.

[3] In the system of Heim and Kratzer (1998), this is accomplished through lambda abstraction followed by Functional Application.

5.1.2.1 Theta links

Links between NPs and pronouns are called *anaphoric* links (or, *bound pronoun* links). Williams introduces two more varieties of links, each of which may be distributing or non-distributing. *Theta links* connect the theta position of a verb with its argument NP; sentence (5.14a) is an example of distributive theta links (each of the seers saw each of the seen). The third kind of link is a "species of theta-linking" that relates the lower VP to the higher one in structure (5.14b).

(5.14) a. They saw us.
d d

b. They [$_{\text{VP}}$ each [$_{\text{VP}}$ left]].
d

Note that Williams takes "floated" each to be VP-adjoined. Its presence forces this link to be distributing. The higher VP has a plural subject (necessarily, otherwise a distributing link would not be licensed), and the lower VP has a singular subject. This function of linking a VP with a proper subpart of itself is unlike the other two types of links, which relate complete NPs; so it is not always clear from Williams's sketch how it should behave.

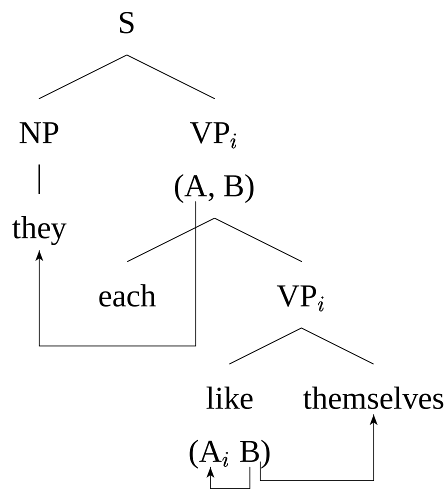
More generally, the plethora of links leads to the prediction of massive ambiguity even for simple sentences, but Williams does not clarify how exactly these readings should be understood. Sentence (5.11) has two theta-links between the pronouns and their verbs, plus the anaphoric link between the two pronouns. Since any of these can be distributing or non-distributing, there are a total of eight configurations; Williams states that "the restriction that a distributive link cannot be headed by singular cuts the number of readings down to four," but does not list the possible or impossible readings. There is a potential source of ambiguity here from whether the thinking is performed individually or collectively, and whether there are many individual sicknesses or just one. (Pragmatic knowledge tells us that thinking and being sick cannot generally be performed jointly—unlike lifting something, for example—but the linking theory ought to represent the possibilities). I confess that I cannot even reproduce Williams's count in this simple example; when I try to list the readings I conclude that there should be six possible readings, not four: John might think that he is sick, that he and Mary are jointly sick, or that they are sick separately. Each of these three readings is compatible with John and Mary doing the thinking either separately or together, giving a total of six readings, all compatible with Williams's linking rules. It is possible that Williams miscounted, or that he had something else in mind. Either way, in reviewing Williams's representations it is more productive to ignore theta-links and limit our attention to anaphoric links.

That said, theta-links are central to Williams's account of the incompatibility of floated *each* with reciprocals and independent readings of reflexives. His explanation makes use of the additional condition that an anaphor must be linked to a theta position of the smallest predicate it belongs to. For example, sentence (5.15a) might mean that John likes John and Mary likes Mary, or that John likes John and Mary and so does Mary; the two readings are apparently rendered as distributing and non-distributing links, respectively, between the subject and object positions of *like.* (The link labels are my own; Williams never illustrates exactly what they should be in such examples). Sentence (5.15b), on the other hand, only has the distributed interpretation: the VP-adjoined *each* causes the embedded VP to be "singular," and the reflexive *themselves* can only link to the singular theta-position of this embedded VP, as shown in (5.16).

(5.15) a. They like(,) themselves.
n n/d

b. They each like themselves.

(5.16)



Being anaphors, reciprocals pattern like reflexives but with the additional requirement that their antecedent must necessarily be plural. Hence sentence (5.17a) has only one reading, while (b) is ungrammatical (the presence of floated *each* means that the only antecedent close enough to the reciprocal, the subject position of the embedded verb, is singular).

(5.17) a. They like each other.

b. * They each like each other.

5.1.3 Dependent reciprocals

We now come to dependent reciprocals, which to a large extent motivated both the scopal analysis and Williams's objections to it. Recall that Williams holds the ambiguity of sentence (5.2) (repeated below) to be entirely due to ambiguity in the interpretation of the pronoun. Since he *apparently* intends to represent the dependent and independent readings of sentence (5.11) via distributing and non-distributing links, respectively, it follows that the dependent reading of (5.2) should be represented as in (5.18a), and the independent reading as in (5.18b).[4]

[4]These structures are simplified representations of the linking relationships that Williams's theory involves. To be fully explicit, the reciprocal should be shown as theta-linked to an argument position of the verb, not directly to an antecedent.

(5.2) John and Mary think they like each other.

(5.11) They think they are smart.

d/n

(5.18) a. John and Mary think they like each other. (dependent)

d n

b. John and Mary think they like each other. (independent)

n n

The reciprocal is never linked to anything outside its local binding domain. In this way Williams's analysis is "non-scopal," since he avoids the scope ambiguity that has been a mainstay of other analyses of this construction.

Although the intent of these representations is clear, already we come upon another problem with Williams's proposal: he intends distributing links to create a singular variable at their lower end, but reciprocals need a plural antecedent; so structure (5.18a) should be illegal. Unfortunately Williams, after illustrating strict and weak distributive links, leaves the link labels out of most of his diagrams; so it is possible that I have misunderstood something. But it is clear that Williams intended to account for the ambiguity in (5.11) by assigning different linking structures to its two readings, which implies that they are distinguished via distributing vs. non-distributing links; so he must have meant the dependent pronoun in (5.18a) to be linked distributively to its antecedent, as shown.

No matter what Williams intended, it seems to make sense for dependent pronouns, which are bound by distributed NPs (not quantifiers) and are compatible with reciprocals, to have a weakly rather than strictly distributive relationship with their antecedent; weakly distributive relations are represented by *non*-distributing links, and both of the NPs they link are "plural," so reciprocal reference is allowed. This interpretation would assign linking structure (5.18b) to *both* readings of 5.2, something Williams probably did not intend. What is really needed, perhaps, is a third link type explicitly for weakly distributive links; in its absence, this may be the best this system can do. Rather than try to further pursue repairs to Williams's system, I will drop the question of link labels and continue to use Williams's notation as an expressive way of representing relationships that have no simple equivalent in terms of indices.

5.1.4 Linking multiple reciprocals

In sections 2.1.7.3 and 2.3.3, I discussed the readings of the following sentences containing multiple reciprocals, and the problems they presented for the Heim et al. (1991a,b) account. This section reviews the linking structures assigned to these readings by Williams's proposal.

(5.19) a. John and Mary read each other's books in each other's languages. (= (2.27))

b. John and Mary told each other that they love each other.

c. They gave each other pictures of each other.

Sentence (5.19c) has the two readings given in (5.20). Reading (a) is represented by linking both reciprocals to the matrix subject, an uncontroversial interpretation considering that the subject is syntactically able to bind either reciprocal by itself (as shown by the grammaticality of examples (5.21a,b)).

(5.20) a. John gave Mary a picture of Mary, and vice versa.

b. John gave Mary a picture of John, and vice versa.

(5.21) a. John and Mary gave each other pictures of Al Gore.

b. John and Mary gave me pictures of each other.

Reading (5.20b) appears to require the higher reciprocal to function as the antecedent of the second, a configuration that Williams's system is designed to allow:

(5.22) They gave each other pictures of each other.

n n

This type of construction cannot be handled correctly by the Heim et al. account, as discussed in section 2.3.3.

Example (5.19b), repeated below, is slightly more involved: According to the Heim et al. analysis, it has a reading that involves two reciprocals with the same antecedent (the matrix subject), the second one being a "long-distance" reciprocal. (That reading says that John told Mary that he loves her). For Williams, this reading is given the structure shown in (5.23), involving the same types of links as in example (5.18a): the lower reciprocal is linked to the embedded pronoun, not to the higher subject, and the pronoun is linked to the matrix subject. These links do not interact with a third link, between the higher reciprocal and the matrix subject.

(5.19b) John and Mary told each other that they love each other.

(5.23) John and Mary told each other that they love each other.

A second reading of sentence (5.19b), parallel to reading (5.20b) above, says that John told Mary that Mary loves him; it requires the higher reciprocal to be used as the antecedent of the embedded pronoun, which in turn is the antecedent of the lower reciprocal:

(5.24) John and Mary told each other that they love each other.

n ? n

The key element in the above representations is the option of letting the antecedent of a reciprocal be a dependent pronoun or reciprocal. Leaving aside the specifics of Williams's linking representations, his intent is that dependent pronouns and reciprocals bear a "weakly distributive" relationship to their antecedents, and that both ends of a weakly distributive link behave as plurals.

But a picture is not a semantic analysis. Recall that the difficulty in assigning a semantic translation to sentences containing dependent reciprocals lies in finding an antecedent that the reciprocal can range over. Although Williams *states* that the embedded pronoun in a sentence like (5.2) is plural, he brings us no closer to an understanding of how its plurality can be expressed in a translation that also captures its dependence on distribution over its antecedent.[5] My own conclusion is that Williams, unencumbered by the expressive limitations of traditional semantic notation, has provided an insightful description of the anaphoric relationships between reciprocals and dependent pronouns; but the linking system he proposes does not help answer the question of how to endow these relationships with a formal semantic translation. The merits of his account are largely complementary to those of the Heim et al. account, which provides a workable semantics of considerable empirical coverage at some cost to the accuracy of the relationships it models.

In the next section, I review the general conclusions we can draw, largely on the basis of Williams's analysis.

5.1.5 Some generalizations

Although I have concluded that the non-traditional linking system that Williams has proposed makes no inroads on the central puzzle of deriving the reciprocal's range, I have found some of the generalizations and distinctions he made to have lasting value. I retain two elements of his approach in particular: (a) the finding that quantification-like expressions divide into two systematically differing groups, one including "real" quantifiers and

[5] The sense in which dependent pronouns are "plural" can be seen more clearly in the treatment of Kamp and Reyle (1993). In their system, dependent pronouns are semantically neutral for number (i.e., can refer to either atomic or non-atomic individuals) and are translated as bound variables; but the discourse referents they introduce are annotated with a diacritic that licenses their use as antecedents of other suitable, morphologically plural pronouns.

Although this system allows dependent pronouns to have non-atomic referents, this is only because their binder (typically a distributor) may range over non-atomic parts of its restrictor. Dependent pronouns are true bound variables, whose values are entirely determined by their binder. Kamp and Reyle's treatment of plurals is presented in more detail in section 3.2.

the other including nominal plurals and other constructions; and (b), that reciprocals should be interpreted primarily, if not solely, with respect to their local antecedent.

Let us consider the class of constituents that may serve as the local antecedent of reciprocals. Definite NPs belong to this class, of course, as do other reciprocals and pronouns dependent on definite NPs; pronouns bound by explicitly quantificational NPs do not, and neither do any antecedents separated from the reciprocal by an intervening "floated" *each.* In Williams's terms, we might say that there is something about reciprocals and dependent pronouns that makes them sufficiently "plural" to serve as the antecedents of reciprocals. Pronouns bound by an overt quantifier, on the other hand, are not "plural" in the relevant sense.

What reciprocals and dependent pronouns have in common is that they are bound not by an overt quantifier but by an implicit distributor. Implicit distributors also pattern unlike overt quantifiers with regard to quantifier raising in pair-list questions, as shown by Krifka (1992) and, independently, by Dayal (Srivastav 1992, Dayal 1996). (See also section 2.2.2.2). It follows that distributors and overt quantifiers should receive different semantic representations, in accordance with each type's characteristics. Overt quantifiers (which do not concern us except by omission) may be represented in the usual way as NP-adjoined Generalized Quantifiers; we may assume that they are assigned scope through quantifier raising. Distributivity, on the other hand, should be represented as a separate, irreducible property along the lines pursued by Sternefeld (1998), Schwarzschild (1992, 1996), Brisson (1998) and others.

Returning now to the issue of dependent reciprocal sentences, Williams makes the strong claim that the interpretation of the reciprocal in the dependent reading of example (5.2) (repeated below) should refer only to its local antecedent, the embedded pronoun *they.* The scopal account of Heim et al. makes the equally strong claim that the semantics of the dependent reciprocal should be derived entirely via its association with the matrix subject, with the dependent local "antecedent" serving only to restrict the possible antecedents that the reciprocal may be bound by. As we saw in chapter 2, the claim is demonstrably too strong, being unable to handle sentences containing chained multiple reciprocals like (5.25), or sentences with dependent pronouns that are not construed as identical with their antecedents, such as (5.1b).

Subsequent scopal accounts were likewise unable to account for sentences like (5.1b). We now turn to two such accounts, those of Sternefeld (1998) and Schwarzschild (1996).

(5.2) John and Mary think they like each other.

(5.25) They want each other to like each other.

(5.1b) The lawyers that represent them$_i$ say they$_i$ will sue each other.

5.2 Cumulation and weak reciprocity: Sternefeld (1998)

Sternefeld's (1998) very elegant system addresses two inadequacies of the representation of distribution as universal quantification.[6] The first is *part* of what Williams (1991) referred to as *weak distributivity:* If John and Mary lifted the couch together and Bill lifted it by himself, sentence (5.26) can be taken as true (provided the context makes its use plausible). But it does not follow that John lifted the couch, or that Mary lifted the couch. (The other half of Williams's concern, the possibility that some members of a plural subject may not actually have participated in the action described by the predicate, is not addressed by Sternefeld). The desired truth conditions for this sentence can be derived if we give it the semantics that I defined in section 1.4 as *weak distributivity.* This is shown in (5.27). (In this section, I use the symbol $\in$ to denote the relation *atomic-part-of*).

(5.26) John, Mary and Bill lifted the couch.

[6] See section 1.4 for a brief review of the relevant issues.

(5.27) $(\forall x \in A)(\exists y \amalg A)\, x \amalg y \;\&\; P(y)$ (= (1.10))

This formula says that every atomic part x of the subject is *part* of some individual y, possibly plural, to which the predicate applies. (Note that this definition does not exclude the possibility that $x = y$).

The second issue involves what I have called the *cumulative* reading (which Sternefeld, inconsistently with my choice of terms and somewhat confusingly, describes as "weak distributivity"). Ignoring for the moment the possibility of group action that we just addressed, sentence (5.28) can be truthfully used even if it is not true that each boy rented all three movies. All that is required under this reading is that each boy rented one or more of the movies, and each movie was rented by at least one boy. The truth conditions of the cumulative reading are given by expression (5.29).

(5.28) The boys rented *Babe, Citizen Kane,* and *Animal House.*

(5.29) $((\forall x \in B)(\exists y \in M)\, xRy) \;\&\; ((\forall w \in M)(\exists z \in B)\, zRw)$

Sternefeld develops an elegant algebraic treatment of distributivity, cumulativity and reciprocity that is based on Link's (1983) plural operator, "*", and Krifka's two-place equivalent (defined in his Ph.D. dissertation), "**". We have seen the plural operator in section 1.3: *P is a predicate whose extension is the closure of the extension of P under the sum operation, in other words, *P is true of all sums of individuals that are in the extension of P. The plural operator can be defined as in (5.30). The two-place "cumulation" operator ** is defined as in (5.31):

(5.30) For any set (i.e., predicate) P, *P is the smallest set such that

a. $P \subseteq {*P}$, and

b. if $a \in {*P}$ and $b \in {*P}$, then $a \oplus b \in {*P}$.

(5.31) For any two-place relation R, let $^{**}R$ be the smallest relation such that

a. $R \subseteq {}^{**}R$, and

b. if $<a, b> \in {}^{**}R$ and $<c, d> \in {}^{**}R$, then $<a \oplus c,\ b \oplus d> \in {}^{**}R$.

(Sternefeld 1998:304; definition due to Krifka)

Recall that a predicate P need not be defined on singular individuals only; hence if b is a plural individual, $P(b)$ does not guarantee that P is true of the parts of b. The plural operator allows such predicates to be interpreted distributively, but with the semantics of *weak distributivity:* distribution "down to individuals" is not required, rather, $^{*}P(b)$ only requires that each atomic part of b must be *part* of an individual, not necessarily atomic, to which P applies. Even collective action satisfies this requirement.

The cumulation operator expresses cumulativity as membership in the "cumulation" of a relation. It can be used to derive the cumulative reading even in cases involving collective action, which formula (5.29) does not cover. For example, consider sentence (5.32) in a situation where some guests might have drunk several bottles by themselves, others might have shared one or more bottles, etc. All that is said is that a total of one hundred guests drank a total of two hundred bottles of wine.

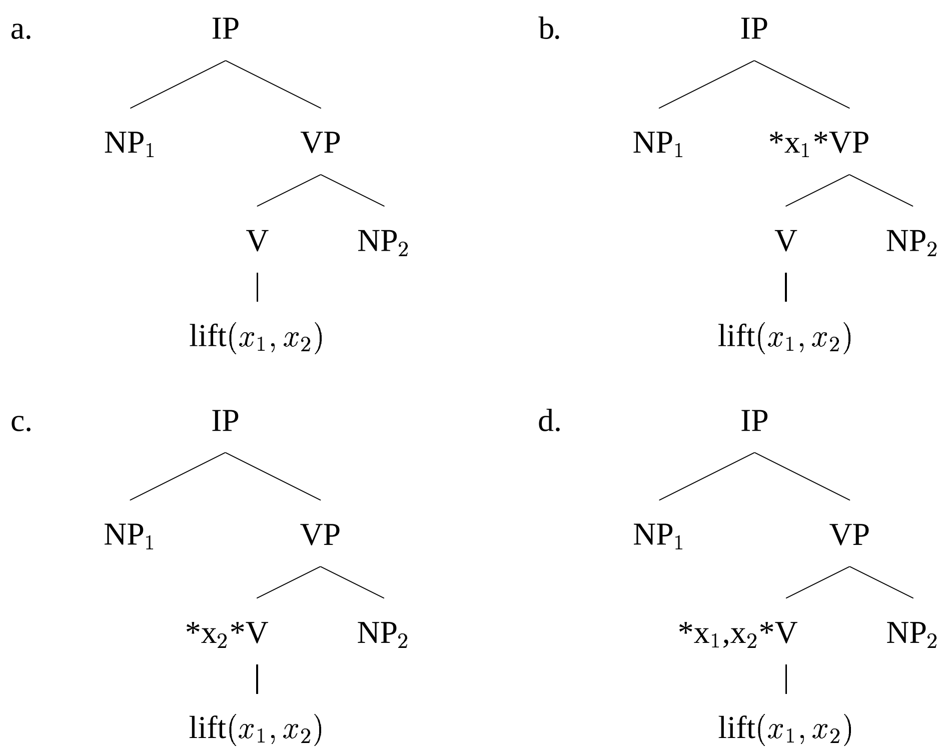
(5.32) One hundred guests drank two hundred bottles of wine.

The original cumulative formula (5.29) can be read as stating that A is the sum of *atomic* parts that appear on the left of the relation R, and B is the sum of atomic parts that appear on its right. The reader may verify that Sternefeld's definition of cumulation (5.31) expresses the same condition, but with the requirement of atomicity removed: it is validated if A is the sum of parts, *not necessarily atomic,* that appear on the left of R, and B is the sum of parts, not necessarily atomic, that appear on its right.

Sternefeld allows the plural and cumulation operators to be inserted freely as "semantic glue” wherever appropriate. Depending on what is inserted where, numerous readings can be derived for sentence (5.33). By way of illustrating Sternefeld’s system, consider the trees given in (5.34), and the corresponding translations and glosses in (5.35). Reading (a), which involves neither distribution nor cumulation, says that five men, acting together, lifted two pianos at once. Reading (b) allows distribution over the subject, while reading (c) expresses distribution over the object. Finally, reading (d) involves cumulation over both positions; it says that a total of five men were involved in lifting a total of two pianos, but no further claims about joint or individual action are made.

(5.33) Five men lifted two pianos.

(5.34)



(5.35) a. $\exists X$ five(X) & *man(X) & $(\exists Y$ two(Y) & *piano(Y) & lift$(X, Y))$

Five men jointly lift two pianos at once.

b. $\exists X$ five(X) & *man(X) & $X \in *\lambda x\,[\exists Y$ two(Y) & *piano(Y) & lift$(x, Y)]$

A total of five men (perhaps in subgroups) lifted two pianos at once.

c. $\exists X$ five(X) & *man(X) & $(\exists Y$ two(Y) & *piano(Y) & $Y \in {}^*\lambda y\,[\text{lift}(X, y)])$

Five men jointly lift a total of two pianos, possibly one at a time.

d. $\exists X$ five(X) & *man(X) & $(\exists Y$ two(Y) & *piano(Y) &

$<X, Y> \in {}^{**}\lambda xy\,[\text{lift}(x, y)])$

A total of five men were involved in acts of lifting a total of two pianos, possibly one at a time.

The two operators can bind a variable, allowing the dependent reading of sentence (5.36a) to be expressed as in (b).[7]

(5.36) a. They love their parents.

b. $Y \in {}^*\lambda x\,[\![x \text{ loves } x\text{'s parents}]\!]$

5.2.1 Reciprocals and weak reciprocity

In a parallel way, as discussed in section 1.4, the semantics of reciprocals are weaker than implied by the translation used by Heim et al. (1991a), which employs universal quantification over the atomic parts of the antecedent. A useful alternative is given by *weak reciprocity,* which Langendoen (1978) defines as follows:

(5.37) $(\forall x \in A)(\exists y, z \in A)\, x \neq y \;\&\; x \neq z \;\&\; xRy \;\&\; zRx$

Like the original version of cumulativity, weak reciprocity is stated in terms of quantification over atomic individuals. Using the cumulation operator, Sternefeld (1998) expresses the semantics of weak reciprocity as the cumulative reading of a derived expression; again, his system removes the requirement that the relation should distribute down to atomic in-

[7] Sternefeld implements variable binding via an adaptation of the system of double indexing (inner and outer indices) defined by Heim (1993).

dividuals. This is accomplished by translating a reciprocal relation A R *each other* as the following expression.

(5.38) $<A, A> \in **\{<x, y> \mid <x, y> \in R \;\&\; x \neq y\}$

Unpacked, (5.38) says that A must be the sum of parts x, not necessarily atomic, such that for each such x there is some part $y \neq x$ for which $x\ R\ y$ holds; and also that A must be the sum of parts y with the corresponding property. Thus if R is the relation of writing a book about someone, as in example (5.39), every atomic part of A collaborated in writing a book (or wrote a book alone), and also was the partial or sole subject of a book.

(5.39) They wrote books about each other.

Using condition (5.38), Sternefeld is able to give reciprocal sentences the more accurate semantics of weak reciprocity. (Recall that the account of Heim et al. relies on universal quantification and expresses strong reciprocity). In a dependent reciprocal sentence the entire reciprocal raises to the matrix clause; the dependent reading of example (5.40a) is given as in (b). For comparison, the independent reading (under which John and Mary each think "we like each other") is given in (c).

(5.40) a. John and Mary think they like each other.

b. $(\exists X)\,(X = \{j, m\} \;\&\; <X, X> \in **\lambda xy\,[x \neq y \;\&\; \text{think}(x, {}^\wedge[\text{like}(x, y)])])$

c. $(\exists X)\,(X = \{j, m\} \;\&$

$X \in *\lambda x\,(\text{think}(x, {}^\wedge[<X, X> \in **\lambda xy\,[x \neq y \;\&\; \text{like}(x, y)]])))$

This analysis depends on the reciprocal's ability to raise to the matrix clause, so that the reciprocal condition can apply at that level. This is the essence of the scopal analysis, and as a consequence, Sternefeld's system fails to give the correct semantics for the dependent reading of examples (5.1), in just the way the Heim et al. system fails. Sentence (5.1b), for example, is translated as follows ($W(x)$ is the translation of the paycheck pronoun *they*, and expresses a function taking lawyers to clients):

(5.1b) The lawyers that represent them$_i$ say they$_i$ will sue each other.

(5.41) $(\exists X)(X = \text{lawyers}_1 \;\&\; <X, X> \in {}^{**}\lambda x\, \lambda y\, [x \neq y \;\&\; \text{say}(x, {}^\wedge[\text{sue}(W(x), y)])])$

This is the same meaning given by the Heim et al. (1991a) analysis: it claims that each lawyer x said that his or her client, $W(x)$, will sue one or more of the other *lawyers*, y. The correct semantics would require replacing the last occurrence of y with $W(y)$, but that is something that Sternefeld's system has no way of doing. Again, the essence of the scopal analysis is that the reciprocal ignores its local antecedent as far as its semantic translation is concerned.

5.3 Covers for distribution: Schwarzschild (1996)

5.3.1 Plurals and covers

Schwarzschild's (1992, 1996) treatment of plurals allows the pragmatics to play a role in how we construe a distributively interpreted sentence. The distributive operator he proposes is a refinement of the VP-adjoined distributor shown in (5.42):

(5.42) $\lambda P\, \lambda N\, \forall y(\text{atomic}(y) \;\&\; y\; \Pi\; N)\, P(y)$

This operator essentially has the semantics of the distributor employed by Heim et al. (1991a), but is adjoined to VP rather than to the subject.[8]

[8]The distributor in (5.42) also differs from the Heim et al. distributor in that it does not require that y be a *proper* part of N.

Schwarzschild is concerned with allowing the distributive operator to describe a situation intermediate between collective and individual action, in which some parts of the subject satisfy a predicate in combination with others, but not individually. Such intermediate groupings are sometimes contextually forced by the context. E.g., consider a context where we want to weigh numerous baskets filled with vegetables, but there is only a fine and a bulk scale. Suppose that individual vegetables can be weighed on the fine scale, and all the baskets together can be weighed on the bulk one, but that neither scale is suitable for weighing the baskets of vegetables. In this case the following sentence could be truthfully asserted:

(5.43) The vegetables are too heavy for the fine scale and too light for the bulk scale.

In order to get a translation that is true under these conditions, we can let the distributor range not over the atomic parts of its subject, but over the elements of a partition of the vegetables into basketfuls.

A partition of a set X is a collection of non-empty, *disjoint* subsets of X such that the union of all sets in the partition is equal to X. Schwarzschild argues that in general, we must allow for a distribution operator that allows overlap between the parts. His argument is based on the oft-discussed example of Rodgers, Hammerstein and Hart, who wrote several musicals in pair-wise collaborations, but none individually or all together.[9] In this case, the following sentence can be truthfully asserted:

(5.44) The men wrote musicals.

On the force of such examples, Schwarzschild adopts *covers* rather than simple partitions: a cover of X is a collection of non-empty subsets of X whose union is equal to X. The dis-

[9] The example is due to Gillon (1987), who proposed an analysis involving covers. (See also the response by Lasersohn (1989)).

tributor he proposes, written *Part,* is a VP-adjoined operator that ranges over the elements of a contextually supplied cover *Cov.* Its meaning can be given as follows:

(5.45) A (plural) individual x is in the extension of *Part(Cov)*(P) iff $[\![\mathrm{Cov}]\!]$ is a cover and $\forall y(y \in \text{Cov} \;\&\; y \;\Pi\; x)\; P(y)$.

For example, the plural individual representing the sum of Rodgers, Hammerstein and Hart is in the extension of the predicate *Part(Cov)(wrote-musicals),* provided *Cov* denotes a cover that includes the pairs that actually wrote a musical (and none of the three as an atomic individual, nor all three of them).

In this definition, *Cov* is a variable over covers of the entire domain of individuals in our model, not just of x. This makes it possible, for example, to treat VP elision using the same cover for the elided VP as for its antecedent. Consider a situation where a number of computers were bought one at a time, and a bunch of diskettes were bought all together. Each purchase was paid for in two installments. It is then possible to say:

(5.46) The computers were paid for in two installments and the diskettes were too.

We can represent the desired situation by choosing a single, universal cover that individuates the computers, but lumps the diskettes together.

The mechanism of covers can express the collective reading of a predicate simply by choosing a cover that puts its entire subject in a single cell. For example, a cover that puts all the salient bottles in a single cell would account for the collective reading of (5.47), while one that puts each bottle in a separate cell would account for its distributive reading.

(5.47) The bottles are light enough to carry.

Since the collective reading is thus expressible by the *Part* operator, Schwarzschild assumes that *all* plural VPs include *Part* in their translation; the collective construal is only distinguished from the distributive by the context.

Different VPs in a complex sentence may share a cover or may have different covers, as necessary. The index of the *Part* operator can bind an open variable in its VP complement; in this way it is possible to treat dependent pronouns as variables bound by the cover operator. The dependent reading of sentence (5.48a) is analyzed as in (b).

(5.48) a. The men dropped their babies on their beds.

b. (the-men') [Part_1(Cov) (drop-on(x_1's baby's bed)(x_1's baby))]

Sentences involving two non-dependent definite NPs, on the other hand, are interpreted by means of ordered-pair covers and the *paired-cover* operator *PPart.* This mechanism is Schwarzschild's method of capturing cumulative readings, as well as the effect that Sauerland (1995a) calls *codistribution.*[10] This involves the pairing between corresponding parts of two definite NPs, as in the following examples:

(5.49) a. The men dropped the babies on their beds.

b. The sides of rectangle R_1 run parallel to the sides of R_2.

c. Even though the couples in our study were not married, the men did display aggressive behavior toward the women.

We are interested in the reading of (5.49a) that involves each man dropping his baby or babies on their beds. Similarly, example (c) is easily understood as referring to aggressive behavior between the members of each couple. Example (b) is well-known, and is due to Scha (1984). Schwarzschild demonstrates that in order for it to be judged as true, the sides of the two rectangles must be paired up in a systematic way similar to the pairing of men and babies.

[10] Sauerland's treatment of codistribution was described in section 3.3.2.

The paired-cover distributor *PPart* applies to two-place predicates. It takes a "paired-cover" consisting of sets of ordered pairs of individuals, defined as follows:

(5.50) T is a **paired-cover** of $<A, B>$ iff:

there is a cover of A, $C(A)$, and there is a cover of B, $C(B)$, such that:

i. T is a subset of $C(A) \times C(B)$.

ii. $\forall x \in C(A)\, \exists y \in C(B) : <x, y> \in T$

iii. $\forall y \in C(B)\, \exists x \in C(A) : <x, y> \in T$

If T is a paired-cover of $<A, A>$, then T is a paired-cover of A.

Note that a paired-cover is not a cover over the set of ordered pairs $A \times B$; if it was, every ordered pair would need to be represented in the cover. The definition of paired-covers only requires that each element of A appear on the left side of *some* ordered pair that is in the cover, and similarly for B. The paired-cover operator can then be defined as in (5.51a), where $\amalg_2$ is the paired-subset (i.e., paired-part) relation defined in (5.51b).

(5.51) a. An ordered pair of individuals X is in the extension of *PPart*(*PCov*)(α) iff

$\forall Y (Y \in \text{PCov}\ \&\ Y \amalg_2 X)\, P(Y)$.

b. $<a, b> \amalg_2 <c, d>$ iff $a \amalg c\ \&\ b \amalg d$

Using this operator, the desired meaning of sentence (5.49a) is generated by providing a cover that places each man in an ordered pair with his baby or babies.

(5.49a) The men dropped the babies on their beds.

a. (the-men')(the-babies')[PPart$_{2,3}$ (PCov)(drop-on'(x_2's bed))]

b. $\forall x \forall y[<x, y> \in f(\text{PCov}_1)$ and x is a man and y is/are his babies] $\rightarrow$

[x dropped y on y's bed]

Similarly, example (5.49c) is generated by providing a cover that places the members of each couple in an ordered pair of their own.

5.3.2 Schwarzschild's semantics for reciprocals

As with distribution, Schwarzschild allows a large role for context in the interpretation of the reciprocal relationship. Once again, he replaces reciprocation among atomic subparts of the subject with reciprocation among the elements of a contextually supplied cover. For example, sentence (5.52a) under the reading where the cows were separated from the pigs induces a two-element cover of the set of animals, consisting of the set of cows and the set of pigs. Sentences (b) and (c) induce a cover that groups the animals by age, regardless of species. Depending on suitable context, other covers might group each animal individually, or group the animals by color, by owner, etc. The resulting representation is sketched in (d).

(5.52) a. The cows and the pigs were separated (from each other).

b. The animals were separated from each other by age.

c. The cows and the pigs were separated from each other by age.

d. (The cows and the pigs) Part(Cov$_1$) (be separated from e.o.)

The truth conditions of a reciprocal predicate are defined in terms of a context-provided relation *Recip,* a two-place predicate that is required to have the properties of some type of reciprocal relation (strong, weak, or some special pairing) that is appropriate to the context of evaluation. The *Recip* relation induces a function *EachOther,* which takes each subplurality x to the sum of all subpluralities that x is in the *Recip* relation with:

(5.53) $[\![\text{EachOther}]\!](x) = \bigcup\{s : <x, s> \in \text{Recip}\}$

Schwarzschild defines the meaning of the reciprocal directly by means of this function, augmented with an initial argument that gives the range argument from which the s elements are drawn. It is represented in a manner based on Cooper's (1979) treatment of donkey pronouns: It is a free variable that combines with its antecedent NP to give a contraindexing function, whose domain and range are the cover itself. The argument of the contraindexing function is bound by the quantifier in the *Part* operator, so that it ranges over the elements of the cover. Specifically, Schwarzschild (1996:120,125,134) gives the following rules.

(5.54) a. for any indices i,j:

each other$_{i,j}$ translates as *EachOther*$(x_i)(x_j)$.

EachOther is a variable of type <e, ee>.

b. for all M,g:

1. $\forall u \forall v$ $[\![\text{EachOther}]\!]^{M,g}(u)(v) \subset u$.
2. $\forall u \forall v$ $[\![\text{EachOther}]\!]^{M,g}(u)(v) \neq v$.
3. $\forall u \forall v$ $[\![\text{EachOther}]\!]^{M,g}(u)(v)$ is undefined if $v \not\subseteq u$.

c. The domain and range of $g(\textit{EachOther})(g(x_j))$ is identical to the value assigned to *Cov* in the *Part* operator that binds the reciprocal.

The conditions under (5.54b) establish that the first argument of *EachOther* is the range argument, and the second is the contrast argument.

A sentence like (5.55a) is then translated an in (b), (c).

(5.55) a. The boys saw each other.

b. The boys$_2$ Part$_1$[saw EachOther(x$_2$)(x$_1$)].

c. $\forall x_1 (x_1$ Π boys$_2$ & $x_1 \in$ Cov) saw(x_1)(EachOther(boys$_2$)(x_1))

This mechanism intentionally leaves a lot of things to the context. In particular, the first argument of the reciprocal (its range argument) is not sufficient to determine the legal values of the reciprocal function; condition (5.54c) requires the legal values to be only those parts of the range that are in the cover *Cov.* The reciprocal function itself can be chosen so as to match each member of the cover with a set of mates according to a contextually appropriate reciprocal relation.

Note that Schwarzschild's translation of reciprocals will assign to each boy b the predicate *see(Y),* where *Y* = *EachOther*$(R)(b)$ is the set of all boys that b saw. Since a boy may have seen several other boys one at a time, this involves a sort of distribution over the object that the reciprocal translation does not address. As I will discuss in section 5.3.4, it is possible to extend Schwarzschild's semantics by adopting Lasersohn's (1998) system of *generalized distributivity,* which allows distribution over the object. Another desirable refinement concerns the distinctness condition given as (5.54b-2). Note that this is given as simple non-equality; as I showed in section 4.8, some type of stronger condition is appropriate to rule out pairing of certain non-disjoint individuals. The distinctness condition I adopted there was as follows:

(5.56) The contrast argument c is distinct from a derived argument z if and only if $c \not{\mathrm{I\!I}}\ z$.

(= (4.48))

5.3.3 Dependent reciprocals

Schwarzschild's analysis can handle simple dependent pronouns by simply letting the reciprocal be bound by the matrix (and presumably only) *Part* operator. This is equivalent to the long distance analysis of Heim et al. (1991b), and suffers from the same shortcomings.

(5.57) They think they like each other.

They$_2$ Part$_1$ [think they$_1$ like EachOther$(x_2)(x_1)$].

It is easy to see that this analysis does not extend to more complicated examples involving dependent pronouns. Schwarzschild (1996:132ff) discusses the indexing constraints in the following example:

(5.58) a. The shoppers parked the cars near each other.

b. <shoppers$_4$, cars$_1$> PPart$_{2,3}$ [parked near EachOther$(x_1)(x_3)$]

The translation given in (b) binds the reciprocal arguments to the second member of the subject and cover operator pairs, with the result that each car is parked near other cars. There is also a reading where each shopper parked some cars near some other shopper or shoppers, which corresponds to the first member of the ordered pairs (indices 2 and 4). The remaining possible indexings are impossible: *EachOther*$(x_1)(x_2)$, for example, would be a function from shoppers to cars. Condition (5.54b-1) was adopted precisely to rule out this type of indexing. But since *EachOther* is constrained to range over the values in the cover, the problem returns if we replace *the cars* above with the dependent *their cars*. Now the only reading that the theory can generate is the one where each shopper parked his car(s) near some other shopper; since cars are no longer included in the cover, there is no way to generate the reading that says each car was parked near other cars.[11]

The same problem arises with any of the other constructions that were presented as problematic for the long distance analysis of the reciprocal. For example, consider the following sentence under its dependent reading:

(5.59) a. The lawyers that represent John and Mary advised them to sue each other.

[11] Roger Schwarzschild (personal communication) points out that the right analysis for examples (5.58a) and (5.59a) can be generated by using paired-covers. This would be tantamount to adopting the cumulative analysis of such sentences, and treating the NP *their cars* as a referential plural (interpreted cumulatively) rather than as an expression containing a bound variable. As I argued in detail in section 3.3, I do not find the cumulative analysis tenable in general. I believe that the arguments presented there should be sufficient to rule out the cumulative analysis here as well. (Schwarzschild (1996:p. 114f) argues against applying the cumulative analysis to a number other of examples of NPs containing dependent pronouns).

b. Lawyers$_2$ Part$_1$ [advised client-of(x_1) to sue EachOther(x_2)(x_1)]

In the translation in (b), I have translated the pronoun *them* as a paycheck function mapping lawyers to their clients, whose argument is bound by the *Part* operator. Since the range argument of *EachOther* is x_2, the reciprocal ranges over lawyers, not clients. This is the same problem we encountered with the Heim et al. (1991a,b) account.

5.3.4 The benefits of the covers approach

Thanks to its flexibility and reliance on pragmatics, Schwarzschild's system provides a framework that can address several issues that arise in a close look at the interpretation of distributivity and reciprocity. Most directly, as we saw, it makes it possible for the distributor to range over parts of the subject that are not necessarily atomic. This lends substance to the observation that distributive entailments are not necessarily, as Williams (1991) puts it, "down to individuals." In addition, the system of paired-covers allows the context to control the way that a cumulative or reciprocal relation is realized.

Schwarzschild's system can be easily extended in various ways in order to solve problems not addressed by Schwarzschild himself. One such issue is that of exceptions, or "nonmaximality" as Brisson (1998) calls it. Williams (1991) notes that the following sentence would be accepted as true even if there were some non-hitters:

(5.60) They were hitting Bill.

If this effect is to be represented in the semantics,[12] we need some way to allow for a

[12]Kroch (1979) argues that is a Gricean effect, since he finds that sentences like the following do not allow exceptions:

(i) Although the men in this room are angry, there are some who aren't.

There is some disagreement about whether distributive sentences allow exceptions; for example, Brisson (in preparation) gives (ii) as an example of a well-formed exception construction:

predicate to be truthfully applied to an individual even if it does not apply to parts of it. One promising approach is pursued by Brisson (1998), who makes use of so-called "ill-fitting" covers. Recall that in Schwarzschild's system, the cover variable is defined over the entire universe of individuals, not just over the subject of each predicate. Consider a situation where the NP *the girls* refers to Faith, Hope and Charity, and consider the truth conditions of sentence (5.61a) under the cover given in (b).

(5.61) a. The girls jumped in the lake.

b. Cov = $\{\{$Faith$\}, \{$Hope$\}, \{$Charity,Tom$\}\}$

c. $\forall y(y \in$ Cov & y Π girls$')$ jumped-in-the-lake(y).

According to the definition of the *Part* operator given in (5.45), sentence (5.61a) is translated as in (c). Note that Charity only appears in the set that also contains Tom, who is not in the denotation of *girls*$'$. This means that the condition y Π *girls*$'$ will exclude this set, and therefore the truth of (5.62c) does not depend on whether or not Charity jumped in the lake. Schwarzschild considers such covers "pathological" and assumes that the pragmatics would rule them out as a matter of principle. Brisson, on the other hand, argues that they give us a way to formalize nonmaximality; cover (5.62b) would be appropriate in a situation (probably involving more than three girls) where most of them jumped in the lake but Charity did not.[13] Brisson's approach to exceptions has been extended to reciprocals and

(ii) The girls jumped in the lake, but not Carrie.

On the other hand, collective predicates are known to allow exceptions more easily. Kroch (1979) gives the following example (his explanation makes use of the distinction, non-existent in Schwarzschild's system, between group individuals and sets of individuals).

(iii) Although the soldiers surrounded the town, not all of them participated.

[13]This is not the only possible way to account for nonmaximality. Sauerland (1995a, 1998) uses the operator ENOUGH (based on the homonymous operator defined by Roberts (1989) as suggested by Emmon Bach), which allows a predicate to be considered true of a subject if it is true of a "substantial part" of it:

(i) $\text{ENOUGH}(P)(y) = 1$ iff there is an x such that x is a substantial part of y and $P(x) = 1$.

relational plurals by Beck (2000).

Another interesting extension is the *generalized distributivity* operator proposed by Lasersohn (1998), which allows distribution over the object only. Recall that Schwarzschild's translation for the reciprocal matches each salient part of the subject with the sum of all parts that it bears the predicated relation to, (e.g., each boy with the sum of all boys that he kicked), but does not further analyze this relation into parts. Distribution over the object allows us to treat that part of the predicate on a par with the subject, including the treatment of exceptions via ill-fitting covers.

Generalized distributivity is built on the *generalized conjunction* operator defined in (5.62). A *conjoinable type* is any type that yields <t> when all its arguments are saturated (i.e., any type ending in *t*). Generalized conjunction over a set X of expressions of some conjoinable type gives an expression of the same type, which is interpreted by passing each argument to all elements of the set X and taking the logical conjunction of the eventual truth values. (With respect to clause (a), note that a set of saturated expressions which all evaluate to 1 is exactly the set { 1 }).

(5.62) a. If $X \subseteq D_t$, then $\sqcap X = 1$ if $X = \{1\}$; $\sqcap X = 0$ otherwise.

b. If $X \subseteq D_{<\mathrm{a,b}>}$ (where <a, b> is a conjoinable type), then $\sqcap X$ is that function $f \in D_{<\mathrm{a,b}>}$ such that for all a, $f(a) = \sqcap\{f'(a) \mid f'(x) \in X\}$.

Lasersohn then defines generalized distributivity over all "distributable types" (of the form <e, κ>, for κ a conjoinable type), and extends Schwarzschild's *Part* operator so that it accepts such types as its second argument:

A very different approach is proposed by Lasersohn (1999), whose system of *pragmatic halos* extends the extension of predicates with a collection of objects that should be considered, for practical purposes and relative to the context at hand, "close enough" to being true members of some predicate's extension. This allows us to speak loosely when the situation allows, and regulates *how* loosely we may speak.

(5.63) Where α is an expression of some distributable type <e, a>, x is any individual, and $[\![\text{Cov}]\!]^{M,g}$ is a cover of the universe of discourse:

$[\![\text{Part(Cov)}(\alpha)]\!]^{M,g}(x) = \sqcap\{[\![\alpha]\!]^{M,g}(y) \mid y \in [\![\text{Cov}]\!]^{M,g} \;\&\; y \leq_i x\}$

In this way, Schwarzschild's approach can be extended to address some of the complications attendant to distributive readings. As noted, the *Recip* variable in the translation of reciprocal predicates allows the context to supply the proper type of reciprocal relationship. In addition, I show in section 5.4.1 that his translation of the reciprocal itself as a function turns out to be useful in providing an analysis for the "chained reciprocals" we discussed in section 2.3.3. Schwarzschild also states some conditions (not reviewed here) that restrict the ways in which the reciprocal's indices can be bound by the indices of the *Part* and *PPart* operators, appropriately expressing the limited binding options available to reciprocals. For these reasons, I find Schwarzschild's system a promising basis for a treatment of reciprocals. Still, it suffers from the same problem as the Heim et al. account with respect to dependent reciprocals. Since the latter are my primary focus, I have based my discussion on the Heim et al. account. In the next section, I show briefly that my proposals can be easily adapted to Schwarzschild's framework.

5.4 Putting it together: Dependent reciprocals and covers

Chapter 4 showed that the problem of expressing the proper semantics of dependent reciprocals in the system of Heim et al. (1991a,b) has a straightforward solution, at least in principle: We should amend their analysis so that it derives the range argument of the reciprocal from the function encoded by its local antecedent, the dependent pronoun. As we saw, Schwarzschild's treatment of reciprocals shares with Heim et al. their "scopal" approach to dependent reciprocals, and therefore its shortcomings as well. Fortunately, this also means that it is straightforward to revise his solution in the same manner as I proposed for the Heim et al. account. Let us develop the resulting account in some more detail.

I retain the essential elements of the proposal developed in chapter 4: dependent pronouns are translated as *restricted anaphoric functions,* and the range argument of the *EachOther* operator is determined not via binding by the *Part* operator, but from the range of the reciprocal's antecedent. As we have seen, it is never possible for the reciprocal to take values from a set other than the set determined by its (local) antecedent, hence this approach will always give us the right results.

Given an appropriate restricted reference function $\lambda x\, r(x)$, we give a dependent pronoun the bipartite structure [$\lambda x\, r(x)\ u$], where u is a variable bound by the distributing operator.

(5.64)

$$
\begin{array}{c}
r(u)_{1,2} \\
\diagup \quad \diagdown \\
\lambda x\, r(x)_1 \qquad u_2
\end{array}
$$

I will assume that we can always derive a restricted reference function from a referring NP by restricting the identity function to the subparts of that NP. Of course if it denotes an atomic individual, the range of the function will have only one element and no reciprocation will be possible.

Several alternative formalisms for obtaining the reciprocal's range from the domain of its antecedent's function were discussed in section 4.4. I will continue to use the range sum operator *RS,* defined as follows:

(5.65) RS(r) = $\sigma z\, (\exists y\, r(y) = z)$ = The maximal z in the range of r (= (4.21b))

We can now modify the reciprocal operator so that it takes as its first argument the reference function corresponding to its antecedent; the second argument is still the contrast argument, i.e., the antecedent itself. This makes *EachOther* of type <ee, <e, e>>. As in Schwarzschild's system, it will be a free variable that can be assigned any contextually appropriate function. To define it, we modify part (b) of Schwarzschild's definition (5.54) as follows:[14]

(5.66) for all M,g:

1. $\forall r_{e,e} \forall v_e \, [\![\text{EachOther}]\!]^{M,g}(r)(v) \subset RS(r)$
2. $\forall r \forall v \; v \not\mathrm{II} \; [\![\text{EachOther}]\!]^{M,g}(r)(v)$
3. $\forall r \forall v \, [\![\text{EachOther}]\!]^{M,g}(r)(v)$ is undefined if $v \not\subseteq RS(r)$

This definition allows us to interpret the reciprocal without the need for non-local binding, on the basis of its local antecedent only. It specifies that the reciprocal takes values among the same set that the contrast argument ranges over.

Recall that the range set operator RS, applied to a function r, gives us a plural individual representing the sum of all elements that r ranges over. This means that if the values of r are non-atomic, this information is not preserved in $RS(r)$. However, condition (5.54c) in Schwarzschild's definition of the *EachOther* function requires that the values generated by *EachOther* must be in the cover used with the *Part* operator that binds the reciprocal. This will restrict the possible values of *EachOther* in cases involving partition into non-atomic subgroups.

Given these definitions, we derive the dependent reading of (5.67a). First we translate the pronoun *them* as the paycheck pronoun mapping lawyers to their clients; the restricted version of this is a function *CL* that maps John and Mary's lawyers to their clients. Then we can apply Schwarzschild's *Part* operator (with a cover containing John's lawyers and Mary's lawyers), to get (5.67b):

[14]Note that the distinctness condition in part *2* of (5.66) has been modified so as to conform to formula (5.56). This change is aimed at having the proper truth conditions when there is overlap between elements of the cover, and is not related to the issue of the range of dependent reciprocals.

(5.67) a. The lawyers that represent John and Mary advised them to sue each other.

b. CL$(x) = \lambda x\, \iota z\, (x\ \Pi$ ANT$_2$ & client-of$(x)(z))$

c. (lawyers)$_2$ Part$_1$ [advised CL(x_1) to sue EachOther(CL)(CL(x_1))]

5.4.1 Chained reciprocals

Schwarzschild translates the reciprocal as the variable *EachOther,* whose value (a function) is provided by the context. This has the natural consequence that in a sentence with two reciprocals, they will both be interpreted as the same function unless other arrangements are made (such as writing the *EachOther* function with a subscript to allow separate interpretations). The default prediction dovetails nicely with a phenomenon that Heim et al. (1991b) pointed out, but could only account for by stipulation: In a sentence like (5.68), they judge that the two reciprocals must operate in tandem, so that each wife was told lies about her own husband.

(5.68) They told each other's wives lies about each other.

If *EachOther* is a functional variable as in Schwarzschild's system, it follows that if it is bound only once it will cause the two reciprocals to vary in tandem. Schwarzschild (p. 123) points this out but expresses reservations about the correctness of such an analysis since, as he says, sentence (5.69) "allows that X put pictures of Y in Z's album, where Y $\neq$ Z," suggesting that the two reciprocals must be translated as different functions.

(5.69) They put pictures of each other in each other's albums.

Given our revised analysis, however, there is another way to account for this reading. Recall that a reciprocal may take another reciprocal as its antecedent. If we treat example (5.69) as an instance of chained reciprocals, we get exactly the distinctness condition that

Schwarzschild describes: that each picture must be inserted in an album that does not belong to the person that it depicts.[15]

(5.70) They put pictures of each other in each other's albums.

The question is whether our semantics for reciprocals can derive such readings. In the scopal analysis the higher reciprocal takes values that are smaller than the whole of its antecedent, since in each instance they must exclude the current value of the contrast argument. For this reason, the values taken by the higher reciprocal are too small to serve as the range argument of the lower reciprocal.[16] In our revised, "non-scopal" version, however, the higher reciprocal can serve as the antecedent function for the lower reciprocal. Since the range of the reciprocal is equal to its domain, applying the range operator to the function *EachOther(r)* represented by the higher reciprocal gives us the correct range argument for the lower one. The resulting translation can be roughly rendered as follows. (The expression *EachOther* has been abbreviated as *EO* for typographical convenience). Here r is the identity function over parts of the subject, *they*$_2$*;* hence *EO(r)* is a function of type $<$e, e$>$ defining reciprocation over parts of the range of r.

(5.71) They$_2$ Part$_1$ put pictures of EO$(r)(x_1)$ in EO(EO$(r))($EO$(r)(x_1))$'s albums.

In this way, the use of restricted functions allows us to immediately derive the correct semantics for chained reciprocals.

[15] Note that the antecedent of the second reciprocal must be *each other,* not *pictures of each other.* I will not delve into the syntactic issues involved, but note that structurally such dependence is definitely possible:

(i) I put pictures of the boys in each other's albums.

[16] This argument is given in more detail in section 2.3.3. Heim et al. (1991b) proposed an account of chained reciprocals along such lines, but section 2.3.3 shows that for the reason discussed here, the semantic translation they propose is incorrect.

Chapter 6

The Variable-Free Approach

This rather technical chapter revisits the analysis of dependent reciprocals developed in chapter 4. Until now, I have only given an ad hoc treatment of examples in which the local antecedent of a dependent reciprocal is an NP containing a possessive pronoun. In this chapter I show that by adopting the framework of Jacobson (1999a,b) Variable Free Semantics, we derive a double benefit: We can provide a much cleaner version of the analysis developed in chapter 4, which in addition applies straightforwardly to this type of examples.[1]

6.1 A short review: The range of dependent reciprocals

Let us begin by recalling the essential problem that the scopal analysis of reciprocals is faced with. It is brought out by the existence of dependent reciprocal readings for sentences such as those in (6.1):

(6.1) a. Their$_i$ coaches think they$_i$ will defeat each other.

[1] An earlier version of this chapter was presented at the 1999 Amsterdam Colloquium (see Dimitriadis 1999a).

b. The lawyers that represent them$_i$ say they$_i$ will sue each other.

c. John and Mary think their mothers like each other.

The dependent reciprocal reading of (6.1c), for example, says that John thinks that his mother likes Mary's mother, and Mary thinks that her mother likes John's mother. In sentences (6.1a) and (b) the intended antecedent of the embedded subject does not c-command it. This necessitates a translation of the embedded pronoun as a paycheck (functional) pronoun, in the style of Engdahl (1986): the pronoun *them* in (6.1b) is translated as $W(u)$, where W is an open variable of type <e, e>, bound by the context to the function taking lawyers to their clients, while u is a variable bound by the distributor that ranges over the matrix subject, *their coaches*. In this way we account for the variable-like behavior of the embedded pronoun even though its intended antecedent does not c-command it.

Even with the paycheck analysis, however, a "scopal" analysis of the reciprocal cannot give the correct readings of these sentences. To see why, consider the dependent reading of example (6.1b). A correct translation should state that each lawyer, x, says that x's client will sue the other clients (or: the other lawyers' clients). To begin with, the system of Heim et al. (1991a,b) predicts that this sentence should be ungrammatical, since it violates their condition that the local antecedent (the pronoun *they*) should be bound by the distributor that binds the pronoun.[2] If, contrary to the restrictions in their system, we were to allow a long-distance reciprocal in this sentence, it might receive the following interpretation:

(6.2) $(\forall x_2\ \Pi\ \text{lawyers}_1)\ \text{say}(\hat{}[\forall x_3(x_3\ \Pi\ X_1\ \&\ W(x_2) \neq x_3)\ \text{sue}(x_3)(W(x_2))])(x_2)$

(W is the function mapping lawyers to their clients).[3]

[2]Heim et al. stipulate that the contrast argument of the reciprocal must be locally A-bound; in effect, that the long-distance binder must be coindexed with the local antecedent. This requirement leads to the prediction that the dependent reading of sentence (6.1b) is impossible: in this example the local antecedent ranges over clients, but the distributor (which is adjoined to the entire subject, not to the embedded pronoun *them*) ranges over lawyers.

[3]This chapter uses a notation for function arguments different from that used in earlier chapters. In this chapter, arguments are written in separate parentheses and saturated from the inside out. For example,

The range argument of the reciprocal, the free variable X, is required to be coindexed with the set over which the long-distance binder (the matrix distributor) quantifies. Because in this case this is the set *lawyers*$_1$, (6.2) says, incorrectly, that each lawyer expects his or her client to sue some *lawyers*. The correct translation would be generated if we could stipulate that X should be interpreted as the set of clients, instead of being coindexed with *lawyers*$_1$. But given the scopal analysis adopted by Heim et al., there is no principled way to require this; to allow pragmatic considerations to come into play here would imply that the range of the reciprocal could be potentially any set, depending on the context, when in fact it is rigidly determined: it can only be the set of entities that the reciprocal's *contrast argument,* $W(x_2)$, ranges over. In other words, we would get the correct semantics if we could somehow replace X_1 with the range of W.

This problem is not limited to the account of Heim et al.; its equivalent, as we have seen in chapter 5, is also found in other scopal treatments of reciprocals, including those of Sternefeld (1998) and Schwarzschild (1996). Sternefeld, for example, captures the semantics of weak reciprocity via membership in the cumulation of the reciprocal predicate. (In contrast, the account of Heim et al. relies on direct quantification and encodes the semantics of strong reciprocity). In dependent reciprocal sentences the entire reciprocal raises to the matrix clause, so that the dependent reading of (6.1b) would be as follows:

(6.3) $(\exists X)(X = \text{lawyers}_1 \;\&\; {<}X, X{>} \in {**}\lambda x\, \lambda y\, [x \neq y \;\&\; \text{say}(^\wedge[\text{sue}(y)(W(x))])(x)])$

This is the same meaning given by the Heim et al. analysis: it claims that each lawyer x said that his or her client, $W(x)$, will sue one or more of the other *lawyers*, y. The correct semantics would require replacing the last occurrence of y with $W(y)$.

---

the property *ate* is written as $\lambda x\, \lambda y$ ate$(x)(y)$, from which we get the proposition *ate*(apple$_1$)(Mary). This notation is more convenient when manipulating higher-arity functions over several steps; it is also the notation Jacobson uses.

In the dependent reciprocal examples we have been considering, the local antecedent of the reciprocal is an expression containing a variable bound by a non-local distributor but ranges over entities different from those that the distributor ranges over (for example, clients vs. lawyers). The problem with such examples is that, as these examples show, the reciprocal's range argument is actually determined by its *local* antecedent: In sentence (6.1c) this is the NP *their mothers,* so the dependent reading is that John thinks his mother likes Mary's *mother.* However, the scopal analysis of reciprocals predicts that the range argument is determined not by the local antecedent, but by the distributor binding the reciprocal: wide scope binding in (6.1c) predicts, at best, that John's mother likes Mary and vice versa. Such interpretations are never possible: reciprocals always range over the same elements their local antecedent ranges over, regardless of what *that* may be dependent on.

What is needed is some way for the object of the reciprocal predicate to range over the elements of the embedded, not the matrix, subject; for example, in sentence (6.1c) that would be the set of mothers. But if the "long-distance" reading involves the translation of the embedded subject as a bound variable, the expression *their mothers* ranges over one mother at a time—and there is no potential antecedent anywhere in (6.1c) that evaluates to the set of mothers!

The failure to find the right range for the reciprocal, then, is not a problem specific to the analysis of Heim et al. The core of the problem is that the range of the dependent reciprocal depends on its local antecedent, but the scopal analysis cannot properly take its contribution into account. Indeed, it can be argued that the essence of the scopal analysis lies precisely in excluding the local antecedent: the core claim of the scopal analysis is that in a dependent ("long distance") reciprocal sentence, the reciprocal enters into a binding relationship with the matrix subject *instead* of with the local subject. In the case of traditional dependent reciprocal examples such as (6.4), the two antecedents have identical ranges, and the correct interpretation results. But examples such as those in (6.1) demonstrate that the scopal analysis derives the range of the reciprocal from the wrong place.

(6.4) John and Mary think they like each other.

6.1.1 Deriving the range argument

In chapter 4, I represented paycheck pronouns as *restricted* reference functions; their domain was taken to be all parts of the plural individual that their binder ranges over. In the examples we have been considering, the domain is the matrix subject. Given a restricted function of this sort, its domain and range can be recovered by applying a maximality operator. In chapter 4 I used the *range sum* operator defined as in (6.5). It gives the sum of all individuals in the range of r.

(6.5) $\sigma z\,(\exists y\, r(y) = z)$ = RS(r) = The maximal z in the range of r (= (4.21b))

For concreteness, consider how the translation of Heim et al. (1991b) needs to be modified in order to account for a sentence like (6.1b). Their translation of the reciprocal can be written as follows:

(6.6) $\lambda P\, \lambda y\, \forall x_k (x_k \, \Pi \, X_i \;\&\; x_k \neq y)\, P(x_k)(y)$

This would combine first with a transitive verb and then with the bound variable representing the embedded subject (the dependent pronoun). We replace the range argument, the free variable X_i, with the expression $RS(r)$, in which r is a free variable that is required to match the reference function of the antecedent. After we also adjust the distinctness condition $x_k \neq y$ to allow evaluation over non-atomic individuals, we obtain:

(6.7) $\lambda P\, \lambda y\, \forall x_k (x_k \in RS(r) \;\&\; y \not\amalg x_k)\, P(x_k)(y)$

The restrictor clause lets x_k range over all elements that are in the range of r and are not supersets of, or equal to, the contrast argument y. The distinctness condition $y \not\Pi x_k$ says that the contrast argument y may not be a part of x_k, the argument that occupies the object position in the reciprocal predicate; the definition of Π considers every individual to be "part" of itself, hence this is condition equivalent to $y \neq x_k$ when both y and x_k are atomic.

In this way we have accounted for dependent reciprocals with paycheck pronouns by replacing one free variable (the X_i argument) with another. Although the range of the reciprocal is now based on the domain of its antecedent's reference function, its determination still involves a free variable, the reference function r. Given that the reciprocal range is rigidly determined, it would be more appropriate for the reference function to be passed to the reciprocal as an argument. But although the paycheck pronoun containing the reference function is visible as the argument y, the function is only part of the translation of the paycheck pronoun. The pronoun itself is an expression $r(u)$ of type <e>, consisting of a variable r representing the reference function (type <e, e>) plus a variable u of type <e> that saturates the argument of the reference function. The pronoun is then visible to the reciprocal predicate only as an individual, albeit an assignment-dependent one; the reference function itself cannot be retrieved for use by the reciprocal.

Because the value of the free variable can be specified through reference to the sentence structure (r is the reference function embedded in the reciprocal's local antecedent), this account is no less constrained than that of Heim et al. (1991b); but it is no better, either.

What we really need is a way to let the reciprocal predicate combine directly with the reference function of its local antecedent. This can be done in a natural way in the Variable-Free Semantics of Jacobson (1999a,b), in which pronouns are represented as elements of type <e, e>, without the saturating argument. The resulting account is cleaner in its combinatorics, and allows us to treat sentences involving NPs with dependent possessive pronouns (such as example (6.1c)) on a par with those involving paycheck pronouns.

6.2 Variable-free semantics

The goal of Jacobson's program is to eliminate from the semantics all expressions that are dependent on the variable assignment function, $[\![\cdot]\!]^g$. As Jacobson points out, an assignment-dependent expression of type α is implicitly a function from the variables that it depends on to expressions of type α. Jacobson represents such expressions as explicitly having this implicit type. In particular, pronouns are represented as functions (of type <e, e>), not as variables over individuals. Typically, a pronoun represents the identity function on individuals, $\lambda x\, x$.

A sentence containing a single pronoun might be dependent on the value of a single open variable of type <e>; in Jacobson's framework, the sentence will receive an assignment-independent translation of type <e, t>. From this we derive a proposition by applying it to a salient, context-supplied individual. This is no worse, and is more explicit, than relying on the proper choice of assignment function to supply us with the desired individual. (Jacobson argues that her approach is actually more natural, since it is counter-intuitive to speak of a "salient" assignment function).

This means that syntactically identical expressions can have different semantic types, depending on whether they contain pronouns etc. Jacobson adopts a rich system of syntactic types that distinguishes, e.g., a transitive verb (type (S/NP)/NP) from a VP containing a pronoun (type $(\mathrm{S/NP})^{\mathrm{NP}}$, a function from NPs to S/NP s).[4] Semantically both forms are of type <e, et>, but $(\mathrm{S/NP})^{\mathrm{NP}}$ does not license an object.

The operators g and z allow unsaturated functions to combine properly with other sentence elements. The operator g achieves the equivalent of function composition, while z, applied to a transitive expression like *love,* replaces one of its arguments with a function applied to a higher argument. Here I will not give fully explicit definitions of Jacobson's

[4] Jacobson's types also indicate the direction from which an argument is expected by means of a subscript on the slash: $/_R$ vs. $/_L$. I will mostly ignore directionality in this exposition.

operators, but will provide simplified representations that are sufficient for our purposes. The complete definitions can be found in (Jacobson 1999a).

(6.8) **The *g* operator** allows a function argument to be passed up; it is parametrized (implicitly, in our simplified representation) so that it can bypass any number of *semantic* arguments of its complement and apply to the first syntactic argument.

Syntax: g_C(B/A) = B^C/A^C

Semantics: <a, b> $\mapsto$ <ca, cb>

Formula: $g_c(\lambda a\, t(a)) = \lambda f\, \lambda c\, [t(f(c))]$

For example, g can be employed to transform $\lambda x\, \lambda y\,$love$(x)(y)$ into either of the expressions $\underline{\lambda f\, \lambda c}\, \lambda y\,$love$(f(c))(y)$ or $\lambda x\, \underline{\lambda f\, \lambda c}\,$love$(x)(f(c))$. The first translation would apply to the ordinary transitive verb, while the second case corresponds to the VP *loves him* (type (S/NP)$^{\mathrm{NP}}$).

The operator g achieves the effect of function composition. For example, an expression of type VP/NP, say *sleeps*, cannot be combined directly with one of type NP$^{\mathrm{NP}}$ (say, *he*), since the types do not match. But g(sleeps) is of type VP/NP$^{\mathrm{NP}}$, and can be combined with the pronoun. Since g(sleeps) is $\lambda f\, \lambda c\,$sleeps$(f(c))$, combining it with a pronoun $\lambda x\, h(x)$ via Functional Application will give us $\lambda x\,$sleeps$(h(x))$, which is the same as the result of function-composing $\lambda c\, h(c)$ with $\lambda x\,$sleeps(x).

(6.9) **The operator *z*** applies to some transitive expression B, replacing an argument x of type X with an argument f of type <e, X>; occurrences of x in B are replaced with $f(y)$, where y is some higher argument of B.

Syntax: z((B/NP)/A) = (B/NP)/A$^{\mathrm{NP}}$

Semantics: <X, eY> $\mapsto$ <eX, eY>

Formula: $z_b(\lambda x\, \lambda y\, p(x)(y)) = \lambda f\, \lambda y\, p(f(y))(y)$

For example, $z[\lambda x\, \lambda y\, \text{love}(x)(y)]$ is $\lambda f\, \lambda y\, \text{love}(f(y))(y)$. Applied to a propositional attitude verb, z will replace a propositional (type <t>) argument with one parametrized by the verb's subject (type <e,t>). Thus z(believe) is $\lambda P\, \lambda x\, \text{believe}(P(x))(x)$.

In addition, there is a family of explicit type-lifting operators:

(6.10) **The type-lifting operator $\boldsymbol{l_B}$** maps any expression to a generalized quantifier over a type B.

Syntax: l_B(A) = B/(B/A)

Semantics: <a> $\mapsto$ <ab, b>

Formula: $l_b(a) = \lambda P\, P(a)$

The lifting operator allows non-arguments, including adjuncts, to be treated like arguments that can undergo g; similarly, an NP can be lifted so that a complement VP becomes combinatorially its argument. This operation allows constituents containing pronouns to combine with higher sentence elements, since arguments introduced by lifting can undergo g, effectively passing the pronoun across the higher constituent. For example, the subject NP *Mary* (type NP) can be combined with a VP containing a pronoun (type (S/NP)$^{\text{NP}}$) by first being lifted to type S/(S/NP) and then undergoing g to type S$^{\text{NP}}$/(S/NP)$^{\text{NP}}$. An example of how this works is given below in (6.14).

With the aid of z, the functions corresponding to bound pronouns are eventually supplied with the correct antecedent as their argument. A simple derivation involving a bound pronoun proceeds as follows:

(6.11) John thinks he lost.

(6.12) a. he = $\lambda x\, x$ (type NP$^{\text{NP}}$)

b. lost = $\lambda x\, \text{lost}(x)$ (type S/NP)

c. g(lost) = $\lambda f\, \lambda y\, \text{lost}(f(y))$ (type S$^{\text{NP}}$/NP$^{\text{NP}}$)

d. he g(lost) = λy lost$([\lambda x\, x](y))$ = λy lost(y)

e. think = $\lambda P_{<\mathrm{t}>}\, \lambda x$ think$({}^\wedge P)(x)$ (type (S/NP)/S)

f. z(think) = $\lambda Q_{<\mathrm{e,t}>}\, \lambda x$ think$[{}^\wedge Q(x)](x)$ (type (S/NP)/S^{NP})

g. z(think) + he g(lost) = λx think$[{}^\wedge$lost$(x)](x)$

h. John thinks he lost = think$[{}^\wedge$lost$(j)](j)$.

Discourse-referring pronouns remain unbound; the sentences they occur in translate into functions from individuals to sentences. For example, sentence (6.13) is translated as in (6.14). In this derivation, note the application of the lift operator l followed by g in steps (d, e); these operations allow the unbound pronoun in *thinks he lost* to be passed across the subject *Mary.*

(6.13) Mary thinks he lost.

(6.14) a. he g(lost) = λx lost(x) <as before>.

b. g(think) = $\lambda Q_{<\mathrm{e,t}>}\, \lambda c\, \lambda x$ think$[{}^\wedge Q(c)](x)$ (type $(\mathrm{S/NP})^{\mathrm{NP}}/\mathrm{S}^{\mathrm{NP}}$)

c. g(think) + he g(lost) = $\lambda c\, \lambda x$ think$[{}^\wedge$lost$(c)](x)$ (type $(\mathrm{S/NP})^{\mathrm{NP}}$)

d. l(Mary) = $\lambda P_{<\mathrm{e,t}>}\, P(M)$ (type S/(S/NP))

e. $g(l(\text{Mary}))$ = $\lambda Q_{<\mathrm{e,et}>}\, \lambda y\, (Q(y))(M)$ (type $\mathrm{S}^{\mathrm{NP}}/(\mathrm{S/NP})^{\mathrm{NP}}$)

f. $g(l(\text{Mary}))$ + g(think) he g(lost) =

$\lambda y\, ([\lambda c\, \lambda x$ think$[{}^\wedge$lost$(c)](x)](y))(M)$ =

λy think$[{}^\wedge$lost$(y)](M)$ (type S^{NP})

The context then supplies a salient individual as the referent of *he.*

It is crucial that z will only replace an argument with a function whose argument is a *higher* argument of the expression z transforms. This condition serves to rule out weak crossover configurations, which could only be generated by allowing a pronoun to be bound by the base position of an operator that crosses over it (see (Jacobson 1999a) for the details). In addition, z is defined so that it only applies to pairs of *syntactic* arguments of the expression it operates on. For example, it may be applied to an expression of type (S/NP)/NP, which has two syntactic arguments, but not to one of type (S/NP)$^{\mathrm{NP}}$.

6.2.1 Dependent pronouns in larger NPs

An immediate benefit of the variable-free approach is that lexical NPs containing a dependent pronoun are compositionally translated as functions, and can therefore be handled just like simple pronouns. In section 2.3.2.2 we saw that sentences like (6.1c), in which the dependent pronoun was embedded in an NP, are problematic for the standard treatment since the local antecedent of the reciprocal is not coreferential with the matrix subject. But such sentences also pose a challenge for an analysis that appeals to the domain of the paycheck pronoun function, since there is no paycheck pronoun here.[5] In the variable-free system the NP *their mothers* in (6.1c), repeated below, is translated compositionally into the function $\lambda x\, \iota y$ *mother-of$(x)(y)$. Combinatorially it behaves just like a pronoun, being of syntactic type NP$^{\mathrm{NP}}$ and semantic type <e, e>, but it encodes a function distinct from the identity function which pronouns represent.

(6.1c) John and Mary think their mothers like each other.

Let us go over the details of how this works. I will adopt Jacobson's (1999b) tentative analysis of genitives, but as she points out, any analysis along the same general lines would work as well. To begin with, Jacobson distinguishes between inherently relational nouns like *mother, wife,* translated as in (6.15a), and non-relational nouns like *woman* that are translated as in (6.15b). The root type of relational nouns is N/$_R$PP, i.e., the underlying syntax is of the form *wife of X*.[6]

[5]The problem posed by this construction was explained in section 4.5.2.

[6]Although the semantic type of relational nouns is <e, et>, their syntactic type presents them as one-place predicates (and hence, they cannot undergo z). Similarly, ordinary nouns are of type <e, t> but do not project

(6.15) a. wife: $\lambda x\, \lambda y$ wife[-of]$(x)(y)$ (type N/$_R$PP : <e, et>)

b. woman: λx woman(x) (type N : <e, t>)

Jacobson posits type-shifting rules that convert the above forms into expressions that accept a prenominal genitive (NP$_{[GEN]}$). The rule for relational nouns is given in (6.16a), and for non-relational nouns in (6.16b).

(6.16) a. Given an expression α of type N/$_R$PP and translation α', there is a homophonous expression β of type NP/$_L$NP$_{[GEN]}$ and translation $\lambda x\, [\iota y\, [\alpha'(x)(y)]]$.

b. Given an expression α of type N and translation α', there is a homophonous expression β of type NP/$_L$NP$_{[GEN]}$ and translation $\lambda x\, [\iota y\, [\alpha'(y)\ \&\ R(y, x)]]$ for some contextually salient relation R.

For example, rule (6.16a) means that there is a form of the relational noun *mother* that takes a prenominal genitive and means the same as *the-mother-of.* Rule (6.16b) might be used, along with the contextually salient relation of possession, to interpret an expression like *Bill's dog* as meaning "the dog belonging to Bill."

To interpret the NP *their mothers* in (6.1c), then, we translate the possessive pronoun *their* in the usual way, as the identity function on individuals. We then use type-shift operation (6.16a) to derive the possessive-compatible form in (6.17a); to this we apply the g operation, which allows it to combine with the pronoun (the identity function ID), giving us (6.17c).

(6.17) a. mothers$_2$ = $\lambda x\, [\iota y$ [*mother-of$(x)(y)]]$ (type $\mathrm{NP}/_{\mathrm{L}}\mathrm{NP}_{[\mathrm{GEN}]}$)

b. g(mothers$_2$) = $\lambda f\, \lambda x\, [\iota y$ [*mother-of$(f(x))(y)]]$ (type $\mathrm{NP}^{\mathrm{NP}}/_{\mathrm{L}}\mathrm{NP}^{\mathrm{NP}}_{[\mathrm{GEN}]}$)

c. their + mothers = $\lambda x\, \iota y$ [*mother-of$(\mathrm{ID}(x))(y)]$ (type $\mathrm{NP}^{\mathrm{NP}}$)

a syntactic argument. This distinction is important to Jacobson's (1999b) treatment of Weak Crossover and i-within-i phenomena, but is not relevant to our present concerns.

In this way, Jacobson's system compositionally translates NPs containing a pronoun (i.e., function) not bound within that NP as functions on their own. The properties of function composition also guarantee that if we treat pronouns as restricted functions as in section 6.1.1, the domain of the NP function is the same as the domain of the pronoun it contains. This is because in general, the domain of the composition $f(g(x))$ of any two functions f and g is just the domain of g. (For the composition to be well-defined, the range of g must be a subset of the domain of f). In the case of an NP containing a pronoun, this means that the domain of the NP's translation will be the same as the domain of the embedded pronoun. In short, the Variable Free framework allows us to treat NPs containing an embedded dependent pronoun just like we'd treat a simple dependent pronoun. This happy state of affairs can be contrasted with the conclusions of section 4.5.2, where only a vague account of such examples could be attempted.

6.2.2 Paycheck pronouns

While ordinary pronouns are treated as the identity map on individuals, paycheck pronouns operate by mapping one individual to another. This suggests that we should allow pronouns to be translated as arbitrary (non-identity) functions; however, to do so would assign to pronouns an infinite number of different possible meanings. The analyses of Cooper (1979) and Engdahl (1986) were more parsimonious: Engdahl translates paycheck pronouns as involving a function-valued open variable, whose value is then supplied by the context. Jacobson adopts the Variable-Free version of this idea: just like the Variable-Free program represents individual-valued variables as the identity map on individuals, function-valued variables are represented as the identity map on *functions*: Jacobson translates paycheck pronouns as the identity map on functions of type <e, e>, $\lambda f_{\text{e,e}}\, f$ (or $\lambda f\, \lambda x\, f(x)$). This is simply the result of applying the g operator to the identity function on individuals, so all pronouns can be given the same underlying representation. To get the sloppy reading of example (6.18), the second sentence is translated as λf hates$(f(B))(B)$. The context supplies the mother function as the argument f.

(6.18) John loves his mother. Bill hates *her.*

The combinatorics of a sentence containing a paycheck pronoun are more complex than those of earlier examples. In this example, the verb *hates* first undergoes z in order to bind the argument of the paycheck function to its subject, allowing the reading *x hates f(x).* Then g is applied to allow the verb to combine with the paycheck pronoun.

(6.19) a. g(her) = $\lambda f\ f$

b. z(hates) = $\lambda f\ \lambda x$ hates$(f(x))(x)$

c. $g(z$(hates)) = $\lambda F_{\text{ee,ee}}\ \lambda h_{\text{ee}}\ \lambda x$ hates$(F(h)(x))(x)$

d. $g(z$(hates)) + g(her) = $\lambda h_{\text{ee}}\ \lambda x$ hates$[[\lambda f\ f](h)(x)](x)$ =

$\lambda h\ \lambda x$ hates$(h(x))(x)$

The subject must then type-lift and undergo g, in order to leave the predicate's argument h unbound. Finally, f is interpreted as the mother function by the discourse context.

(6.20) a. l(Bill) = $\lambda P\ P(B)$

b. $g(l$(Bill)) = $\lambda Q_{<\text{ee,et}>}\ \lambda f\ Q(f)(B)$

c. $g(l$(Bill)) + (hates her) = $\lambda f\ [\lambda h\ \lambda x$ hates$(h(x))(x)](f)(B)$ =

λf hates$(f(B))(B)$

Paycheck pronouns are of type $<$ee, ee$>$, which will from now on be abbreviated as $<$p$>$.

6.3 Reciprocals in the variable-free system

For concreteness, I will use as a starting point the treatment suggested by Heim et al. (1991b), in which the contrast argument of the reciprocal is bound by a covert distributor that is freely inserted, as in simple distributive sentences. For various reasons which need not concern us here (see section 2.3.1), I replace their NP-adjoined distributor with one adjoined to VP, which provides the same universal quantification over the parts of the subject:

(6.21) $D = \lambda P_{<\mathrm{e,t}>}\, \lambda w\, (\forall x\ \Pi\ w) P(x)$

The translation that Heim et al. give to the reciprocal can be written as in (6.6), repeated below. The universal quantifier ranges over the possible objects of the reciprocal predicate, while the argument y, the *contrast* argument of the reciprocal, is bound by a higher universal quantifier introduced by the distributor that ranges over the subject. The *range* argument of the reciprocal is given by the open variable X_i. For the variable-free version we rewrite it as an extra argument of the reciprocal, obtaining the reciprocal translation given in (6.22).

(6.6) $\lambda P\, \lambda y\, \forall x_k (x_k\ \Pi\ X_i\ \&\ x_k \neq y)\ P(x_k)(y)$

(6.22) $\lambda P_{<\mathrm{e,et}>}\, \lambda R_{\mathrm{e}}\, \lambda x\, \forall y (y\ \Pi\ R\ \&\ y \neq x)\ P(y)(x)$

In use the reciprocal combines with a transitive verb, to give a predicate such as the one in (6.23). This is an expression of syntactic type (S/NP)$^{\mathrm{NP}}$: the range argument in effect causes the reciprocal predicate to behave combinatorially as if it contained a pronoun. At any stage, R can be bound via the z operator or remain unbound and be passed up with the help of g. Note that the formulas as given would allow R to remain unbound; but the syntax of reciprocals requires R to be bound in a particular way: in the long-distance analysis of reciprocals, by the individual over which the binder of the reciprocal's contrast argument ranges.

(6.23) $\lambda R_{\text{e}}\, \lambda x\, \forall y(y\ \Pi\ R\ \&\ y \neq x)$ like$(y)(x)$.

In an ordinary distributed sentence, the distributor is directly combined with the VP, and can then be applied to the subject:

(6.24) (John and Mary) + D(ran).

= $[\lambda w\ (\forall x\ \Pi\ w)$ ran$(x)](J \oplus M)$

= $(\forall x\ \Pi\ J \oplus M)ran(x)$

Because of the extra argument of a reciprocal predicate, it cannot combine directly with D; it combines with $z(D)$ instead, as in (6.25b):

(6.25) a. $z(D) = \lambda Q_{<\text{e,et}>}\, \lambda w\ (\forall x\ \Pi\ w) Q(w)(x)$

b. $z(D)$ + (like each other) = $\lambda w\ (\forall x\ \Pi\ w)\ (\forall y\ \Pi\ w\ \&\ y \neq x)$ like$(y)(x)$

In a simple reciprocal sentence like (6.26a), this combines with a plural subject to give (6.26b).

(6.26) a. John and Mary $z(D)$(like each other)

b. $\forall x(x\ \Pi\ J \oplus M)\, \forall y(y\ \Pi\ J \oplus M\ \&\ y \neq x)$ like$(y)(x)$

This is simply the Heim et al. account translated into variable-free semantics, but with VP-adjoined distributors. Because the subject is not buried in the subject-distributor complex as it would be if the distributor was NP-adjoined, the range argument can be recovered from the distributor, as shown, instead of remaining as a free variable. But this account is as incapable of handling dependent paycheck pronouns as the original; if a reciprocal predicate is not combined with a distributor but is instead combined with a dependent paycheck pronoun, as in example (6.27a), we get:[7]

(6.27) a. The lawyers who represent John and Mary say they like each other.

b. g(they) = $\lambda f_{\mathrm{ee}}\, f$ (Paycheck pronoun)

c. $l(g$(they)) = $\lambda P_{<\mathrm{p},<\mathrm{ee},\mathrm{et}>>}\, P(\lambda f\, f)$

d. $g(l(g($(they))) = $\lambda Q_{<\mathrm{e},<\mathrm{p},<\mathrm{ee},\mathrm{et}>>>}\, \lambda R_{\mathrm{e}}\, Q(R)(\lambda f\, f)$

e. (like each other) = $\lambda R\, \lambda x\, \forall y(y\ \Pi\ R\ \&\ y \neq x)$ like$(y)(x)$ (= (6.23))

[7] In this derivation the paycheck pronoun must be used as the subject of a predicate already containing an unsaturated variable (the range argument R). The detailed steps take us too far afield from the points being made in the main text. For the motivated reader they are presented in this footnote with an explanation of each step's function. (Compare the derivation of the paycheck pronoun in section 6.2.2). It is also hoped that my comments on this derivation will help dispel in some readers the growing suspicion that the workings of the Variable-Free calculus are irremediably impenetrable.

In the first three steps, the paycheck pronoun g(they) lifts to $\lambda P_{<\mathrm{p},\,\mathrm{X}>}\, P(\lambda f\, f)$ via l, then undergoes g in order to pass up the reciprocal's range argument R. Recall that <p> is the type of paycheck pronouns, <ee, ee>. The type <X> is an abbreviation for whatever will turn out to be needed here (which we derive below).

(i) The lawyers who represent John and Mary say they like each other.

(ii) a. g(they) = $\lambda f_{\mathrm{ee}}\, f$
Paycheck pronoun.

b. $l(g$(they)) = $\lambda P_{\mathrm{p,X}}\, P(\lambda f\, f)$
P matches an (incomplete) VP expecting a paycheck subject. (<X> = TBA)

c. $g(l(g($(they))) = $\lambda Q_{<\mathrm{e},<\mathrm{p},\mathrm{X}>>}\, \lambda R_{\mathrm{e}}\, Q(R)(\lambda f\, f)$
Q matches an incomplete VP containing a pronoun or variable, *and* expecting a paycheck subject.

d. (like each other) = $\lambda R_{\mathrm{e}}\, \lambda x_{\mathrm{e}}\, \forall y(y\ \Pi\ R\ \&\ y \neq x)$ like$(y)(x)$ (= (6.23))
A VP containing the unbound variable R.

e. g(like each other) = $\lambda R_{\mathrm{e}}\, \lambda f_{\mathrm{ee}}\, \lambda w_{\mathrm{e}}\, \forall y(y\ \Pi\ R\ \&\ y \neq f(w))$ like$(y)(f(w))$
The subject of the VP is a pronoun.

f. $g(g$(like each other)) = $\lambda R_{\mathrm{e}}\, \lambda F_{\mathrm{p}}\, \lambda h_{\mathrm{ee}}\, \lambda w_{\mathrm{e}}\, (\forall y\ \Pi\ R, y \neq F(h)(w))$ like$(y)(F(h)(w))$
The subject of the VP is a *paycheck* pronoun.

The last expression must match the argument Q on line (iic). Since its type is <e, <p, <ee, et>>>, we conclude that the type <X> on line (iic) is <ee, et>.

Finally, we combine pronoun and reciprocal predicate to get:

g. $g(l(g($they))) + $g(g$(like each other)) =
$\lambda R\, [\lambda R\, \lambda F_{\mathrm{p}}\, \lambda h_{\mathrm{ee}}\, \lambda w\, \forall y(y\ \Pi\ R\ \&\ y \neq F(h)(w))$ like$(y)(F(h)(w))](R)(\lambda f\, f)$ =
$\lambda R\, [\lambda h_{\mathrm{ee}}\, \lambda w\, \forall y(y\ \Pi\ R\ \&\ y \neq (\lambda f\, f)(h)(w))$ like$(y)((\lambda f\, f)(h)(w))]$ =
$\lambda R_{\mathrm{e}}\, \lambda h_{\mathrm{ee}}\, \lambda w_{\mathrm{e}}\, \forall y(y\ \Pi\ R\ \&\ y \neq h(w))$ like$(y)(h(w))$

As discussed in the body of the text, this formula applies the paycheck function only to the subject, not to the object, of *like.* This eventually results in the wrong semantics for the sentence.

f. $g(g(\text{like each other})) =$

$\lambda R\, \lambda F_{\text{p}}\, \lambda h_{\text{ee}}\, \lambda w\, \forall y(y\ \Pi\ R\ \&\ y \neq F(h)(w))\, \text{like}(y)(F(h)(w))$

g. $g(l(g(\text{they}))) + g(g(\text{like each other})) =$

$\lambda R_{\text{e}}\, \lambda f_{\text{ee}}\, \lambda w_{\text{e}}\, \forall y(y\ \Pi\ R\ \&\ y \neq f(w))\, \text{like}(y)(f(w))$

In step (e) it can be seen that the paycheck function f is only applied to the subject, not the object, of *like*; when the entire sentence is translated, w will be bound by the matrix distributor, the range argument R will be identified with the denotation of the matrix subject, and the context will assign to f the function mapping lawyers to clients. The result is the non-existent reading of clients liking lawyers.

6.3.1 The remainder of the derivation

To further illustrate the operation of the Variable-Free system, let us walk through the remaining steps of the derivation of (6.27a). The arguments f and w in (6.27g) were contributed by the paycheck pronoun, while the argument R is the range argument of the reciprocal. As the derivation continues, the verb *say* undergoes the z operation once to allow w to be bound by the higher argument of *say*, and the g operation twice on its propositional complement to allow R and f to be passed through. The embedded clause (6.27g) is then passed to the result, giving (6.28e).

(6.28) a. say = $\lambda P_{<\text{t}>}\, \lambda x\, \text{say}({}^\wedge P)(x)$

b. $z(\text{say}) = \lambda Q_{<\text{e,t}>}\, \lambda x\, \text{say}({}^\wedge Q(x))(x)$

c. $g(z(\text{say})) = \lambda P_{<\text{ee,et}>}\, \lambda f\, \lambda x\, \text{say}({}^\wedge P(f)(x))(x)$

d. $g(g(z(\text{say}))) = \lambda Q_{<\text{e},<\text{ee,et}>>}\, \lambda R\, \lambda f\, \lambda x\, \text{say}[{}^\wedge Q(R)(f)(x)](x)$

e. $g(g(z(\text{say}))) + (\text{they like each other}) =$

$\lambda R\, \lambda h\, \lambda x\, \text{say}[{}^\wedge \forall y(y\ \Pi\ R\ \&\ y \neq h(x))\ \text{like}(y)(h(x))](x)$

This expression must now be combined with a distributor, and then with the matrix subject. Because relative clauses are strong islands, a distributor inside the relative clause cannot raise to bind variables outside it.[8] Therefore, the internal structure of the complex NP is of no interest at this time; it is translated as an ordinary plural individual, which we can write as:

(6.29) (The lawyers that represent John and Mary) = lawyers$_1$

Although the semantics might allow the range argument R to remain unbound and be supplied by the discourse context, the syntax of reciprocals requires that R must be bound by the individual over which the binder of the reciprocal's contrast argument ranges, i.e., by the individual over which the matrix distributor ranges. This is accomplished by applying z to the distributor, just as in example (6.26). First, however, the distributor must undergo g in order to allow the paycheck function f to be passed up:

(6.30) a. D = $\lambda P_{<\mathrm{e,t}>}\, \lambda w\, (\forall x\ \Pi\ w) P(x)$

b. $g(D)$ = $\lambda Q_{<\mathrm{ee,et}>}\, \lambda f_{<\mathrm{ee}>}\, \lambda w\, (\forall x\ \Pi\ w)\, Q(f)(x)$

c. $z(g(D))$ = $\lambda N_{<\mathrm{e},<\mathrm{ee,et}>>}\, \lambda f\, \lambda w\, (\forall x\ \Pi\ w) N(w)(f)(x)$

We then combine this with the expression derived in (6.28e), to get:

(6.31) $z(g(D))$ + (say they like each other) =

$\lambda f\, \lambda w\, (\forall x\ \Pi\ w)$

$[\lambda R\, \lambda h\, \lambda x\, \text{say}[\text{^}\forall y(y\ \Pi\ R\ \&\ y \neq h(x))\, \text{like}(y)(h(x))](x)]\, (w)(f)(x)$ =

$\lambda f\, \lambda w\, (\forall x\ \Pi\ w)[\text{say}[\text{^}\forall y(y\ \Pi\ w\ \&\ y \neq f(x))\, \text{like}(y)(f(x))](x)]$

[8] In section 3.4, I discuss the resemblances between these phenomena and the "functional relative clauses" studied by Sharvit (1999). The examples she studies involve overt quantifiers embedded in relative clauses, and her findings are not applicable to the distributed definite NPs at work here.

Note that in the last step the range argument R was identified with the argument w, over which the distributor will range. As in the simple reciprocal case, this results from applying z to the distributor, and satisfies the condition that the range argument of the reciprocal should be the same as the set over which the distributor binding the reciprocal ranges. As I have noted this condition is correct for non-long-distance (i.e., non-dependent) reciprocal examples; it also gives the correct result for the dependent reciprocal examples studied by Heim et al., but not, as we have seen, for examples like (6.27a).

We are now almost done. The remaining steps lift the subject and apply g, in order to pass the function f across it; then the subject can be combined with the distributed predicate in (6.31), giving (6.32c):

(6.32) a. $l(\text{lawyers}_1) = \lambda P\, P(\text{lawyers}_1)$

b. $g(l(\text{lawyers}_1)) = \lambda P\, \lambda f\, P(f)(\text{lawyers}_1)$

c. $\lambda f\, (\forall x\ \Pi\ \text{lawyers}_1)\text{say}[{}^\wedge\forall y(y\ \Pi\ \text{lawyers}_1\ \&\ y \neq f(x))\ \text{like}(y)(f(x))](x)$

The discourse then binds f to the function mapping lawyers to clients. As we have already seen, this formula improperly makes the following claim: Every lawyer x in the set of John and Mary's lawyers says that his or her client $f(x)$ likes every *lawyer* not identical to that client.

6.4 Using the range

What kind of adjustments can be made to give the right interpretation? As discussed in section 6.1.1, in the standard (assignment dependent) semantics the reciprocal translation (6.6) given by Heim et al. can be replaced by one that explicitly uses the range of the pronoun's reference function, as in (6.7).

(6.6) $\lambda P\, \lambda y\, \forall x_k (x_k\ \Pi\ X_i\ \&\ x_k \neq y)\, P(x_k)(y)$

(6.7) $\lambda P\, \lambda y\, \forall x_k (x_k \in RS(r)\ \&\ y \not\Pi\ x_k)\, P(x_k)(y)$

The reference function itself had to be represented as an unbound variable. In the Variable-Free system, we have the option of direct access to the pronoun's function, and a better analysis is possible. If the translation of the reciprocal can expect to be combined with a functional subject, we can dispense with the range argument added in translation (6.22) and derive the reciprocal's range from the range of the antecedent function, translating the reciprocal as in (6.33):[9]

(6.33) $\lambda P_{\mathrm{e,et}}\, \lambda r_{\mathrm{ee}}\, \lambda x\, \forall y (y\ \Pi\ RS(r)\ \&\ y \not\Pi\ r(x))\, P(y)(r(x))$

This approach is the variable-free adaptation of the proposal developed in chapter 4. In order for it to work, it is again necessary that the pronoun functions come with domains (from which the range is easily computed), and that their domain be no larger than required by the pronoun's binder, i.e., the domain of a dependent pronoun should be co-extensive with the matrix subject. Otherwise, the application of the domain operator to an unrestricted function would return too large a set (perhaps, the set of all individuals in our model!) for the range argument of the reciprocal.

6.4.1 Deriving the dependent reading

Let us begin by using formula (6.33) to compute the meaning of example (6.1c), as always under the dependent reading. The meaning of *their mothers* is as given in (6.17c). Since it contains a non-paycheck pronoun, is of the same type as a pronoun. This makes the

[9] In formula (6.33), the argument λr corresponds to the function appearing as the local antecedent of the reciprocal (embedded subject), while λx corresponds to *its* argument. The range is derived from r, rather than being supplied as an additional argument as in (6.22).

The syntactic type of the reciprocal is $(\mathrm{S}^{\mathrm{NP}}/\mathrm{NP}^{\mathrm{NP}})/((\mathrm{S}/\mathrm{NP})/\mathrm{NP})$.

derivation of this example almost trivial: The reciprocal first combines with the verb *like,* then with the reciprocal subject *their mothers.* No type-shifting is necessary.

(6.1c) John and Mary think their mothers like each other.

(6.34) a. each other = $\lambda P_{\mathrm{e,et}}\,\lambda r_{\mathrm{ee}}\,\lambda x\,\forall y(y\ \Pi\ RS(r)\ \&\ y\ \not{\mathrm{I}}\ r(x))\,P(y)(r(x))$

b. like = $\lambda x\,\lambda y\,\mathrm{like}(x)(y)$

c. like each other = $\lambda r_{\mathrm{ee}}\,\lambda x\,\forall y(y\ \Pi\ RS(r)\ \&\ y\ \not{\mathrm{I}}\ r(x))\,\mathrm{like}(y)(r(x))$

d. their mothers = $\lambda x\,\iota y\,[\text{*mother-of}(\textsc{id}(x))(y)] = \lambda x\,\mathrm{m}(x)$ (= (6.17c))

e. (their mothers) + (like each other) =

$\lambda x\,\forall y(y\ \Pi\ RS(\lambda x\,\mathrm{m}(x))\ \&\ y\ \not{\mathrm{I}}\ \mathrm{m}(x))\,\mathrm{like}(y)(\mathrm{m}(x))$

Next, the verb *think* undergoes z in order to bind the pronoun in the embedded sentence to the subject of *think.* The result does not contain any unbound pronouns, and so combines with the distributor directly. Finally, the result is applied to the matrix subject.

(6.35) a. think = $\lambda P_{<\mathrm{t}>}\,\lambda x\,\mathrm{think}(^\wedge P)(x)$

b. z(think) = $\lambda Q_{<\mathrm{e,t}>}\,\lambda x\,\mathrm{think}[^\wedge Q(x)](x)$

c. z(think) + (their mothers like each other) =

$\lambda x\,\mathrm{think}[^\wedge\forall y(y\ \Pi\ RS(\lambda x\,\mathrm{m}(x))\ \&\ y\ \not{\mathrm{I}}\ \mathrm{m}(x))\,\mathrm{like}(y)(\mathrm{m}(x))](x)$

d. $D = \lambda P_{<\mathrm{e,t}>}\,\lambda w\,(\forall x\ \Pi\ w)P(x)$

e. D + (think their mothers like each other) =

$\lambda w\,(\forall x\ \Pi\ w)\mathrm{think}[^\wedge\forall y(y\ \Pi\ RS(\lambda x\,\mathrm{m}(x))\ \&\ y\ \not{\mathrm{I}}\ \mathrm{m}(x))\,\mathrm{like}(y)(\mathrm{m}(x))](x)$

f. (6.1c) =

$(\forall x\ \Pi\ J\oplus M)\mathrm{think}[^\wedge\forall y(y\ \Pi\ RS(\lambda x\,\mathrm{m}(x))\ \&\ y\ \not{\mathrm{I}}\ \mathrm{m}(x))\,\mathrm{like}(y)(\mathrm{m}(x))](x)$

The expression we have derived lets the reciprocal range over the set of values returned by the mother function, giving us the correct semantics as long as this function is properly restricted. Recall that the full form of the function I have abbreviated as $m(x)$ is

$\lambda x\, \iota y\, [{*}\textit{mother-of}(\textsc{id}(x))(y)]$, where ID() is the identity function translating the pronoun *their.* Once again, we assume that this function has the smallest possible domain, the set $\{$*John, Mary*$\}$. It is worth noting here that the translation of the noun *mothers* does not itself need to have a domain that is special to this sentence.[10] As explained in section 6.2.1, the laws of function composition guarantee that $\textit{m}(x)$, the result of composing ID() with the mother function, will have the same domain. Hence the range of $\textit{m}(x)$ is just the set of John and Mary's mothers, as it should be.

What about an ordinary dependent reciprocal sentence, such as (6.4)? Its derivation proceeds exactly as above, except that in place of the mother function $\lambda x\, \textit{m}(x)$, we have the identity function $\lambda x\, x$. Again, this function is restricted to the smallest domain needed. The result is given as (6.36):

(6.4) John and Mary think they like each other.

(6.36) $(\forall x\ \Pi\ J{\oplus}M)$think$[{}^\wedge\forall y(y\ \Pi\ RS(\lambda x\, x)\ \&\ y\ \not\Pi\ x)$ like$(y)(x)](x)$

6.4.2 Paycheck pronouns

For a more complicated example, let us return to example (6.27a). Because it involves a paycheck pronoun, its derivation is somewhat more complicated. First the reciprocal combines with the verb *like* as before, but then the result undergoes g in order to prepare for the paycheck pronoun *they.* Since the reciprocal VP does not have a free range argument that must be passed up as in the derivation of section 6.3, the embedded pronoun does not need to lift; it combines directly with the reciprocal predicate, as shown in (6.37f).

(6.27a) The lawyers who represent John and Mary say they like each other.

(6.37) a. each other = $\lambda P_{\mathrm{e,et}}\, \lambda r_{\mathrm{ee}}\, \lambda x\, \forall y(y\ \Pi\ RS(r)\ \&\ y\ \not\Pi\ r(x))\, P(y)(r(x))$

[10] Since pronouns are anaphoric but common nouns are not, domains for common nouns are conceptually less well motivated. The persistence of pronoun domains under composition means that the present account need not rely on domains for common nouns.

b. like = $\lambda x\, \lambda y$ like$(x)(y)$

c. like each other = $\lambda r_{\text{ee}}\, \lambda x\, \forall y(y\ \Pi\ RS(r)\ \&\ y\ \not\mathrm{Y}\ r(x))$ like$(y)(r(x))$

d. g(like each other) =

$\lambda F_{\text{p}}\, \lambda h_{\text{ee}}\, \lambda x\, \forall y(y\ \Pi\ RS(F(h))\ \&\ y\ \not\mathrm{Y}\ F(h)(x))$ like$(y)(F(h)(x))$

e. g(they) = $\lambda f\ f$

f. g(they) + g(like each other) =

$\lambda h\, \lambda x\, \forall y(y\ \Pi\ RS([\lambda f\ f](h))\ \&\ y\ \not\mathrm{Y}\ [\lambda f\ f](h)(x))$ like$(y)([\lambda f\ f](h)(x))$ =

$\lambda h\, \lambda x\, \forall y(y\ \Pi\ RS(\textsc{id}(h))\ \&\ y\ \not\mathrm{Y}\ h(x))$ like$(y)(h(x))$

Although the result of applying $\lambda f\ f$ to the paycheck function h is just h, the range of the result depends on the range of both: If f is the identity function on functions whose domain is given set of lawyers, only such functions can be provided by the context. For this reason I have written the argument of the *RS* operator as ID(h), rather than just h. The argument h should eventually express the function mapping lawyers to clients (which will be supplied by the context). It will be passed up during the remainder of the derivation, which is otherwise straightforward. First, *say* undergoes z to bind the argument of the paycheck function, then it undergoes g to pass up the paycheck function itself. Then the result combines with the subordinate clause:

(6.38) a. say = $\lambda P\, \lambda x$ say$(\hat{}P)(x)$

b. z(say) = $\lambda P\, \lambda x$ say$(\hat{}P(x))(x)$

c. $g(z(\text{say}) = \lambda P\, \lambda f\, \lambda x$ say$(\hat{}P(f)(x))(x)$

d. $g(z(\text{say})$ + (they like each other) =

= $\lambda f\, \lambda x$ say$(\hat{}\forall y(y\ \Pi\ RS(\textsc{id}(f))\ \&\ y\ \not\mathrm{Y}\ f(x))$ like$(y)(f(x)))(x)$

The distributor also undergoes g in order to pass up the paycheck function. Again, we ignore the internal structure of the matrix subject, translating it as the plural individual *lawyers*$_1$. It must lift to $\lambda P\, P($*lawyers*$_1)$ so that it, too, can pass up the paycheck function:

(6.39) a. $D = \lambda P_{<\mathrm{e,t}>}\, \lambda w\, (\forall x\ \Pi\ w) P(x)$

b. $g(D) = \lambda Q_{\mathrm{ee,et}}\, \lambda f_{\mathrm{ee}}\, \lambda w\, (\forall x\ \Pi\ w)\, Q(f)(x)$

c. $g(D)$ + (say they like each other) =

$\lambda f_{\mathrm{ee}}\, \lambda w\, (\forall x\ \Pi\ w)$ say$({}^\wedge\forall y(y\ \Pi\ RS(\textsc{id}(f))\ \&\ y \not\Pi f(x))$ like$(y)(f(x)))(x)$

d. l(lawyers$_1$) = $\lambda P\, P$(lawyers$_1$)

e. $g(l$(lawyers$_1$) = $\lambda P\, \lambda f\, P(f)$(lawyers$_1$)

f. (6.27a) =

The lawyers who represent John and Mary say they they like each other =

$\lambda f\ (\forall x\ \Pi$ lawyers$_1)$ say$({}^\wedge\forall y(y\ \Pi\ RS(\textsc{id}(f))\ \&\ y \not\Pi f(x))$ like$(y)(f(x)))(x)$

The function written as ID is the identity map on functions defined just on John and Mary's lawyers; hence the context must supply a function restricted to these individuals. In this case this will be the function mapping these lawyers to their clients.

6.4.3 What about non-dependent reciprocals?

The analysis as given expects to find a function immediately above the reciprocal predicate. But what of a simple reciprocal sentence like (6.40), which does not involve a dependent pronoun? In this case the immediate antecedent to the reciprocal is the NP *John and Mary*, not a pronoun.

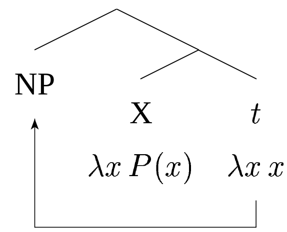
(6.40) John and Mary like each other.

In order to apply the analysis developed in the previous section it is necessary to lift a constituent of type $<$e$>$ to the identity function over itself, and then arrange for the distributor to pass this function to the reciprocal. Although this process renders the derivation of simple reciprocals more complex than the derivation of dependent ones, it is in the spirit of Montague's strategy of assigning the most complex type necessary to all instances of a semantic category. Given the translation we are after, there seems to be no way around this: if we start with an expression that expects a subject of type <e> and then type-shift to bind the pronoun, we end up with plain bound-pronoun semantics, as illustrated for the Heim et al. (1991b) account in section 6.3. Because the initial formula for the reciprocal expected a pronoun, the type-shifting operations treat the local antecedent as an ordinary singular bound variable; there is no way to compute the correct range argument from it. Thus we need to translate the reciprocal in such a way that it assumes a function subject, and either define all NPs as functions (again, following the Montagovian approach), or provide a mechanism for the creation of a suitable identity function when there is none. In this section I flesh out one method, based on a suggestion by Maribel Romero, that could accomplish the latter.

The essence of this solution is to use a trace left behind by raising the subject NP as the function we need. Recall that the trace of moving an NP of type <e> is treated as a variable of type <e> in the standard semantics (e.g., of Heim and Kratzer (1998)), and hence in the variable-free system it must be represented as a function of type <e,e> (and syntactic type $\mathrm{NP}^{\mathrm{NP}}$). Before we go on, it must be noted that one of Jacobson's ultimate goals is to develop a semantics for the direct interpretation of surface structures. Therefore traces, and especially the traces of LF movement, are not welcome in her system. However, nothing in the Variable-Free system itself rules them out, and Jacobson (1999a) incorporates them into her discussion of possible alternative implementations under the Variable-Free umbrella. I will employ traces here because they make a simple analysis possible, and because I have no personal objections to their use.

A trace, then, will be treated like a bound pronoun by both the standard semantics and the Variable-Free system. Consider the following example, constructed purely for purposes of illustration. (It might represent raising to subject of the object of an unaccusative verb X). Here *NP* has moved past some predicate X, leaving behind a trace t that is translated as the identity function.

(6.41)



This structure will be interpreted as usual: First X must undergo g in order to combine with t, then the result is applied to the NP on top. (The result of combining X with t is of type S$^{\mathrm{NP}}$, not S/NP; hence raising of the node NP must involve some type shifting to license the last operation).

(6.42) a. $g(X) = \lambda f\, \lambda y\, P(f(y))$

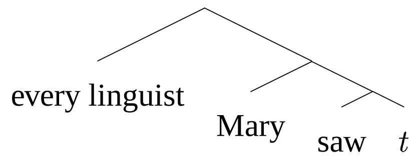
b. $g(X) + t = \lambda y\, P([\lambda x\, x](y)) = \lambda y\, P(y)$

c. $g(X)(t) + \text{NP} = P(\text{NP})$

In this example, the trace position is interpreted exactly as if there had been no movement. This is not always the case: in quantifier-raising constructions, the trace is interpreted as a variable (of type <e>) bound by the raised quantifier. For example, suppose that sentence (6.43) is interpreted by raising the object quantifier rather than by lifting its type; this derivation is slightly complicated by the fact that as usual, *Mary* must undergo l followed by g in order to pass up the open variable in the VP *saw t*. Once this is done, the quantifier will combine with its complement, binding its own trace as a variable that ranges over linguists.

(6.43) Mary saw every linguist.

(6.44)



(6.45) a. $t = \lambda x_{\mathrm{e}}\, x$ (as before)

b. every linguist = $\lambda P\, \forall x$ linguist$(x) \rightarrow P(x)$

c. $g(l(\text{Mary}))g(\text{saw})(t) = \lambda x$ saw$(x)(\text{Mary})$

d. (a) + (b) = $\forall x$ linguist$(x) \rightarrow$ saw$(x)(\text{Mary})$

Next, we must ensure that a subject NP always raises, leaving a trace, when it is the local antecedent of a reciprocal. Rather than invent a syntactic motivation such as Case assignment, I will assume that movement can be triggered in order to generate the proper semantic structure. The reciprocal VP *like each other* has the translation shown in (6.46), which is of type <ee, et>. Therefore it cannot combine directly with a constituent of type <e>, and would justify movement of the subject in order to repair the type mismatch; but to where?

(6.34c) like each other = $\lambda r_{\mathrm{ee}}\, \lambda x\, \forall y(y\ \Pi\ RS(r)\ \&\ y\ \not\Pi\ r(x))$ like$(y)(r(x))$

To answer this question, recall that the subject of reciprocals is (generally) distributively interpreted. In section 6.3 I adopted VP-adjoined distributors (type <et, et>), defined as in (6.21):

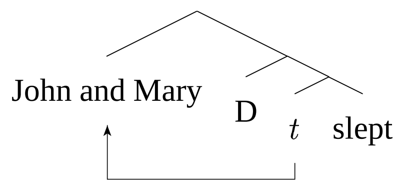
(6.21) D = $\lambda P_{<\mathrm{e,t}>}\, \lambda w\, (\forall x\ \Pi\ w)P(x)$

Now note that in order to obtain the correct semantics, the trace of the subject NP must appear below the distributor:

(6.46) [NP [D t [V each-other]]]

To bring it all together, then, I will assume that subjects are generated VP-internally, and distributors are adjoined to VP. It follows that the subject of a distributively interpreted sentence will always need to raise above the distributor, leaving behind a trace that will be bound by the distributor in step (c):

(6.47)



a. g(slept) = $\lambda f\ \lambda y$ slept$(f(y))$

b. t + g(slept) = λy slept$([\lambda x\ x](y))$ = λy slept(y)

c. D = $\lambda P_{\text{<e,t>}}\ \lambda w\ (\forall x\ \Pi\ w)\ P(x)$

d. D (g(slept)(t)) = $\lambda w\ (\forall x\ \Pi\ w)$ slept(x)

e. $J{\oplus}M$ (D slept) = $(\forall x\ \Pi\ J{\oplus}M)$slept$(x)$

A reciprocal predicate will work the same way:

(6.48) John and Mary like each other.

(6.49) a. (John and Mary) [D [t (like each other)]]

b. (like each other) = $\lambda r_{\text{ee}}\ \lambda x\ \forall y(y\ \Pi\ RS(r)\ \&\ y\ \not{\mathrm{I}}\ r(x))$ like$(y)(r(x))$ (= (6.34c))

c. t = $\lambda x\ x$

d. t + (like e.o.) = $\lambda x\ \forall y(y\ \Pi\ RS([\lambda x\ x]), y\ \not{\mathrm{I}}\ [\lambda x\ x](x))$ like$(y)([\lambda x\ x](x))$

= $\lambda x\ \forall y(y\ \Pi\ RS([\lambda x\ x]), y\ \not{\mathrm{I}}\ x)$ like$(y)(x)$

e. D (t like e.o.) = $\lambda w\ (\forall x\ \Pi\ w)\ \forall y(y\ \Pi\ RS([\lambda x\ x]), y\ \not{\mathrm{I}}\ x)$ like$(y)(x)$

f. J+M like e.o. = $(\forall x\ \Pi\ J{\oplus}M)\ \forall y(y\ \Pi\ RS([\lambda x\ x]), y\ \not{\mathrm{I}}\ x)$ like$(y)(x)$

Note that the trace, like pronouns, is treated as a restricted function. Its domain must be the set of all things that it will range over when used, namely, the sum of *John* and *Mary*.

In this way we allow our definition of reciprocals, developed with dependent reciprocals in mind, to be used with ordinary reciprocals as well.

The preceding account appeals to VP-internal subjects in order to provide a function between the reciprocal VP and the distributor. The assumption that the subject is base-generated lower than the adjunction position of the distributor has implications for the analysis of examples involving VP conjunction. Note that under the VP-internal subject analysis, VP conjunction must be explained as involving either across-the-board extraction of the subject (as in (6.50a)), or the conjunction of constituents smaller than the VP, perhaps V′, so that the position of the subject is not duplicated.

(6.50) a. John [$_{VP}$ [$_{VP}$ t ate] and [$_{VP}$ t slept]]

b. John [$_{VP}$ t [$_{V'}$ [$_{V'}$ ate] and [$_{V'}$ slept]]]

Now consider an example such as (6.51a), which involves the conjunction of a collective and a reciprocal VP. Because only one of the two VPs is interpreted distributively, such examples have been used as evidence that distributors should be adjoined to VP, not to the subject (see section 2.3.1.3, or Lasersohn (1995)). They are analyzed as involving conjunction of two VPs, only the second of which is distributive.[11] This is shown in (b). The distributor must appear within one of the conjoined constituents, else it would apply to both of them. But because we have placed the distributor higher than the underlying position of the subject, we necessarily have duplication of the subject position. Therefore only

[11]Heim et al. (1991b), who use NP-adjoined distributors, discuss the following example of mixed collective-distributive predication:

(i) They criticized each other's driving just after PRO colliding.

Their solution is to raise the distributor in the first clause only as far as its VP-internal subject, so that it does not block c-command of PRO by the subject at its matrix position:

(ii) They$_1$ [$_{VP}$ [e_1 D]$_2$ criticized e_2 each other's driving just after [PRO$_1$ colliding]

Whatever the merits of their analysis for this example, it does not help us with example (6.51a) unless we analyze it as not involving VP conjunction after all, and allow a PRO in the second VP.

the first of the above options, across-the-board extraction, can account for this example.

(6.51) a. They collided and blamed each other.

b. They [[$_{VP}$ $^{\text{COLL}}$collided] and [$_{VP}$ D [$_{VP}$ blamed each other]]]

Augmented with these assumptions, the proposed Variable-Free Semantics treatment of reciprocals can account for the full range of constructions we have considered: Single-clause reciprocals as well as dependent reciprocals whose antecedent can be an ordinary dependent pronoun, a paycheck-style dependent pronoun, or a larger NP containing a dependent pronoun.

6.5 Conclusions

This chapter has explored a new approach to the technical problem of how to extract the range argument of the reciprocal in a compositional way from its antecedent. In the standard semantics, Engdahl's (1986) analysis of paycheck pronouns represents the pronoun as a function-valued open variable, whose argument is saturated by another variable. In Jacobson's Variable-Free Semantics, the pronoun function is present but without a saturating argument. This makes it possible to state a translation for the reciprocal that allows the reciprocal direct access to the pronoun function, and consequently to its domain.

Abstracting away from all the technology, the solution I have proposed works as follows: The reciprocal expects as its local antecedent not a plural NP but a function, possibly the identity function, which is (normally) applied to the parts of some distributively interpreted NP. This function must necessarily come with a domain, from which the range set of the reciprocal is computed.

The Variable-Free framework also makes possible a near-trivial analysis of examples containing a possessive dependent pronoun: in the Variable-Free semantics, a full NP containing a pronoun behaves combinatorially just like a pronoun by itself; even the domain of the corresponding function is induced by the domain of the embedded pronoun. The same applies to any example in which the local antecedent of the reciprocal is a complex NP containing a dependent pronoun. However, the resulting reciprocal semantics *require* the antecedent of the reciprocal to be a function; this complicates the translation of the simplest, non-dependent reciprocal constructions. Therefore this analysis needed to be augmented with the system of VP-internal subjects and traces sketched in the last section. This part of the analysis involves some non-trivial (if not terribly controversial) syntactic claims, which I have doubtless been remiss in not defending by means of additional arguments.

In the interests of simplicity, the Variable-Free treatment of reciprocals developed in this chapter was based on the analysis of Heim et al. (1991b). In principle it should be embedded in a more recent analysis of reciprocity that deals appropriately with issues like the type of reciprocal relationship expressed. As discussed in chapter 6, my choice for a basic framework is Schwarzschild's (1996) treatment of distribution and reciprocals. In that chapter I presented an adaptation of the preliminary, non-Variable-Free version of my proposal within Schwarzschild's (1996) system. My final proposal, then, is an analysis that combines Schwarzschild's treatment with the proposal developed in the present chapter, which derives the range argument of the reciprocal from the range of its antecedent. Although the two systems are readily compatible, the adaptation of Schwarzschild's mechanics to the Variable-Free framework involves a level of notational and derivational complexity that I do not wish to tackle here. Therefore I will leave the details of the complete system as an exercise to the reader.

Chapter 7

Summary and further directions

This dissertation has explored the determination of reciprocal range from a variety of directions, focusing on the central question of how locality constraints on reciprocal interpretation can be reconciled with the compositional derivation of the reciprocal's range argument. This issue, it has long been known, is closely connected to the representation of dependent pronouns.

A dependent pronoun is broadly speaking a bound variable, and therefore logically singular. When a dependent pronoun appears as the antecedent of a reciprocal, it is possible to assign to the reciprocal a range argument that does not appear within the reciprocal's local binding domain. This dissertation has focused on the implications of such "long-distance" reciprocal constructions for the analysis of reciprocals, and of pronouns in general.

Chapter 2 showed that the usual, "scopal" analyses of reciprocals are at odds with the core generalization that arises from considering the full range of dependent reciprocal constructions: that the range argument of the reciprocal is always determined by its local antecedent. Accordingly, in chapter 4 I proposed that the range of the reciprocal is computed from the domain of its local antecedent. In an ideal solution, all required arguments of the reciprocal should be compositionally supplied to it; or better, since the reciprocal is subject to Principle A and overtly depends on only a single antecedent, a single semantic argument to the reciprocal would provide it with all the necessary information. An effective analysis for doing so can be expressed in the framework of Jacobson's Variable Free Semantics, as I showed in chapter 6.

Chapter 3 focused on the interpretation of dependent pronouns, and defended their analysis as "paycheck" pronouns bound by a distributive quantifier. In particular, I argued against an alternative cumulative analysis that would treat such pronouns as having a plural denotation. In chapter 4 the analysis of pronoun functions was enriched to include function domains.

In the course of exploring these issues we have considered some new or neglected constructions: most crucially to the argumentation, reciprocals whose antecedent is a dependent paycheck pronoun, or an NP containing a dependent pronoun. We also examined split-dependent pronouns, whose referent is the sum of a bound and a fixed part, and their behavior as antecedents of reciprocals (sections 3.6 and 4.6); finally, I have attempted to rescue from the realm of curiosity some constructions in which the reciprocal's antecedent is *not* interpreted distributively (see section 4.7).

As much as possible I have attempted to ensure that the mechanics of the proposed analyses reflect the strict locality requirements on the interpretation of the reciprocal's range argument. Consequently I have striven to minimize the role of pragmatics in the selection of the reciprocal's logical arguments. In particular, leaving the selection of the range argument to the pragmatics leaves open the possibility that, if we could only set up the right context, another source for the range argument could be found. In fact such scenarios are never instantiated: to the best of my knowledge, the range argument of a reciprocal is always based on its local antecedent. On the other hand, it is indisputable that the specific reciprocal relation that holds between the parts of the reciprocal's range does depend on the context, and can be left to the pragmatics as proposed by Schwarzschild (1996).

This worthy goal has not entirely been achieved: The proposed analysis does rely on the semantic component, not on the pragmatics, for the determination of the range argument. But the semantic translations I have proposed *could* allow the arguments of the reciprocal to be supplied from outside its local domain. The locality constraints on the interpretation of the reciprocal cannot be expressed in the ordinary semantic notation and must be stated, as usual, in the syntax.

The descriptive focus of this dissertation has been on reciprocals, and its formal focus on the interpretation of pronouns as functions with domains. Inevitably, I have set aside many more issues than I have tackled. To begin with, I have been concerned with a rather narrow range of reciprocal constructions. I only considered reciprocals embedded in tensed complements of a handful of verbs, and invariably with a definite (local or non-local) antecedent. Reciprocals as subjects of ECM constructions, or as adjuncts, were barely touched on. A broader variety of syntactic structures would have enriched the syntactic coverage of this treatment, while the narrow focus on definite NPs leaves unanswered the question of how the phenomena I have studied interact with quantificational and quasi-quantificational binders. Similarly, I have implicitly drawn a distinction between pronouns bound by distributors (which I dubbed "dependent pronouns") and pronouns bound by true quantifiers. While there are real differences in the binding properties of quantifiers and distributed NPs, I have not attempted to characterize them or to further explore their effects on bound pronoun interpretation.

Turning to reference functions, the analysis presented here raises (but regrettably does not attempt to answer) the question of which variables (in the Variable-Free system, which functions) should be analyzed as functions with domains: should that be just variables bound by a distributor, or all variables bound by a quantifier?

During the body of this dissertation I have abstracted away from the very interesting question of which reciprocal relationship (strong vs. weak reciprocity, and numberless variants) is expressed by a given use of the reciprocal. This is a question that does not appear germane to the approach to reciprocals that I have followed, but I suspect that advances in our understanding of this subject may turn out to have implications for the determination of reciprocal range. I have also only touched on the question of the proper distinctness condition between the potentially overlapping parts of the reciprocated-over range set.

Finally, this dissertation has dealt almost exclusively with English. While a cross-linguistic perspective might have enriched our insight on the interpretation of dependent reciprocals and pronouns, I can only acknowledge that I found English reciprocals to be more than sufficient challenge.

Bibliography

Asudeh, Ash. 1998. Anaphora and argument structure: Topics in the syntax and semantics of reflexives and reciprocals. Master's thesis, The University of Edinburgh, Centre for Cognitive Science.

Beck, Sigrid. 1999. Reciprocals and cumulation. In *Proceedings of SALT 9*, ed. Tanya Matthews and Devon Strolovitch, 16–33. Ithaca, NY: Cornell University.

Beck, Sigrid. 2000. Exceptions in relational plurals. Paper presented at SALT 10.

Brisson, Christine. 1998. Distributivity, maximality, and floating quantifiers. Doctoral Dissertation, Rutgers University.

Brisson, Christine. In preparation. Issues in plurality and collectivity. Ms., University of Pennsylvania.

Chomsky, Noam. 1981. *Lectures on government and binding*. Berlin: Mouton de Gruyter, seventh edition. Originally published by Foris. This edition published 1993.

Chomsky, Noam. 1986. *Barriers*. Number 13 in Linguistic Inquiry Monographs. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Kenneth Hale and Samuel J. Keyser, 1–52. Cambridge, MA: MIT Press.

Cooper, Robin. 1979. The interpretation of pronouns. *Syntax and Semantics* 10:61–92.

Dalrymple, Mary, Makoto Kanazawa, Sam Mchombo, and Stanley Peters. 1994. What do reciprocals mean? In *Proceedings of SALT 4*, 61–78.

Dayal, Veneeta. 1996. *Locality in wh-quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer.

Dimitriadis, Alexis. 1999a. Reciprocal interpretation with functional pronouns. In *Proceedings of the Twelfth Amsterdam Colloquium*, ed. Paul Dekker, 97–102. University of Amsterdam.

Dimitriadis, Alexis. 1999b. Reconciling dependent plurals with *each other*. In *Proceedings of SALT 9*, ed. Tanya Matthews and Devon Strolovitch. Ithaca, NY: Cornell University.

Engdahl, Elisabet. 1986. *Constituent questions: The syntax and semantics of questions with special reference to Swedish*. Number 27 in Studies in Linguistics and Philosophy. Dordrecht: D. Reidel.

von Fintel, Kai. 1994. Restrictions on quantifier domains. Doctoral Dissertation, University of Massachusetts.

Fox, Danny. 1995. Condition C effects in ACD. *MIT Working Papers in Linguistics* 27:105–119.

Gillon, Brendan. 1987. The readings of plural noun phrases in English. *Linguistics and Philosophy* 10:199–219.

Hausser, R. R. 1974. Quantification in an extended Montague grammar. Doctoral Dissertation, University of Texas at Austin.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation, University of Massachusetts at Amherst.

Heim, Irene. 1993. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. Technical Report SfS–07–93, Seminar für Sprachwissenschaft, Eberhard-Karls-Universität, Tübingen.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Heim, Irene, Howard Lasnik, and Robert May. 1991a. Reciprocity and plurality. *Linguistic Inquiry* 22:63–101.

Heim, Irene, Howard Lasnik, and Robert May. 1991b. On ‘reciprocal scope’. *Linguistic Inquiry* 22:173–192. Reply to Williams (1991).

Higginbotham, James. 1981. Reciprocal interpretation. *Journal of Linguistic Research* 1:97–117.

Higginbotham, James. 1983. Logical form, binding, and nominals. *Linguistic Inquiry* 14:395–420.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16:547–101.

Hintikka, Jaakko. 1999. No scope for scope? *Linguistics and Philosophy* 20:515–544.

Jacobson, Pauline. 1999a. Towards a variable-free semantics. *Linguistics and Philosophy* 22:117–184.

Jacobson, Pauline. 1999b. Paycheck pronouns, Bach-Peters sentences, and variable-free semantics. Ms., Brown University, August 1999.

Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Number 42 in Studies in Linguistics and Philosophy. Dordrecht: Kluwer.

Krifka, Manfred. 1992. Definite NPs aren’t quantifiers. *Linguistic Inquiry* 23:156–163.

Kroch, Anthony S. 1979. *The semantics of scope in English*. Number 18 in Outstanding Dissertations in Linguistics. New York: Garland. 1974 Doctoral Dissertation.

Langendoen, D. Terence. 1978. The logic of reciprocity. *Linguistic Inquiry* 9:177–197.

Lasersohn, Peter. 1988. A semantics for groups and events. Doctoral Dissertation, Ohio State University.

Lasersohn, Peter. 1989. On the readings of plural noun phrases. *Linguistic Inquiry* 20:130–134.

Lasersohn, Peter. 1995. *Plurality, conjunction, and events*. Dordrecht: Kluwer.

Lasersohn, Peter. 1998. Generalized distributivity operators. *Linguistics and Philosophy* 21:83–93.

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75:522–551.

Lebeaux, David. 1983. A distributional difference between reciprocals and reflexives. *Linguistic Inquiry* 14:723–730.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use and interpretation of language*, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 302–323. Berlin: de Gruyter.

Lønning, Jan Tore. 1997. Plurals and collectivity. In *Handbook of logic and language*, ed. Johan van Benthem and Alice ter Meulen, 1009–1053. Amsterdam: Elsevier.

Massey, G. 1976. Tom, Dick and Harry, and all the King's men. *American Philosophical Quarterly* 13:89–107.

Moltmann, Friederike. 1992. Reciprocals and *same/different*: Towards a semantic analysis. *Linguistics and Philosophy* 15:411–462.

Perry, John. 1993. The problem of the essential indexical. In *The problem of the essential indexical and other essays*, 33–52. Oxford University Press. First published in *Nous* 13 (1979), pp. 3–21.

Pritchett, Bradley L. 1990. A note on scope interaction with definite plural NPs. *Linguistic Inquiry* 21:646–654.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Roberts, Craige. 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12:683–721.

Romero, Maribel. 1999. Intensional choice functions for *which* phrases. In *Proceedings of SALT 9*, ed. Tanya Matthews and Devon Strolovitch. Ithaca, NY: Cornell University.

Sauerland, Uli. 1994. Codistributivity and reciprocals. Ms., MIT, December 1994.

Sauerland, Uli. 1995a. Codistributivity and reciprocals. In *Proceedings of WECOL '94*.

Sauerland, Uli. 1995b. The representation of reciprocals in grammar. In *Proceedings of ESCOL '94*, ed. Janet Fuller, Ho Han, and David Parkinson. DMLL Publications, Ithaca.

Sauerland, Uli. 1998. Plurals, derived predicates, and reciprocals. In *The interpretive tract*, ed. Uli Sauerland and Orin Percus, number 25 in MIT Working Papers in Linguistics, 177–204. MIT Press.

Scha, Remko J. H. 1984. Distributive, collective and cumulative quantification. In *Truth, Interpretation and Information: Selected Papers from the Third Amsterdam Colloquium*, ed. Jeroen Groenendijk, Theo M. V. Jannsen, and Matrin Stokhof, 131–158. Dordrecht: Foris.

Schwarzschild, Roger. 1992. Types of plural individuals. *Linguistics and Philosophy* 15:641–675.

Schwarzschild, Roger. 1996. *Pluralities*. Number 61 in Studies in Linguistics and Philosophy. Dordrecht: Kluwer.

Sharvit, Yael. 1999. Functional relative clauses. *Linguistics and Philosophy* 22:447–478.

Srivastav, Veneeta. 1992. Two types of universal terms in questions. In *Proceedings of NELS 22*.

Stalnaker, Robert C. 1999. *Context and content: Essays on intentionality in speech and thought*. Oxford Cognitive Science Series. Oxford University Press.

Sternefeld, Wolfgang. 1998. Reciprocity and cumulative predication. *Natural Language Semantics* 6:303–337.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1:235–284.

Williams, Edwin. 1986. A reassignment of the functions of LF. *Linguistic Inquiry* 17:265–299.

Williams, Edwin. 1991. Reciprocal scope. *Linguistic Inquiry* 22:159–173.