OUTSTANDING DISSERTATIONS IN
LINGUISTICS

Edited by
Jorge Hankamer
University of California, Santa Cruz

A Garland Series
Modal Subordination, Anaphora, and Distributivity

Craige Roberts

GARLAND PUBLISHING, INC.
New York • London
1990
Acknowledgements

This book is a revision of my 1987 Ph.D. dissertation written under Barbara H. Partee in the Linguistics Department at the University of Massachusetts at Amherst. I have added an introduction and revised Chapters 1, 3 and 5. However, the ideas developed here are substantially those found in the dissertation.

It was a privilege to study at the University of Massachusetts; there I had the opportunity to work with respected scholars from a variety of theoretical persuasions, and the cross-pollination of ideas from these different theories has resulted in the present work. I was especially fortunate to be able to work with Barbara Partee, whose scholarship and excellent teaching still inspire me. She was always generous with her time and ideas, and gave me valuable comments at every stage of my work.

The idea of modal subordination was first developed in a paper written in 1983, after participating with Hans Kamp in a seminar on Discourse Representation Theory. Charles Jones, a participant in that seminar, first pointed out to me the example of modal subordination which spurred my research. If a farmer has a donkey, he beats it. It usually dies of its wounds. The version of the theory which appears here as Chapter 1 is considerably revised and extended since its appearance in my dissertation. In much of its present form it appeared in Linguistics and Philosophy 12:683-721, 1989. In the course of its evolution, I benefited particularly from discussions with Barbara Partee, Angelika Kratzer, and Fred Landman, as well as Irene Heim, James Higginbotham, Roger Higgins, Nirit Kadmon, Hans Kamp, Stanley Peters, Peter Sell, and two anonymous reviewers for Linguistics and Philosophy.

My approach to the Binding Theory got out of this work on discourse anaphora, beginning with a paper written in Amherst in the spring of 1984 for a working group on logical form. Conversations with Emonn Bach, Jae-Wong Cloos, Roger Higgins, and Nirit Kadmon were helpful in developing this approach. David Pesetsky and Edwin Williams offered valuable suggestions and criticism of various versions of Chapter 2, and of the approach to the representation of quantifier scope proposed in Chapter 4.

The members of my dissertation committee, Barbara Partee, Emonn Bach, Hans Kamp, Angelika Kratzer, and David Pesetsky were very diligent in reading various drafts of Chapters 3-5, and the final product is a great deal better for their help. Barbara Partee was particularly helpful in the exploration of distributivity and the dependent plural, generously offering me a number of examples from her collection. Fred Landman was an unofficial committee member; he gave me detailed and extremely helpful comments on Chapter 3. I also benefited from discussions about distributivity with Godbeard Link and Mats Rooth.

Since 1987, I have had valuable discussions about the problems discussed here and/or comments on this manuscript from Stephen Levinson, Jan Tore Lattning, Stephen Neale, John Nerbonne, Alice ter Meulen, and Johan van Beekem. Miles Russell helped to prepare the present revision with a careful proofreading.

Thanks to the faculty, staff and students in the Linguistics Department at the University of Massachusetts at Amherst, especially Kathy Adams-czyk, Lynne Ballard, Nina Dabek, Alison Heutner, and Scott Myers, who made things so much easier.

I also want to give special thanks to my good friend and colleague Nirit Kadmon. She and I have worked together closely on a number of topics related to the work here, and I have benefited from her detailed comments on various drafts. I am also enormously grateful for her personal generosity and concern.

I am grateful to Colin Gordon for all the help, patience, and love he has shown me during the years I have worked on this book. He also undertook the huge task of preparing the manuscript of the dissertation, using the \TeX\typesetting program developed by Leslie Lamport as a special version of Donald Knuth's \TeX program.

And thanks to my parents, LaVerne Bennon Roberts and William H. Roberts, Jr., who have always supported my aspirations, and to Laura, for the paper airplanes.

My research at the University of Massachusetts was supported in part by the System Development Foundation, grant SDF \#650: "Research on the Formal Foundation of Semantics," awarded to Barbara H. Partee, as well as by the National Institutes of Health Development Grant \#18707: "Comprehending Sentences with Long Distance Dependencies," awarded to Lyn Frazier and Charles Clifton. Further research was conducted at the Center for the Study of Language and Information under a grant from the System Development Foundation, and also in the Linguistics Department of The Ohio State University. I gratefully acknowledge the support of these individuals and institutions, without which this work would not have been possible.
# Contents

Acknowledgements vi

Introduction 1

1 Modal Subordination 3
   1.1 Mood, modality, and modal subordination 4
   1.2 Modal subordination and constraints on anaphora 11
       1.2.1 The insertion approach 11
       1.2.2 The accommodation of the missing antecedent approach 16
   1.3 The interpretation of DBNs involving modality 26
       1.3.1 Subordination with non-epistemic modality and counterfactuals 27
       1.3.2 A formal theory of DBNs with modality 31
   1.4 Generalized subordination in discourse 34
   1.5 Conclusion 37

2 Anaphora, Coreference and the Binding Theory 39
   2.1 Problems in the theory of binding 41
       2.1.1 Syntactic problems 41
       2.1.2 Problems of interpretation 45
       2.1.3 Reinhart’s proposal 50
   2.2 An alternative proposal 52
       2.2.1 Reconstruction 59
       2.2.2 Disjunctive reference 62
       2.2.3 Crossover 65
   2.3 Indirect binding 67

3 Distributivity 77
   3.1 Previous approaches to the group-distributive distinction 80
       3.1.1 Distributivity as quantifier lowering/raising 80
       3.1.2 The distributive-group distinction as a function of predicate type 88
       3.1.3 Determiners and the distributive-collective-cumulative distinction 95
       3.1.4 Summary 113

4 The Representation of Scope 183
   4.1 The representation of quantifier scope 184
       4.1.1 Quantifying in and quantifier raising 185
       4.1.2 Scope indexing at S-Structure 198
   4.2 The operator 215
   4.3 Scope and distributivity in Discourse Representations 219
       4.3.1 Determiners and the mapping onto Discourse Representations 219
       4.3.2 Scope indexing, D, and the mapping from S-Structure to Discourse Representations 223

5 Remarks on Plural Anaphora 235

Bibliography 243
Introduction

The intimate relationship between anaphora and operator scope has long been recognized in the linguistic literature. Two assumptions are fundamental to much of the work on anaphora in the past two decades: The first is that the scope of quantifiers and other natural language operators may be restricted to some extent by the linguistic structures in which they occur; one instance of this generalization is that the scope of an operator is restricted to the sentence in which it occurs. The other fundamental assumption is that in order for a Noun Phrase (NP) to serve as antecedent for a pronoun, any operator which has scope over the NP must have scope over the pronoun as well. Given these assumptions and a theory of anaphora, one can use anaphora as a tool to explore both the behavior of operators and the nature of the linguistic structures in which they occur.

What follows is an exploration of certain aspects of the relationship between anaphora and operator scope, as reflected in English. It is intended to further the development of the theory of anaphora which takes the two assumptions above as fundamental. I begin by considering a class of apparent counterexamples to these assumptions above, involving the phenomenon I call modal subordination. This is illustrated in (1):

(1) (a) Anna should make a garden.
    (b) She could fill it with apple trees.

The apparent antecedent for it in (1b) is a garden in (1a). Yet on the most plausible reading of (1a), a garden is under the scope of the modal operator should. If we wish to hold to the second assumption, should in (1a) must have scope over it and hence over (1b); but this conflicts with the assumption that the scope of the modal is restricted to (1a). In Chapter 1, I develop a theory of modal subordination which permits us to retain both assumptions about the scope/anaphora relationship, and also accounts for some of the restrictions on the occurrence of this type of anaphora.

The account which I develop takes as its point of departure the theories of Kamp (1981) and Heim (1982) about anaphora in discourse. I propose an extension of Kamp's theory of Discourse Representation Theory to include modality, and use modals to reflect the notion of mood which is central to understanding
Chapter 1

Modal Subordination

Modal subordination is a phenomenon which stems from the relationships between propositions in a discourse. It is reflected in two kinds of related problems, which are illustrated in the discourses in (1) - (4):

(1) The birds will get hungry (this winter).
(2) (a) If Edna forgets to fill the birdfeeder, she will feel very bad.  
(b) The birds will get hungry.
(3) (a) If John bought a book, he’ll be home reading it by now.  
(b) # It’s a murder mystery.  
(4) (a) If John bought a book, he’ll be home reading it by now.  
(b) It’ll be a murder mystery.

The first type of problem involves the effect of context on the inferences which we may draw from a given sentence. In (1), for example, a proposition is asserted. But the same sentence in the context of the discourse in (2) seems more likely not to be asserted, but only to be asserted as following from the antecedent of (2a). The birds need not actually get hungry for the whole discourse to be true, so long as Edna is filling the birdfeeder so that the proposition expressed by the antecedent of the conditional isn’t true. The other type of problem involves anaphora, where the apparent antecedent is a quantified expression and the anaphor is not within its scope under standard assumptions about quantifier scope. In (3a) and (4a)


1Intended anaphoric relations are indicated by underlining antecedent and anaphor. The symbol “#” here and below indicates that the sentence is infelicitous in this context.
we see an example of the classical "doxey" sentences of Geach (1962), which have posed problems for theories of anaphora because the indefinite noun phrase a book serves as antecedent for a pronoun outside its scope, in the consequent. The discourse theories of Hans Kamp (1981) and Irene Heim (1982) provide an account of the felicity of anaphora in examples of this type. However, they do not account for several facts about the potential for anaphora in succeeding sentences. The discourse in (3) is indefinite. (b) is only interpretable as an independent assertion in this context, and the discourse is indiscernible because we have no available antecedent for the pronoun it. But the syntactically similar discourse (4) is felicitous. (4a) may be interpreted as a sort of continuation of the conditional in (a), as if it were coordinated with the consequent. I will show that the facts about inference and anaphora which are displayed in (2) and (4) are the consequences of a phenomenon I call modal subordination. Note that this notion of subordination is not the traditional syntactic notion. For example, in (4a), although the antecedent clause is syntactically subordinate, the main, or consequent clause is modally subordinate to it in discourse. This will become clear below.

I will develop a formal theory of examples such as (1)-(4) as an extension of Kamp’s (1981) Discourse Representation Theory. In Section 1.1, I will relate the notion of the mood of a sentence to a theory of modality in model-theoretic semantics and show how this is relevant for modal subordination. In Section 1.2, I will argue that in order to account for the anaphoric phenomena, we need a theory of discourse which provides discourse referents intermediate between syntactic noun phrases (NPs) and their model-theoretic interpretation. I will sketch informally the way in which I propose to account for modal subordination by extending Discourse Representation Theory to include modality. In Section 1.3, I will consider cases of modal subordination where the modality is not epistemic, and I will present a formal theory which accounts for both the epistemic and non-epistemic cases. In Section 1.4, I will discuss how we might extend the analysis of subordination in discourse to cover cases involving non-modal operators, and in Section 1.5, I will summarize my conclusions.

1.1 Mood, modality, and modal subordination

Before we can discuss the relation between mood and modality, we must consider what it means to make an assertion in a discourse. Following Robert Stalnaker (1978), I will characterize this notion in a possible worlds semantic framework. Because we as individuals are not omniscient and do not know everything about the world in which we live, we do not know which of the possible ways that things may be, or possible worlds, is the actual world. However, as participants in a conversation, we assume a set of propositions about the way the world is.

margin note 6: Heim (1982) discusses the relevance of Stalnaker’s theory of assertions for her theory of discourse. What is new here is the discussion of the relationship of mood and modality to what is asserted.

These may be introduced explicitly in the course of the conversation and mutually agreed upon, or they may be implicit presuppositions which it is assumed that all participants share. These explicit and implicit assumptions Stalnaker calls the common ground of the conversation, and this is enough to rule out quite a few possible worlds, those in which any of the propositions in the common ground are false. The larger the common ground in a given conversation, the smaller the set of possible worlds compatible with all the propositions presupposed, that is, the closer we come to being able to fully characterize the actual world. The set of possible worlds compatible with the common ground of a conversation is called the context set. These are the remaining candidates for the actual world. Given this framework, Stalnaker characterizes assertions as follows (1979, p.339):

To make an assertion is to reduce the context set in a particular way, provided that there are no objections from the other participants in the conversation. The particular way in which the context set is reduced is that all of the possible situations incompatible with what is said are eliminated. To put it in a slightly different way, the essential effect of an assertion is to change the presuppositions of the participants in the conversation by adding the content of what is asserted to what is presupposed. This effect is avoided only if the assertion is rejected.

So, every time we accept some assertion about the actual world we come closer to being able to completely characterize that world, and the context set of remaining possibilities becomes smaller. For example, suppose you and I are discussing the poet Lorine Niedecker. You mention that she was born near Lake Koshkonong in southern Wisconsin in 1903, a fact which I already know. This fact is then in our common ground, and the context set determined by that common ground contains no worlds in which Lorine Niedecker was not a poet, or in which she was born in Tulsa, Oklahoma. But if this is all the information we share about Lorine Niedecker, there will be worlds in the context set where she never left Lake Koshkonong, worlds where she left there at age 20 and rarely returned, worlds where she had children, worlds where she did not, and so forth. If I tell you that Lorine Niedecker left her birthplace for only a few years in the late forties and that she never had children, and if you accept what I say as true, then we add these propositions to our common ground. We now eliminate from the context set all worlds incompatible with this information.

This characterization of the notion of common ground and the discussion in Stalnaker’s work from which it is drawn tend to suggest that the common ground is epistemic in character and is shared by all participants. However, things are

margin note 8: For simplicity, I will assume throughout this chapter that the world which we attempt to characterize in discourse is always the actual world. However, it is often the case that we assert propositions to be true not of the actual world but of some fictional or fantasy world, as in a novel, a play, children’s play, etc.
somewhat more complex than this. In actual practice, there may be as many versions of the common ground of a given conversation as there are participants. This is because we typically have different ideas about what propositions are implicitly presupposed, as well as failing to communicate or understand properly those propositions which are explicitly presupposed or asserted. We may even deliberately mislead a hearer into thinking we share certain assumptions, so that the conversation proceeds on the basis of what one of us believes to be false premises. In any case, each of the participants in a conversation adopts a certain fiction which is crucial for effective communication: that the common ground as he or she knows it is in fact common to all the participants. Many restrictions on what constitutes felicitous discourse are designed to make sure that in crucial respects our common grounds do match. Some of these restrictions are rhetorical in nature, concerned with more or less conscious strategies for the presentation and flow of information and the development of argumentation; but some seem to be more deeply integrated into linguistic rule systems such as that governing anaphora. We will consider how such restrictions affect discourse anaphora in the following section.

Let us assume that Stalnaker's characterization of what it is to make an assertion is essentially correct. It remains to clarify when a particular utterance is an assertion. Even if we rule out consideration utterances which clearly function as questions, commands, or other non-assertive speech acts, not all declarative sentences uttered in a discourse are asserted to be true, as we have seen in (26). It is the mood of an utterance which tells us whether or not it is asserted. The sense of mood which interests us here is what Jespersen (1965) calls modal mood. This does not concern a grammatical feature of verbs, but rather describes a feature of sentence use: it relates to the speaker's commitment to the truth of a sentence in the actual world. If a speaker indicates by conventional means that a sentence or clause is to be interpreted as true in the actual world, we say that the sentence or clause was uttered in the factual mood. Utterances in the factual mood are asserted, in Stalnaker's sense. Sentences in the indicative grammatical mood, such as (1), are generally interpreted as factual where there is no context to suggest otherwise. But if a clause, such as the antecedent of a conditional, expresses a hypothetical assumption, or if there is otherwise some question about the actual truth of the clause, we say that it is uttered in a nonfactual mood. Nonfactual mood is expressed by a variety of conventional means. The subjunctive grammatical mood is one means; for example, I have been told (Roger Higgins, p.c.) that in German journalistic style the subjunctive may be used in main clauses to indicate that the proposition expressed is hearsay, and that the writer does not necessarily subscribe to its truth. Other, unrelated languages use similar morphological devices, for example Japanese (Karina Williamson, p.c.) and Finnish (Anne Vainikka, p.c.). Nonfactual mood may also be indicated by expressions like suppose that ..., or if ... then .... It may involve the use of modal auxiliaries like would or could, or adverbials like probably, supposedly, etc. And nonfactual mood may also be suggested by the sequence of tenses in a discourse.

I.1. MOOD, MODALITY, AND MODAL SUBORDINATION

below we will consider how this works in examples such as (2).

I propose that in a possible worlds semantic framework, mood should be interpreted in terms of modality. If we make a hypothetical supposition, in a nonfactual mood, we are not committing ourselves to its truth in the actual world. But for the purpose of exploring the consequences if that supposition were in fact true, we temporarily add it to our common ground. This temporarily eliminates some possible worlds in the original context set — those in which the hypothetical assumption is not true. Since we do not necessarily know whether our assumption holds true, the reduced context set which results may or may not include the actual world, that is it may or may not be realizable.  

In order to express this formally, I will adopt a theory of modality in natural language which has been proposed by Angelika Kratzer (1977,1979,1980,1981) and is compatible with Stalnaker's functional characterization of assertion. Kratzer points out that the force of modal operators such as necessarily, possibly, would, and the like in natural language is not absolute, as are the necessity and possibility operators in modal logic, but is relativised to a contextually determined set of propositions. In Kratzer (1980), there are two sets of propositions involved in such relativization, given by two functions which she calls the modal base, and the ordering source. In this section, I will simplify the informal discussion of Kratzer's theory by only considering the relativization of the force of modal operators to a single set of propositions, those in the common ground; for the examples under consideration, the common ground will play the role of the propositions given by Kratzer's modal base. Later, in Section 1.3, we will see more complex examples where the common ground and the set of propositions given by the modal base are not the same, and we will consider the importance of the ordering source in characterizing non-epistemic modality and counterfactuals. As an example of how the relativization of modal force works in Kratzer's theory, consider (6).

(6) Ella might lift that refrigerator.

Here, the modal force of might is that of possibility. If we translate this utterance into a sentence of a modal predicate calculus and then interpret it in the standard fashion, it means roughly, "there exists at least one member of the set of all possible worlds in which Ella lifts that refrigerator". Now, if we assume that this set of possible worlds contains not just those situations which we regard as reasonable in the actual world, but all possibilities, including, for example, a world in which ordinary women such as Ella easily lift two ton trucks, the utterance is then trivially true. It would come as no surprise that in that world such women also lift refrigerators. But this flies in the face of our intuitions about the proposition, which seems much stronger than this.

*This is an extension of Kratzer's (1980) usage of realizable.
CHAPTER I. MODAL SUBORDINATION

In ordinary conversation, a hearer is likely to assume that the person uttering (6) is making a claim in view of what is physically possible and normal in the actual world. Following Kratzer, we will relativize the modal force of (6) to the set of possible worlds where the actual facts about this sort of thing are all true. In general, speakers tend to assume that the common ground of a given discourse includes a set of propositions which we might characterize as 'what is physically possible' (or, more properly in a doxastic common ground, 'what we in common assume about what is physically possible'). Thus, in a lay conversation, speakers do not generally assume that the common ground includes propositions about quarks or the ultimate nature of light. This determines a context set which does not contain all possible worlds, and in particular, won’t contain worlds where ordinary women easily lift two-ton trucks. Relativizing the modal force of (6) to this context set, the proposition has truth conditions which are closer to our intuitions about its meaning: it will be true in case there is a possible world in which the actual facts about human strength, gravity, etc. are true, and in that world, Ella lifts that refrigerator.

This relativization of the modal force to a narrowed context set is very similar to domain selection in quantification. For example, if a speaker says "Everyone seems happy", he doesn’t usually mean that absolutely all individuals whatsoever seem happy, but only those whose domain is suitably narrowed: for example, ‘those individuals in this room’, or the like. We often are not explicit about how to select this domain, assuming that our hearers will guess what we intend from the context. Similarly, we often assume that our hearers will understand how we intend to restrict the context set of possible worlds over which modal operators range. This was the case in the refrigerator example just discussed.

Sometimes, though, a speaker is more explicit about at least some of the propositions which she wants the hearer to add to their shared common ground. Consider again example (2):

1.

(a) If Edna forgets to fill the birdfeeder, she will feel very bad.
(b) The birds will get hungry.

It has been noted by several authors (including Kratzer (1988) and Heim (1982)) that conditional sentences often have modal force, either explicit, as when the consequent contains a modal auxiliary such as might, or implicit, in which case the modal force is that of necessity. The antecedent clause is hypothetically added to the common ground, narrowing the context set against which the modal force is evaluated. (2a) is an indicative conditional without overt modal operators. Because of the future tense in the consequent, we are not tempted to interpret the present tense in the antecedent as a generic present, quantifying over times when

9Actually, if-classes may serve to modify various operators, not just modals. Farkas & Sagisaka (1983) point out examples like the following:

(i) Mary is usually friendly, if she’s not in a hurry.
(ii) All cats like to use scratching posts, if they haven’t had their claws removed.

In (i), the operator is a temporal adverb of quantification; in (ii), a universal quantifier. Often, conditionals without any explicit operator seem to have a modal flavor. However, this is not always the case. Steve Berman (p.c.) points out that the following example seems to involve quantification over times or events:

(iii) If Edna forgets to fill the birdfeeder, she feels bad. The birds get hungry.

Such examples may involve situations, however these are to be defined. See Section 1.4 for further discussion. In all the cases which interest us in this section, treating the if-clause as modifier of a modal operator presents no problem.
true that the birds have gotten hungry.

However, the same sentence occurs in a different context in (2). The reference time of the nonfactual antecedent of (2a) is most likely that of some future time which the participants have been discussing, a time when the feeder is supposed to be filled. The consequent is then interpreted as future with respect to its reference time, the latter given by the antecedent, in keeping with the plausible implication that the state referred to in the consequent will result from the event referred to in the antecedent. Now consider the interpretation of (2b) in this context. Since we have already been talking about Edna, and since Gricean conversational principles (Grice 1967) require that we generally seek to determine the relevance of a proposition to its context in discourse, we consider how (2b) may be related to (2a). In this case, the sequence of tenses and the plausibility of narrative continuity suggest that the reference time of (2b) is given by (2a); more precisely, that (2b) has the same reference time as the consequent of (2a), that which is given by the antecedent.

However, note that although the entire conditional in (2a) is factual (that is, its truth is a condition on membership in the context set of worlds which are candidates for actuality), neither the antecedent nor the consequent by themselves are factual. This means that the event denoted by the antecedent may not occur in the actual world. We are likely to view (2b) as a consequence of the antecedent (and possibly the consequent of (2a) as a consequence of (2b)). But we don’t have actual effects of non-actual causes. That is, since the consequent of (2a) and (2b) take the antecedent as giving their reference time, they must be non-factual as well. In terms of the account I am proposing, this means that although (2b) contains no overt modal operators, it is to be interpreted against a possibly non-realistic common ground which includes the antecedent of the preceding conditional. If we do this, we get an interpretation like ‘If Edna forgets to fill the bird feeder, she will feel very bad and the birds will get hungry’. This seems to give the right truth conditions.

In this informal discussion of the relationship of mood to reference time, there are a number of important questions which remain unexamined. For example, I have not discussed here what I mean formally by reference time. Is it an interval or an event? Why does taking a non-factual clause α as giving the reference time for a clause β entail that β is non-factual as well? In the formal theory presented in Section 1.3, I will have nothing to say about times or the temporal relations between sentences. However, I believe that this type of example may be used to argue for a particular kind of theory of the relationship between temporal reference and subordination in discourse, and this will be one of the central arguments of Roberts’ forthcoming work. What is important here is that if we take (2b) as interpreted against a partially nonfactual common ground which includes the proposition denoted by the antecedent of (2a), we can explain the inference problem discussed in the introduction.

There is another way of interpreting (2b) in this context, a factual interpretation where the future time referred to by its tense is existentially quantified:

1.2. CONSTRAINTS ON ANAPHORA

There is some future time when the birds will get hungry. This is an assertion, factual in mood. We take it to be a fact of life that the birds will get hungry in any case. On this interpretation, (2b) is only indirectly relevant to the utterance of the preceding conditional, and I take it that this is why the nonfactual interpretation is preferred, since we generally seem to prefer the most relevant interpretation of an utterance with respect to its immediately preceding context.

This discussion of examples (1) and (2) suggests that the notion of mood in conjunction with a theory of modality along the lines suggested by Kratzer can provide an account of the inferential properties of discourses with intersegmental modal subordination. Now we will turn to consider the relevance of anaphora for a theory of modal subordination.

1.2 Modal subordination and constraints on anaphora

1.2.1 The insertion approach

To account for the facts about anaphora illustrated in examples (3) and (4), the Stalnaker-Kratzer approach I have just sketched will not alone suffice. This is because this approach by itself has no provision for describing structural relations between NPs above the sentential level; hence, we do not have the means to discuss formal constraints on anaphora in discourse. Irene Heim’s File Change Semantics (1993) and Hans Kamp’s Discourse Representation Theory (1981) were initially developed to deal with problems in anaphora, and extending them to include the theory of modality I have just described will provide us with the basis of a theory of anaphora in modal subordination. In what follows, I will use Discourse Representation Structures (DRSs) to illustrate the theory; however, a very similar theory may be developed using Heim’s File.

Both discourse theories utilize variable-like discourse referents to serve as intermediate representations of syntactic noun phrases on the DRS (or File) level.6

For Kamp, the relative location of two discourse referents in the DRS determines whether one of these may serve as antecedent to the other.7 If a clause is in a factual mood, it is mapped onto the top level of the DRS. Consider the DRS for the simple sentence John bought a book in (7):

6Other researchers before Kamp and Heim had realized the need for something intermediate between syntactic NPs and their real world denotations. For example, see Karttunen (1976), from whom the term discourse referent is borrowed, and Webber (1978).

7Chierchia & Roath (1984) point out that this is really a matter of the scope of the operators involved in the interpretation of the DRSs and not of configurationality. However, we can ignore this point here.
Kamp proposes an algorithm for mapping from sentential syntactic structures onto DRSes. In Chapter 1 I proposed a different approach, which permits us to retain the insights of the Binding theory of the Government and Binding framework of Chomsky (1981) and associates. For our purposes here, the important feature which these approaches share is that each NP in a discourse will be mapped onto a discourse referent in the DRS. Both the content of the original NP and the content of anything predicated of that NP in the sentence are then entered as conditions on its discourse referent. In the case of (7), each NP in the original sentence is correlated with a distinct discourse referent, John with x, and a book with y. The proper name John and the common noun book have become conditions on the discourse referents correlated with their respective NPs, and a further condition specifies that the relation bought holds between x and y.

The resulting DRS is still an uninterpreted formal structure, logically syntactic in Carnap’s sense. To interpret this structure, we use another algorithm to embed it in a truth conditional model. Formally, an embedding is a function from discourse referents onto individuals in the model, such that the individual which a given discourse referent r is mapped onto displays each property corresponding to a condition on r. A DRS, and thus, indirectly, the discourse which induced it, is true in a model iff there is an embedding of the DRS in the model. Note the meta-language existential quantification over embeddings. Although indefinite and definite NPs aren’t treated as inherently quantificational in this theory, this meta-language quantification has the truth conditional effect of existentially quantifying over all of the discourse referents on the top level of a representation. (7) may be interpreted in a model as asserting that there is an individual to which x may be mapped, John, and an individual to which y may be mapped, which is a book, such that the individual corresponding to x bought the individual corresponding to y.

Universally quantified NPs induce a more complex DRS. Compare Kamp’s representation for the generic interpretation of sentences such as (8), in (9):

\begin{center}
<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>John(x)</td>
<td>book(y)</td>
</tr>
<tr>
<td>bought(x, y)</td>
<td></td>
</tr>
</tbody>
</table>
\end{center}

Here, the symbol between the antecedent and consequent boxes is reminiscent of the symbol for the material conditional. The lefthand subordinate box contains the representation of all the material in the common noun of the subject, here the noun farmer and the relative clause which modifies it, including the representation for the NP a donkey within that relative clause. The righthand subordinate box represents the material predicated of that subject. In such a representation, the righthand box is said to be subordinate to the lefthand box. When we introduce a discourse referent for a pronoun, such as w in the righthand box in (9), we must indicate an antecedent for the pronoun, or else the resulting DRS will be ill-formed. To find an antecedent, we locate an accessible discourse referent, that is, one which is on the same or a superordinate level of structure as the pronoun’s discourse referent. In (9), the discourse referent for the pronoun, w, may take any accessible discourse referent as its antecedent; here, we have selected the discourse referent for a donkey, y. This is symbolized by equating w with y. The extensional embedding conditions for such a representation specify that for any embedding of the antecedent box there must be an extension of that embedding to provide an embedding of the consequent box as well. Thus, any farmer/donkey pair which stands in the own relation must stand in the beat relation as well. Again, this metalanguage universal quantification explains the seemingly universal force of the indefinite in the relative clause without treating that NP itself as inherently universally quantified.

If a clause is in a non-factual mood, the common ground against which it is interpreted is not realistic. Entities which are introduced with an indefinite NP in that section of discourse may not actually exist. So the clause is mapped onto a subordinate level of the DRS, reminiscent of the way in which hypothetical assumptions are indented in a proof. Consider the sentence which occurs in (3a) and again in (4a):

\begin{center}
(3) (a) If John bought a book he’ll be home reading it by now.
\end{center}

This is another example of modal subordination in a conditional sentence, as with example (2a), but here anaphora is involved. As before, although the entire conditional is asserted, neither the antecedent nor the consequent is factual in mood. Because the antecedent is not factual, it must be entered into the DRS on
a level subordinate to the top level, as in the DRS in (10).\footnote{Kamp does not use the symbol $\Box$ in his representation of sentences like (3), using instead the symbol $\Rightarrow$, as in (9). He doesn't consider modality in his account.}

(10)

\[
\begin{array}{ccc}
  x & y \\
  \text{John}(x) & \text{book}(y) & \Box \\
  \text{bought}(x, y) & \text{reading}(x, w) & \text{z} = x \\
  & & \text{w} = y \\
\end{array}
\]

Here, the representations of the antecedent and consequent are each in a box which is subordinate to the top level. The lefthand box, representing the antecedent, is as in (7). The righthand box represents the consequent; as in the example with a universal quantifier in (9), the righthand box, that representing the consequent, is subordinate to the lefthand, or antecedent, box. (Note again how this notion of subordination differs from the syntactic one.) We enter discourse referents for the pronouns in the consequent of (3a) on the appropriate level of the DRS, \(z\) for he and \(w\) for \(it\) in the righthand box. These may then take the accessible discourse referents for \(John\) and a \(book\), \(x\) and \(y\), as their antecedents.

The symbol $\Box$ in (10) is mnemonic for the necessity operator. In interpreting the DRS, the antecedent box will serve as a restriction on the necessity operator, in much the same way as in Kratzer’s theory. In terms of the formal theory in Section 1.3, its truth conditions can be paraphrased as ‘For all worlds in which there is an individual \(z\), and a \(book\) \(y\), and \(x\) bought \(y\), then \(z(=x)\) is reading \(w(=y)\).’ As an assertion, this utterance instructs the participants in the discourse to reduce the context set as follows: ‘Consider all worlds in the present context set. (These are all candidates for the actual world. That is, they are all worlds in which all the presuppositions in our common ground so far are true.) Now consider only those worlds in that set where there is an individual \(x\), and there is a \(book\), \(y\), and \(x\) bought \(y\). In each of those worlds, you should find that there is an individual \(z(=x)\) and an individual \(w(=y)\) and that \(z\) is reading \(w\). If not, discard that world from the context set.’ We see that the notion of modal subordination which we expressed using Kratzer’s theory of modality in the birdfeeder example translates readily into Kamp’s configurational subordination. Just as in that example, after I have modified the original context set, containing candidates for the actual world, by the hypothetical addition of the antecedent of the conditional to the common ground, I no longer know whether the resulting context set contains the actual world. We haven’t asserted the antecedent, but only assumed it temporarily. So we have only temporarily assumed the existence of the entities referred to in the antecedent. Since the consequent is subordinate to the antecedent, we may continue to assume the existence of those referents, and

1.2. CONSTRAINTS ON ANAPHORA

their discourse referents may serve as antecedents for anaphors in the consequent.

Now consider (3b):

(3) (a) If John bought a book, he’ll be home reading it by now.
(b) # It’s a murder mystery.

There is no overt modal in (3b), and no plausible dependency relation between the reference time of its simple present tense and the conditional present of the consequent of (3a). Hence, we are not tempted to interpret the two clauses as in the same, nonfactual mood, and there is no evidence of modal subordination. Further, a conditional such as (3a) is only felicitous where we do not know the truth of its antecedent, and in this case, this entails not knowing whether John bought a book. But (3b) seems to be about some actual thing which is a murder mystery. Since (3b) appears to be in the factual mood, we enter it on the top level of the DRS, assigning a discourse referent \(r\) to the pronoun \(it\), and adding the condition \(\text{murder-mystery}(r)\), as shown in (11):

(11)

\[
\begin{array}{ccc}
  x & y & r \\
  \text{John}(x) & \text{book}(y) & \Box \\
  \text{bought}(x, y) & \text{reading}(x, w) & \text{z} = x \\
  & & \text{w} = y \\
\end{array}
\]

But now the discourse referent for a \(book\), \(y\), is in a box which is subordinate to \(r\), and so \(y\) is not an accessible antecedent for \(r\). The discourse is not felicitous unless there is another discourse referent in prior discourse which would be a plausible antecedent for the pronoun, or where the pronoun is deictic.\footnote{These two cases may be the same from the point of view of discourse theory. For some discussion, see Heim (1982).}

(4b), on the other hand, contains the same modal auxiliary \(\text{will}\) as the consequent of the preceding conditional sentence, and it is readily interpreted as an extension of the nonfactual mood:
CHAPTER 1. MODAL SUBORDINATION

(4) (a) If John bought a book, he’ll be home reading it by now.
(b) It’ll be a murder mystery.

One way to represent the resulting modal subordination would be to simply add the representation of (4b) to the consequent box of the DRS (10), as in (12):

(12)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>John(x)</td>
<td>book(y)</td>
</tr>
<tr>
<td>bought(x, y)</td>
<td></td>
</tr>
</tbody>
</table>

Here, a new discourse referent $r$ has been added to the consequent of (10), along with the condition murder-mystery on $r$, and $r$ has been equated with the accessible discourse referent $y$. Let us call this way of representing modal subordination the insertion approach. In this example, the result gives us the correct truth conditions. (12) would be interpreted as 'In all worlds where there is a book which the individual named John bought, you will find that John is reading the book and that it is a murder mystery.' As an assertion, the conditional is an instruction to remove from the context set determined by the prior common ground any worlds in which the antecedent is true and the consequent false.

1.2. CONSTRAINTS ON ANAPHORA

In (13a), might leads us to interpret the material in its scope as nonfactual. Uttering this sentence does not commit the speaker to the existence of an event in the actual world. Rather, it asserts that among the candidates for the way things are, the possible worlds in the present context set, there is at least one where a thief breaks into the house. Just as in Kratzer’s theory of modality, might will be interpreted as the possibility operator in modal logic, relativized to the context set. The DRS of (13a) is shown in (14):

(14)

The modal might is translated by the diamond operator on the left. The nonfactuality of the remainder of (13a) is expressed by entering its representation on a subordinate level of the DRS, the box on the right in (14). The modal force here is possibility, indicated in the DRS by the diamond operator. We might utter (13) in a situation where all the participants in the discourse already have certain assumptions, such as ‘Given that there has been a lot of theft in this neighborhood’, and ‘Given that this house has poor security’. These propositions in our common ground, rather than propositions denoted by any nonfactual clauses, might alone serve to relativize the force of the modal. Because the hypothetical common ground is realistic in such a case, I have not included a subordinate box to the left of the diamond operator, though in other contexts this might be necessary, as in “If you left a Mercedes parked out front, a thief might break into the house”.

In this example, I think the most natural sense of possibility is that of future possibility. An adequate embedding algorithm for this DRS will require us to examine each world in the context set to determine whether among the possible futures which branch out from the present moment in that world there exists at least one in which the box on the right in (14) may be truthfully embedded. If this is so, the world is retained in the context set after the utterance of (13a). If not, it is removed. Informally, its truth conditions might be paraphrased, ‘There exists some possible future in which a thief breaks into the house’.

Now, if we try to represent (13b) by inserting material into the righthand box in (14), as in (15), we will get the wrong truth conditions:

(13) (a) A thief might break into the house.
(b) He would take the silver.
Here, I have added a discourse referent, \( y \) for he, and the appropriate condition \( \text{take-silver on} \ y \), equating \( y \) with the accessible antecedent \( x \). The model theoretic interpretation of this DRS could be paraphrased, ‘It’s possible that a thief will break into the house and take the silver.’ But this is not our understanding of (13). In uttering this discourse, I don’t simply assert that a thief \( \text{might} \) take the silver. I’m saying something stronger: ‘It’s possible that a thief will break into the house, and if he \( \text{does} \), he will undoubtedly take the silver.’ The problem, of course, is that (13b) has a different modal force than (13a). \( \text{Might} \) in (13a) has the force of possibility; whereas, \( \text{would} \) in (13b) has the force of necessity. Instead of inserting the non-factual (13b) under the scope of the possibility operator in (13a), we must treat the modal auxiliary \( \text{would} \) as indicating that (13b) is a modally subordinated clause which is, like the consequent of a conditional, in need of an antecedent. The approach I suggest, which I will call the accommodation of the missing antecedent approach to modal subordination, is the pragmatic accommodation of a contextually given hypothetical common ground to be the antecedent of the modally subordinated clause. I use the term accommodation in an extension of the David Lewis’ (1979) sense, where to accommodate a presupposition is basically to add it to our common ground because without that presupposition, we cannot assign a truth value to an utterance, i.e., cannot make sense of it. Here is Lewis’ definition (p.340):

If at time \( t \) something is said that requires presupposition \( P \) to be acceptable, and if \( P \) is not presupposed just before \( t \), then — \( \text{ceteris paribus} \) and within certain limits — presupposition \( P \) comes into existence at \( t \).

Stalnaker (1979) identifies the set of propositions in the common ground of a discourse as the presuppositions of that discourse. We do not always introduce such presuppositions explicitly. Rather, we often assume that we share certain knowledge about the world with other members of the discourse, that is, we presuppose the propositions expressing that knowledge. In discussing (13), for example, I said that it would be felicitously uttered where the participants in the discourse already had certain assumptions about the actual world which made it probable. But these assumptions might not have been explicitly introduced into the discourse. Perhaps we’ve both lived in this neighborhood all our lives. I know there are thieves out there and you know it too, so we don’t need to say it. It is presupposed. That is, in a Kratzer-type theory where we relativize the force of a natural language modal element to some context set, the propositions which determine that context set may be implicit in whole or in part. So, when we encounter the modal \( \text{would} \) in (13b), we may assume that certain relevant assumptions should be added, at least hypothetically, to our common ground. To illustrate how accommodation works in modal subordination, consider the DRS for (13) in (16):

The upper portion of (16) contains the representation of (13a) which we saw in (14). The diamond possibility operator is intended to have scope only over this top box. The lower portion contains the representation for (13b). The necessary modal force of \( \text{would} \) induces the form for the representation of necessity which we have seen above. (13b) itself is represented in the righthand, or consequent box, while in the lefthand box we have accommodated the contextually available representation of the proposition a thief \( \text{breaks in} \) as a hypothetical common ground, narrowing the context set over which the necessary force of \( \text{would} \) will range. Notice that this accommodation is very naturally licensed by our assumptions of the relevance of (13b) to its context. This representation now gives the correct truth conditions when interpreted in a model. That is, something like ‘Given what we already know in common about the actual world, it is possible that a thief will break into the house. In all such worlds where a thief breaks into the house, he takes the silver.’
In general, then, the antecedent of conditional sentences serves as an explicit hypothetical addition to the common ground against which the consequent is to be evaluated. In sentences which are not conditional in form, modal subordination involves the pragmatic accommodation of a contextually salient proposition (or propositions) to serve as antecedent for the nonfactual clause.

I propose that we generalize this treatment of modal subordination via accommodation to examples such as (4), replacing the representation in (12), involving insertion, with that shown in (17):

\[
\begin{array}{|c|c|}
\hline
x & y \\
\hline
\text{John}(x) & \text{book}(y) \\
\hline
\text{bought}(x, y) & \text{reading}(z, w) \\
\text{□} & z = x \\
& w = y \\
\hline
\end{array}
\]

Here we have taken the preceding antecedent box as our accommodated hypothetical common ground for the representation of (4b). The interpretation of (17) might be paraphrased, 'In all worlds where there is a book which the individual named John bought, you will find that John is reading the book. And in every world in which there is a book which the individual named John bought, the book is a murder mystery'. One motivation for treating (4) in this way is that it gives a more compositional account of the mapping from sentences to DRSes, in that here the modal auxiliary in (4b) is treated as itself inducing a modal operator, the lower instance of □. An open question is whether we should treat modally subordinate sentences without overt modal operators, such as (2b), as inserted into the prior representation of material under the scope of a modal, such as the representation of (2a), or as containing an implicit modal operator themselves, for which we then accommodate relativizing material as we have done in (17). In what follows, I will assume for simplicity that all the examples of modal subordination involve accommodation.

There is one worry about the use of accommodation: It is a very powerful device. What prevents us from simply accommodating an appropriate common ground, including a potential anaphoric antecedent, in every case where we have a pronoun in discourse with no apparent antecedent. If this were possible, we would have no account of the infelicity of examples such as (3b). We seem to need constraints on the power of accommodation. One which is already clear from prior discussion is that modal subordination, and thus the accommodation which it triggers, requires nonfactual mood. Further, it must be plausible that the modally subordinate utterance has a hypothetical common ground suggested by the immediately preceding context. The examples which we have examined so far occur after a conditional or contain an explicit modal operator such as \textit{would} or \textit{might} to trigger the subordination. Another type of example involves \textit{or}, the disjunction operator. (18) is due to Barbara Partee:

\[
\begin{array}{|c|c|}
\hline
x & y \\
\hline
\text{John}(x) & \text{book}(y) \\
\hline
\text{bought}(x, y) & \text{□} \\
& \text{murder-mystery}(r) \\
\end{array}
\]

(18)

Either there's no bathroom in this house or it's in a funny place.

Here we find no overt modal in the second conjunct, yet the quantified noun phrase \textit{no bathroom} appears to serve as a sort of antecedent for a pronoun, \textit{it}, which is outside its scope under standard assumptions about quantifier scope.

On pragmatic grounds, we may assume that neither disjunct of a disjunction is asserted, and hence that both are nonfactual. We have seen above that any sentence uttered in a nonfactual mood may justify the accommodation of a hypothetical common ground. Further, we have also noted that relevance to context often dictates the choice of such an accommodated common ground. We often take the two disjuncts in a propositional disjunction to be alternative answers to the same topic of discussion. In (18), we may naturally assume that that topic is whether there is a bathroom in the house. The first disjunct entertains a negative answer to that question, so it seems perfectly natural to assume that the second disjunct pertains to the possibility of a positive answer to that same question. Thus, the accommodation of the portion of the representation of the first disjunct which is under the scope of the negation operator may be seen as the most natural means of providing an antecedent for the second disjunct, and hence for the pronoun \textit{it} within it.\footnote{This is very similar to a sentence in Evans (1977):}

\[
(1) \quad \text{Either John does not own a donkey, or he keeps it very quiet.}
\]

\footnote{Higginbotham (p.c.) also points out that this procedure can be iterated for multiple disjunction operations, as in (1):}
1.2. Constraints on Anaphora

(p.c.) points out that though (20) is logically equivalent to (18), (20) does not appear to license the anaphora we find in (18):

(20)  

Either every bathroom does not belong to this house, or it’s in a funny place.

This follows from the general algorithm for mapping from sentences with universally quantified NPs onto DRSes, as we saw, e.g., in (8), (9) above. Using this algorithm, the first disjunct in (20) would map onto the DRS in (21).

(21)  

If we then represent the second disjunct as we have done for (18) in (19), we would derive (22):

(22)  

In (22), (21) refers to the entire DRS in (21) above. Here, the discourse referent for every bathroom, \( x \), is in a subordinate box within the sub-DRS (21), which is itself under the scope of a negation, so \( x \) is not an accessible antecedent to the discourse referent for it, \( y \); the box in which \( y \) occurs is only subordinate to the top level of the DRS (21). That is, the more complex logical structure of the first disjunct does not provide us with an appropriate representation to be accommodated as antecedent of the second disjunct, and so anaphora seems infelicitous.

The derivation of the DRS for (18) in (19) is not intended to be algorithmic. That is, it is not always the case that where we have disjunction we accommodated...
date the negation of the first disjunct as antecedent for the second. For example, consider Steve Berman's (p.c.) (23):

(23) Either there's a bathroom on the first floor, or it's on the second floor.

Here, accommodating the negation of the first disjunct would not achieve the desired results. The discourse referent for a bathroom would be under the scope of a negation operator, and hence not accessible to it. However, (23) seems most felicitous when there is no intonational pitch accent on bathroom. (Compare this with (18), which always has a pitch accent on bathroom, and may even have a phrase boundary after it.) This may be taken as a signal that the speaker presupposes there is a bathroom, and it is this conventionally indicated presupposition which licenses the hearer to introduce a representation for there's a bathroom on the highest level of the DRS, providing a discourse antecedent for it. I assume that relevant contextually supplied and conversationally implicated material, as well as accommodated material, may be introduced into the same DRS as the explicit text. See Kadmon (1987), for independent arguments to this effect.

The contrast between (18) and (20) points to an important constraint on the type of accommodation we have been using when anaphora is involved: it requires the explicit prior representation of potential antecedent discourse referents. We may not simply infer their existence. Heim (1982, Chapter III) discusses the accommodation of antecedents for definite noun phrases and shows that it is constrained by the requirement that new file cards introduced under this type of accommodation must be cross-referenced to some pre-existing file card. But she claims in Section 5.3 that antecedents for pronouns cannot generally be accommodated in this fashion, due to their relative lack of descriptive content. The examples under discussion here provide further evidence for this claim. In these cases, we do not accommodate antecedents for pronouns directly; rather, it is the independently required accommodation of appropriate hypothetical common grounds for nonfactual utterances which supplies pronominal antecedents. Though independently motivated, this type of accommodation may not serve to introduce previously unmentioned discourse referents, new file cards in Heim’s terms, to serve as antecedents for pronouns. Although we may infer the first disjunct of (18) from the logically equivalent first disjunct of (20), (20) still may not license the same anaphoric relations as (18).

In the formal theory developed here, where accommodation takes place at the DRS level of representation, the requirement under consideration might be expressed as a stipulation that if the accommodated material includes the antecedent of a pronoun in the modally subordinate clause, that material must be

\[ \text{borrowed from a prior representation. But I think that careful consideration will show, as Heim's discussion suggests, that in fact the requirement follows from Gricean cooperative principles for conversation and hence should not be regarded as a stipulation. That is, since pronouns have no descriptive content, the speaker must take great pains to make sure that their intended antecedent discourse referent is in the common ground of the conversation. Introducing a discourse referent through an explicit utterance is the surest way to guarantee this.} \]

This same constraint on accommodation is shown clearly by another example due to Barbara Partee, (24):

(24) (a) Nine of the ten marbles are in the bag.
(b) # It's under the couch.

Here, although (24a) conversationally implicates that there exists a tenth marble which is not in the bag, we may not accommodate this information directly into the DRS, for if we did we would have a potential antecedent for it in (24b), and (24b) seems infelicitous precisely because there is no available antecedent for it. Notice, however, that (24b) seems more felicitous after a long pause, especially if after uttering (24a) the speaker notices that the hearer is looking for something. A solution to this problem was suggested by Lyn Frazier (p.c.). She points out that in order to infer that there is one marble which is not in the bag the hearer must perform a mathematical calculation: she must subtract nine from ten. Even though this calculation seems quite trivial to us, it introduces a factor which was not involved in the previous examples, a nonlinguistic operation. Notice also that in previous examples the accommodated information was simply copied from portions of a pre-existing DRS. This would not be the case here, where the representation for (22a) does not involve a discourse referent for a single marble, but only plural referents for the groups of nine and ten marbles. We cannot automatically assume that the hearer has performed the necessary calculation. However, as Heim (1982) discusses, once it is clear that some entity is salient in the context for all participants in a discourse, as is the case with deixis, for example, we may accommodate a discourse referent for that entity. In the case where the speaker of (24a) notices the hearer looking for something, the speaker may assume that the hearer has performed the calculation, has realized that there is a missing marble, and is looking for it. In our terms, the speaker may assume that there is a discourse referent for the missing marble in the common ground, which then may serve as antecedent for the utterance of a pronoun such as it in (24b).

The account proposed for this example should suffice as well for the contrast in anaphoric potential of the following sentences from Isard (1975):

\[ \text{See Pierrehumbert (1980) and Selkirk (1983) for discussion of the phonology of pitch accent placement, and Ladd (1980), Selkirk (1983) for discussion of the relationship of 'old information,' including presupposition, to lack of pitch accent.} \]
First square 19 and then cube it.

First take the square of 19 and then cube it.

For most speakers, in (25) where the square of 19 has not been explicitly mentioned as such, it may only refer to 19, while in (26), it may either refer to 19 or to the referent of the NP the square of 19.

There is a further point which I would like to note: The phenomenon of anaphora licensed by modal subordination provides an independent argument for a level of discourse representation intermediate between syntactic representation and model theoretic interpretation. Cases such as Landman's (13) and Partee's (18) show that pragmatic accommodation is required, so that neither a syntactic representation such as S-Structure in Government and Binding theory, nor a simple transform of S-Structure such as LF, would suffice to explain the data. An approach which posits operators with discourse scope in an extended version of LF, such as the one that Heim explores in a preliminary way in Chapter II of Heim (1982), would run into the same problems with mixed modals as the insertion approach considered earlier. On the other hand, given the assumptions we have made here about the form of a grammar and its interpretation, we may not explain modal subordination in terms of direct interpretation in a model. Standard models fail to provide discourse referents, which we have seen to be crucial to this account. In addition, consider the discourse in (27):

(a) One of the 10 marbles is out of the bag.
(b) It's under the couch.

If we assume that (24a) conversationally implicates that one marble is not in the bag and that (27a) implicates that nine marbles are in the bag, then (24a) and (27a) will be true in exactly the same worlds. Yet unlike (24a), (27a) provides an antecedent for it in the (b) sentence, and the discourse is felicitous. Hence, I believe that the phenomenon I have described here argues for an intermediate level of Discourse Representation.

1.3 The interpretation of DRSes involving modality

In this section I will present a formal theory of the interpretation of DRSes involving modality. In the examples of modal subordination which were considered in Sections 1.1 and 1.2 above, the explicit or implicit hypothetical extensions of the common ground which provided the antecedents for anaphora were all live possibilities in the discourse up to that point: that is, they were doxastically plausible given the common ground of the participants (at least, given the common ground which the participants as a group assume they share, though one or more may have had evidence to the contrary). But many utterances denote propositions which contradict "the facts", i.e. are incompatible with propositions in the common ground of the conversation. In Section 1.3.1, I will briefly consider a few examples of this type in order to illustrate what is involved in incorporating non-epistemic modals into a general account of modal subordination.15 The discussion of these examples is intended to motivate some of the complexities of the formal theory which follows, in Section 1.3.2.

1.3.1 Subordination with non-epistemic modality and counterfactuals

Consider the following examples:

(a) If I had brought a book with me to Georgia, I could have read it on the plane.
(b) I would probably have finished it by now.

(a) You should eat a bagel.
(b) It would fill you up.

(28a) may be uttered in a context in which I did not bring a book with me to Georgia, and in which this fact is part of the common ground shared by the participants in the discourse. The addition of the antecedent of (28a) to this common ground, then, would leave us with the empty context set. Given the interpretive principles we have discussed so far, this would make the whole conditional vacuously true, since in all the worlds in this empty context set, the consequent would be true. But the conditional seems to be much more informative than this, so that the truth conditions thus derived are too weak. The modal auxiliary should in (29a) has the force of necessity. Its utterance in a given context should not be taken as a suggestion by the speaker to remove all worlds from the context set in which the hearer does not eat a bagel, i.e. as the assertion that it is impossible in this context that the hearer not eat a bagel. Rather, it is a suggestion that it would be ideal in some sense if she ate a bagel, though we generally acknowledge that ideals will not necessarily be realized. So, (28b) and (29b) seem to be parallel to the examples involving anaphora licensed by modal subordination, yet the

15In incorporating non-epistemic modality into this account of modal subordination, I benefited from closely related work by Irene Heim (1985) on modality and propositional attitudes in File Change Semantics.
account I sketched above using the common ground to relativize modal force is not adequate for these cases.

Kratzer (see references) develops compelling arguments for a unified theory of modal with conditionals, including counterfactuals. And we have just seen that modal subordination seems to be possible in non-epistemic conditionals. In both epistemic and non-epistemic cases, the speaker first establishes a set of worlds (possibly not including the actual world) in which some individual \( a \) is said to exist. So long as we continue to talk about those worlds, we may continue to assume \( a \)'s existence and to refer anaphorically to the discourse referent with which \( a \) was originally introduced into the conversation. Given this parallel, we would like to find a more general approach to modal subordination which can account for (28) and (29) as well as the cases already discussed.

As I mentioned in Section 1.1, Kratzer relativizes the modal force of a modal or conditional using two distinct functions, each of which is a function from possible worlds to sets of propositions. The first is called the marital base; the marital base may be epistemic (or doxastic), as in the case of most of the examples of modal subordination we examined above, or circumstantial (facts, of course, need not be known). In the cases of epistemic modality we have considered, the marital base assigns to each world in the context set of the conversation at that point exactly those propositions which are in the common ground, so that the intersection of the propositions assigned to a given world by the marital base equals the context set. What this means is that the usual accessibility relation of modal logic will hold in such cases between just those worlds which are in the context set. However, this is not the case in the non-epistemic examples, so we need some more general terminology. Given a world \( w \) and a marital base \( m \), I will use the expression derived context set for \( w \) to describe the intersection of the propositions in \( m(w) \), i.e. \( \bigcap m(w) \).\(^{16}\)

The other function used to relativize modal force is the ordering source. The propositions assigned to a given world \( w \) by an ordering source \( o, o(w) \), reflect what would be true under (possibly counterfactual) ideal circumstances. These propositions are used to induce an ordering on the worlds in the derived context set for \( w \) given by the marital base: only those worlds in \( \bigcap m(w) \) which also come closest\(^{17}\) to realizing the ideal given by \( o(w) \) will be in the domain of the marital operator involved. A world \( w' \) is closer to the ideal given by \( o(w) \) than a world \( w'' \) if more of the propositions in \( o(w) \) are true in \( w' \) than in \( w'' \).

Kratzer (1980) argues that we interpret counterfactuals using a particular type of marital base and ordering source. The marital base assigns the empty set of

propositions to each world \( w \) in the context set of the conversation at that point. Following the general rule for interpreting conditionals, we take the union of \( m(w) \) with the proposition expressed by the antecedent, call this union \( m^+(w) \). Then with the special type of marital base used for counterfactuals, the derived context set \( \bigcap m^+(w) \) is the set of worlds in which the antecedent is true, whether or not those worlds were in the original context set of the conversation at the point at which the conditional was uttered. But the marital operator may not range over all the worlds in \( \bigcap m^+(w) \); rather, it only ranges over those worlds in the derived context set which come closest to the ideal expressed by the ordering source. For counterfactuals, we use an ordering source which is totally realistic, that is, one which assigns to any world \( w \) just those propositions which are true in \( w \); hence the set of worlds in the intersection of \( o(w) \) equals \( w \). We use this ordering to consider only those worlds where the antecedent is true and which are most like \( w \) in all other respects.

For example, consider the set of all those worlds in which the antecedent of (28a) is true, i.e. in which I did bring a book; in the discourse context described above where I didn't bring a book, this set will be disjoint from the context set, and in fact, the derived context set will be the same for each world in the original context set. We then impose an order on worlds in the set, using the totally realistic ordering source \( o \). So, for any world \( w \) in the original context set, we only consider those worlds in \( \bigcap m^+(w) \) which are most like \( w \). In order for the whole conditional to be true in \( w \), the consequent must be true in each such world. If the consequent is false in any of these most ideal worlds in \( m^+(w) \), then \( w \) is eliminated from the original context set. Otherwise, it is retained. The domain of the marital in (28b) is then restricted using the derived context set and ideal of (28a), where we already know that for any world \( w \) in the context set it is true in all the most ideal worlds in its derived context set that I brought a book and read it on the plane. This will be reflected in the DRS for (28b) by the accommodation of the marital for the antecedent of (28a), as in the epistemic examples we have considered.

In (29a), should is deontic; imagine that the common ground in this case includes propositions such as the fact that the hearer is hungry, that bagels are readily available here, and that eating a bagel will satisfy the hunger. However, it may also include the fact that the hearer is anorexic and refuses to eat anything. That is, in all of the worlds in the corresponding context set, the hearer will not eat a bagel. Let us assume that the marital base used in interpreting the marital in (29a) assigns to each world in the context set \( w \) a subset of the propositions in the common ground; for example, the propositions that the hearer is hungry, that bagels are available, etc., but not the fact that the hearer is anorexic. This marital base will determine a derived context set which is a superset of the context set at that point in the conversation. Assume further that for each world \( w \) in the original context set, the ordering source, \( o \), characterizes what are considered good nutritional practices in \( w \); for example, it might assign to \( w \) the proposition that someone who is hungry will do something to satisfy the hunger. We can call

\(^{16}\)I borrow the term derived context set from Stalnaker (1985), following Heim (1985).

\(^{17}\)Actually, there may be no 'closest' set of worlds; cf. Lewis (1973) for discussion. Kratzer (1980) takes this into account in her formal definition of the ordering source. I will ignore this complication in the informal discussion in the interest of conveying the basic intuition more clearly. However, it will be incorporated into the formal theory which follows.
the worlds in the derived context set ∩m(w) which come closest to o(w), the **nutritional ideal** for w. The speaker’s assertion amounts to an instruction to discard any worlds w from the context set whose nutritional ideal includes a world in which the hearer doesn’t eat a bagel, whether she will in fact do so in w or not. We might paraphrase this interpretation of (29a) as, ‘In view of the fact that you’re hungry, that bagels are available and would satisfy the hunger, and that one who is hungry would ideally do something to satisfy the hunger, you should eat a bagel.’

Note that for any given world w, m(w) need not include all the propositions which were in the common ground prior to uttering (29), and in particular not the propositions that the hearer is anorexic and will refuse to eat anything. Hence, we can extend (29a) as in (30), without contradiction:

(30) You should eat a bagel, but you won’t.

The second conjunct in (30) is factual in mood. Unlike the first conjunct, in interpreting it we take into account all we know about the actual world, including the hearer’s anorexia, etc.

The material in (29a) which is under the scope of the modal, ‘you eat a bagel’, is non-factual; that is, the modal should signals that we must consider the truth of ‘you eat a bagel’ in each world in the possibly nonrealistic nutritional ideal of w, for each w in the context set. Would in (29b) then continues the nonfactual mood; we need to relativize its modal force to some contextually salient set of propositions. In this context, the paraphrase which most plausibly captures this relativization is ‘if you ate a bagel, it would fill you up’. Since the antecedent of this conditional is at least implausible, given what we know about anorexia, we interpret (29b) as if it were the antecedent of a counterfactual. Now the modal base is empty, and the ordering source is totally realistic, as in (28). That is, for each world w in the context set, we consider all worlds in which the hearer ate a bagel which are closest to the ideal, that is, to the way things actually are in w. If the bagel fills the hearer up in all those ideal worlds, then (29b) is true in w, and it is retained in the context set which results. Otherwise, it is eliminated.

In translating this intuitive analysis into DRS terms, we need to make a decision about the role of a modal base and ordering source in deriving such representations. Heretofore, we have assumed that we relativize modal operators both to material in the DRS to their left and also to material in superordinate DRSes. But the non-epistemic examples show that this would not give the correct results in all cases, since they may involve a non-realistic modal base or ordering source; a modal base m is non-realistic if for some w it is not the case that w ∈ m(w). Further, a modal base or ordering source may assign a different set of propositions to each world in the context set, so that we cannot simply introduce a single set of relativizing propositions into the DRS to the left of the modal operator. And more than one modal base or ordering source may play a role in the interpreta-

1.3. DRS Interpretation

1.3.2 A formal theory of DRSes with modality

Given this informal discussion of the interpretation of non-epistemic modals, we turn now to consider a formal theory of Discourse Representation Structures which include modality. In this chapter, I will not present a general theory of the mapping from syntactic structures onto DRSes. See the proposal in Chapter 4 of DRS construction for sentences without modals; see also Heim (1982) for a somewhat different approach within File Change Semantics. As we have seen, examples involving modal subordination raise new questions about the non-trivial way in which pragmatic factors enter into DRS construction. Heim (1982), Partee (1984), and Kadmon (1987) discuss other types of examples where non-syntactic input is crucial in constructing appropriate files or DRSes. However, in order to respect the constraint on accommodation discussed at the end of Section 1.2, we must stipulate that discourse referents may only be accommodated when a) they are to serve as the antecedent of a definite description (as discussed by Heim (1982)), b) they are licensed by deixis or other clear contextual factors (also discussed by Heim), or c) they are borrowed from the prior representation of explicitly uttered NPs.

Syntax:

DRS, the language of discourse representation structures (DRSes), is based on a set VAR of variables, a set of n-place predicates (for all n), and the relation

---

18In developing this, I benefited from Landman (1987), which lays out the main parameters to be considered in formalizing Discourse Representation Theory, and develops several versions of a simpler, extensional form of the theory.
symbols ¬, ∨, ⇒, ⊢, and ◊. DRL is the set of all DRSes, where DRSes are defined as follows:

A DRS K is a pair \((X_K, C_K)\), where \(X_K\), the local domain of K, is a finite set of variables and \(C_K\), the set of conditions in K, is a finite set of conditions.

Conditions are all and only the following:
1. If \(P\) is an n-place predicate and \(x_{i_1}, \ldots, x_{i_n}\) are variables, then \(P(x_{i_1}, \ldots, x_{i_n})\) is an atomic condition.
2. If \(K\) is a DRS, then \(\neg K\) is a condition.
3. If \(K_i\) and \(K_j\) are DRSes, then \(K_i \lor K_j\) is a condition.
4. If \(K_i\) and \(K_j\) are DRSes, then \(K_i \Rightarrow K_j\) is a condition.
5. If \(K_i\) and \(K_j\) are DRSes, then \(K_i \land K_j\) is a condition.
6. If \(K_i\) and \(K_j\) are DRSes, then \(K_i \land K_j\) is a condition.

The following syntactic notions may be defined on (occurrences of) DRSes:
- \(\preceq\) is accessible from; \(\preceq\) is the smallest partial order on DRSes such that for any DRS \(K\), if \(\neg K_i \in K\), or \(K_i \Rightarrow K_j\) in \(C_K\), or \(K_i \land K_j\) in \(C_K\), then \(K \preceq K_i\) and \(K \preceq K_j\); and if \(K_i \lor K_j \in C_K\), then \(K \preceq K_i\) and \(K \preceq K_j\).
- The accessible domain of \(K\), \(A_K\), is the set of all variables in (local) domains of DRSes accessible from \(K\): \(A_K = \bigcup_{K \preceq K} X_K\).

We then impose the following condition on DRSes:

No free variables: if \(x\) occurs in an atomic condition in \(C_K\) then \(x \in A_K\).

Semantics:

A model \(M\) for DRL is a structure \((W, A, i)\), where \(W\) is a set of possible worlds, \(A\) is a nonempty set of individuals, and \(i\) is the interpretation function mapping pairs of an n-place predicate and a world into \(pow(A^n)\).

An assignment function, \(f\), is a total function from \(VAR\) to \(A\).

Given two assignment functions, \(f\) and \(g\), \(f\) varies from \(g\) at most with respect to \(X\), \(f(X) = g(X)\), if \(\forall y (\neg (y \in X) \rightarrow f(y) = g(y))\).

A proposition is a set of possible worlds.

A modal base or ordering source is a function from possible worlds to sets of propositions.

The truth of a DRS with respect to a world and an assignment function is defined recursively, as follows:
For all worlds \(w, u, v, w', u'\), assignment functions \(f, g, h\), modal bases \(m\), ordering sources \(o\), models \(M\), DRSes \(K, K_i, K_j\), sets of conditions \(C\), n-place predicates \(P\), and variables \(x\):

1. \(\langle w, f \rangle \models K\) iff \(\forall v \in C_K (\langle w, f \rangle \models_C c)\)
2. (a) \(\langle w, f \rangle \models P(x_{i_1}, \ldots, x_{i_n})\) iff \(\langle f(x_{i_1}), \ldots, f(x_{i_n}) \rangle \in \mu(P)(w)\)
   (b) \(\langle w, f \rangle \not\models \neg K_i\) iff \(\exists g (\langle X_{K_i}, f \& \langle w, g \rangle \models K_i\)
   (c) \(\langle w, f \rangle \models (K_i \lor K_j)\) iff \(\exists g (\langle X_{K_i}, f \& \langle w, g \rangle \models K_i\)
   \lor \exists g (\langle X_{K_j}, f \& \langle w, g \rangle \models K_j\))
   (d) \(\langle w, f \rangle \models (K_i \Rightarrow K_j)\) iff \(\forall g (\langle X_{K_i}, f \& \langle w, g \rangle \models K_i\)
   \Rightarrow \exists h (\langle X_{K_j}, g \& \langle w, h \rangle \models K_j\))
   (e) \(\langle w, f \rangle \models (K_i \land K_j)\) iff \(\forall u, g (\langle X_{K_i}, f \& \langle u, g \rangle \models K_i\)
   \land \exists u' (\langle X_{K_j}, u' \models K_j\) \land \langle w, g \rangle \models K_i)\)
3. \(K\) is true in a world \(w\) iff \(\exists f (\langle w, f \rangle \models K)\).

The first clause of the recursive definition of the truth conditions for a DRS \(K\) tells us that \(K\) is embeddable, or is verified in a model relative to a world and an assignment function iff all the conditions in \(K\) are satisfied relative to those same elements. This notion of satisfaction is then defined in the second clause. Clauses (2a) through (2d) are fairly simple and require no special comment. The complexity of clause (2e), which is intended to reproduce in DRS terms the truth conditions for conditional necessity in Kratzer (1980), is necessitated by the fact that in the derived context set, \(\cap \{m(w) \cup \{\langle v, g \rangle \models K_i\}\}\) (which is the formal specification of the \(m^*(w)\) of the previous section), there may be no set of worlds 'closest' to \(w\) under the ordering induced by \(o(w)\). Instead, we require that there be some world \(w'\) in the derived context set such that all worlds \(w'\) in the context set which are at least as close to the ideal as \(w'\) are such that \(K_i\) is embeddable in them. If we were only considering epistemic examples, where the ordering source is generally irrelevant, we could simplify this clause considerably, ignoring the ordering source as follows:

\(((2e') \langle w, f \rangle \models (K_i \land m, K_j)\) iff \(\forall u, g (\langle X_{K_i}, f \& \langle u, g \rangle \models K_i\) \land \exists h (\langle X_{K_j}, g \& \langle w, h \rangle \models K_j\))\)

Since in the epistemic examples \(u \in \{m(w)\} \text{ iff } u\) is in the context set prior to interpreting the condition, this clause amounts to saying that the conditional is true iff in all worlds in the context set in which the antecedent is true (i.e., worlds in the hypothetical context set), the consequent is true as well. Now we can see
that the clause for the material conditional, (2d), just gives truth conditions for a subcase of (2e') where the modal base is what Kratzer calls totally realistic, that is, where for any \( w \), the only world in \( m(\omega) \) is \( w \) itself.

### 1.4 Generalized subordination in discourse

The examples of subordination in discourse which we have examined so far all involve modality in some form or other. However, there are examples displaying anaphoric phenomena very much like those considered above, but where the subordination appears to involve only non-modal operators. In this section, I will present a few such cases and discuss informally how they may be related to modal subordination.

Karttunen (1976) noticed cases like (31):

\begin{enumerate}
\item[(31)]
\begin{enumerate}
\item (a) Harvey courts a girl at every convention.
\item (b) She always comes to the banquet with him.
\item (c) The girl is usually very pretty.
\end{enumerate}
\end{enumerate}

Subordination in (31) is induced by adverbs of quantification (Lewis 1975): at every convention in (a) establishes a limited set of cases or situations which are contextually salient and serve to restrict the domain of always in (b) and usually in (c).

Something similar seems to license the anaphoric relation between the definites and the preceding indefinites in (32), from Stenning (1978):

\begin{enumerate}
\item[(32)]
\begin{enumerate}
\item (a) In each room, there was a cat and a goldfish.
\item (b) [The goldfish] dived.
\item (c) [The cat] caught [it].
\end{enumerate}
\end{enumerate}

Here, the adverbial in each room quantifies over locations, and (b) and (c) are implicitly offered as instantiating the situation in such a location.

Temporal parallels were noted in Sells (1985). His example (29b) is given in (33):

\begin{enumerate}
\item[(33)]
\begin{enumerate}
\item (a) A train leaves every hour for Boston.
\item (b) It always stops in New Haven.
\end{enumerate}
\end{enumerate}

The relevant reading of (33a) is that in which a train is under the scope of the temporal quantifier every hour. Partee (1984) has offered an analysis of temporal quantifiers, introducing discourse referents for events into DRSes. The scope of such operators seems to be limited to the sentential domain, as with other types of operators. Yet, it in (33b) seems to refer back to a train. Our intuition that (33b) is temporally subordinate to (33a) is confirmed by the most natural interpretation of the optional temporal adverb always. We do not take it to quantify over all times; rather, as in the modal cases we have considered above where the operator's domain is restricted by the accommodation of contextually salient material, the domain of always here is restricted to the salient set of times: those in which a train leaves for Boston.

We can easily imagine DRSes for the examples in (31)–(33) which parallel the representations I have proposed for examples involving modal subordination, but use non-modal operators. But how would we interpret such representations? When operators range over cases, over locations of a particular sort, or over temporal intervals, are they quantifying over entities which differ essentially from the possible worlds over which modal operators range?

Recent work on the use of partial models in semantic interpretation, such as Barwise and Perry (1983), Kratzer (1985), and Landman (1986), suggests that these operators do not differ essentially in their range. Rather, they all range over situations, however these are defined in a particular theory. For example, in this type of approach, we might paraphrase the truth conditions for (31) informally as, 'In every situation which is a convention, Harvey courts a girl, and in every such situation, she comes to the banquet with him.' In the context of the theory developed here, we might think of situations as partial worlds (though, of course, this is not their character in all of the theories mentioned). Possible worlds themselves can be thought of as just the limit case of total situations. So, when we move to using partial models, I believe that we will find that we can provide a unified analysis of the generalized discourse subordination involved in the cases in (31)–(33), as well as in the modal examples considered earlier. In each case, the second sentence in a discourse is interpreted as involving an operator (explicit or implicit) whose force is relativized so that it ranges only over the type of situation given in part by the first sentence. The development of a formal theory of DRS interpretation which incorporates this idea goes beyond the scope of this chapter; however, see Roberts (forthcoming) for such a theory, as well as for a detailed consideration of apparent restrictions on sequences of tense in discourses involving modal subordination.

There is another type of case which seems closely related to discourse subordination and involves universal quantifiers. (34) is due to Barbara Partee (p.c.), (35) to Sells (1985):\(^{19}\)

---

\(^{19}\)Sells offers (35) [=his 31] as an example of temporal subordination. The reference time of (b) is that provided by (a); however, this does not necessarily mean that temporal subordination is involved. In Sells' example (33), on the other hand, the quantifier in (a) under whose scope (b) seems to fall is a quantifier over times.

Sells also offers (i) [=his 33b] as an example of temporal subordination:
1.5. Conclusion

In these cases, we begin a narrative with a statement about a class of individuals, then we zoom in on one instantiation of that class to continue the narrative. The problem with (37) seems to be that for some reason the second sentence does not comfortably continue the narrative in this fashion. At present, the precise character of the constraints on telescoping eludes me, so that here I cannot do more than note that these cases exist and that they may be related to the phenomenon under consideration here.

1.5 Conclusion

Summarizing, modal subordination is a phenomenon wherein the interpretation of a clause α is taken to involve a modal operator whose force is relativized to some set β of contextually given propositions. We say that in that context, α is interpreted as modally subordinate to the propositions in β, or, rather loosely, that α is modally subordinate to the clauses (if any) used to express β. Intrasententially, the consequent of a conditional is modally subordinate to the antecedent clause, as in conditional donkey sentences. Intersententially, modal subordination involves the accommodation of material from prior utterances to serve the role of an antecedent for the subordinated clause. In both cases, the antecedent clause, whether or not accommodated, plays a crucial part in determining the derived context set which is used to restrict the range of the (explicit or implicit) operator over the modally subordinate clause. In cases which involve pronoun anaphora, accommodation is constrained by the requirement of explicit prior representation of potential anaphoric antecedents.

Note that under the conception just sketched, not all cases of intersentential anaphoric relations where the anaphoric antecedent occurs in a nonfactual utterance should be analyzed as involving modal subordination. Consider, for example, (39) (due to Fred Landman, p.c.) and (40) (due to Jerry Morgan, p.c.):\(^ {20}\)

(i) (a) The author claims that Vul\(\text{can}\) exists after all.
(b) It has circled around Mercury for ages without us ever noticing it.

(39) (a) Last night I dreamed I got a red Pors\(\text{che}\) for my birthday.
(b) I drove it all over the countryside and loved every minute of it.
(c) This morning I woke up and much to my surprise found it parked in my driveway.

(38) (a) The author claims that Vul\(\text{can}\) exists after all.
(b) It has circled around Mercury for ages without us ever noticing it.

Although the adverb of quantification usually is sometimes temporal, I don’t believe it is in this case. Rather, here it seems to quantify over something like cases or instances of the rice-grower/wooden cart pairs introduced in (a). Hence, I take (i) to illustrate a case of subordination to the universal quantifier.

\(^ {20}\)Landman and Morgan don’t necessarily share my analysis of these examples.
Chapter 2
Anaphora, Coreference and the Binding Theory

Here are two kinds of theories of anaphora in the current literature on natural language: One deals with constraints on anaphora in discourse. This is the sort of theory I explored in the discussion of modal subordination in Chapter 1. The other is primarily a sentence-level theory and is exemplified by the Binding Principles of Government and Binding Theory (see Chomsky (1981)). This theory seeks to explain anaphora in terms of configurational notions such as command and governing category. While there is some overlap in the concerns of the two types of theory, such as in the relation of quantifiers to the variables they bind, neither seems to address the full range of anaphoric phenomena. As a result, I will argue, there is no clear delineation of which anaphoric constraints are dependent on sentence-internal configurations and which are best explained in terms of the structure of discourse.

In this chapter I will suggest how this might be clarified, proposing that although the basic thrust of Binding Theory is correct, at least for English, many problems can be resolved by reconsidering the role of indices in semantic interpretation. Some of the problems which will concern me are exemplified in (1) – (3). Each case poses a problem for the interpretation generally assumed of the randomly assigned indices — that coindexed noun phrases are coreferential, while non-coindexed noun phrases are not:

1. Alan and Margaret ate their dinner.
2. Mary thought she had the mumps and Alice did too.
3. (a) Only Reagan voted for himself.
   (b) Only Reagan voted for Reagan.
2.1 Problems in the theory of binding

2.1.1 Syntactic problems

In this section I will examine the Binding Theory as it is outlined in Chomsky (1981), Lectures on Government and Binding (hereafter LGB). The theory pertains to relationships between NPs in A-positions (intuitively, argument positions in matrix clauses, as opposed to preposed or extraposed positions). There are two basic notions in this theory. The first is that of government, defined as follows (LGB, p.250):

(4) Consider the structure (i):

\[ \beta \ldots \gamma \ldots \alpha \ldots \gamma \ldots \], where

(a) \( \alpha = X^0 \) or is coindexed with \( \gamma \)

(b) where \( \phi \) is a maximal projection, if \( \phi \) dominates \( \gamma \)

\[ \text{then } \phi \text{ dominates } \alpha \]

(c) \( \alpha \) c-commands \( \gamma \)

In this case, \( \alpha \) governs \( \gamma \)

The derived notion of a governing category is defined in terms of (4) (LGB, p.211):

(5) \( \beta \) is a governing category for \( \alpha \) iff \( \beta \) is the minimal category containing \( \alpha \), a governor of \( \alpha \), and a SUBJECT accessible to \( \alpha \).

The other basic notion is that of binding, defined as follows for NPs in A-positions (cf. LGB, p.184):

---

1This idea has been foreshadowed elsewhere in the literature. In particular, Partee (1978) suggests that there are two kinds of pronoun, a bound variable and one whose reference is pragmatically determined. Bach & Partee (1980), Partee & Bach (1981) and Reinhart (1983) all assume that there are pronouns which should not be treated as bound variables, but are otherwise related to their antecedents. With the advent of discourse theories such as Heim’s and Kamp’s, the nature of this other type of relation can now be clarified and its constraints explored, as we have done in Chapter 1. Note that unlike Cooper (1979), I am not here proposing that pronouns which are bound in the two different fashions receive different semantic interpretations: every pronoun will be represented as a discourse referent in the discourse representation and ultimately behave as if it were a bound variable in a model theoretic interpretation.

---

2See LGB, pp.209-211 for discussion of the notion of SUBJECT.
\[\alpha \text{ is } A\text{-bound by } \beta \text{ if and only if } \alpha \text{ and } \beta \text{ are coindexed, } \beta \text{ c-commands } \alpha, \text{ and } \beta \text{ is in an } A\text{-position.}\]

For our purposes here we may assume that a node \(\beta\) c-commands a node \(\alpha\) iff the first branching node which dominates \(\beta\) dominates \(\alpha\). If an NP is not bound, then it is free. The Binding Theory is then as follows:

1. **Binding Theory**
   - (A) An anaphor is bound in its governing category.
   - (B) A pronoun is free in its governing category.
   - (C) An R-expression is free.

Anaphors in this theory include reflexives and the traces of NP-movement. All other pronouns are pronominals. Other NPs and traces of wh-movement are R-expressions. In order to avoid the confusion which may arise between the more traditional uses of these terms and their technical sense in GB theory, I will refer to reflexives and reciprocals as A-pronouns (pronouns governed by Principle A) and to other pronouns as B-pronouns (those governed by Principle B).

The focus of inquiry here will be on the problem of determining at which level of the syntax to apply the Binding Theory. In LGB (p.169f.) Chomsky adopts the position that the Binding Theory applies at S-Structure (SS). Brody (1979) pointed out that applying Principle C at Logical Form (LF) gave incorrect results. For example, consider the sentences in (8) and their LFs in (9) (LGB pp.196–7):

\[
\begin{align*}
(8) & \quad \text{(a)} & \text{which book that John read did he like} \\
(b) & & \ast \text{ he liked every book that John read}
\end{align*}
\]

\[
\begin{align*}
(9) & \quad \text{(a)} & \text{for which book x that John read, he liked x} \\
(b) & & \text{for every book x that John read, he liked x}
\end{align*}
\]

The binding possibilities for (8a) and (8b) differ at SS, since in (b) John, an R-expression, is bound, violating Principle C of the Binding Theory. But their LFs are essentially identical — in neither case does he c-command John, and hence both would be predicted to be grammatical if all the binding principles applied at LF.

Binding cannot apply at Deep Structure (DS) because of the effect of move-\(\alpha\) on binding possibilities, as in (10):

\[
\begin{align*}
(10) & \quad \text{(a)} & \ast \text{it seemed to themselves that the athletes were fit for the race} \\
(b) & & \text{The athletes seemed to themselves to be fit for the race.}
\end{align*}
\]

However, the analysis of crossover (cf. pp.193-4) seems to require the application of the Binding Theory, or at least Principle C, at LF. The typical strong crossover paradigm is as in (11):

\[
\begin{align*}
(11) & \quad \text{(a)} & \ast \text{who said he say Mary had seen ti} \\
(b) & & \ast \text{he said Mary had seen everyone} \\
(c) & & \text{who, ti said Mary had kissed him}
\end{align*}
\]

In terms of the Binding Theory, (a) is ungrammatical because \(t_i\), an R-expression, is bound by \(h\), violating Principle C. After Quantifier Raising (QR) at LF, the (b) case is ungrammatical for the same reason, the trace of every everyone, bound by \(h\), while (c) is grammatical because the R-expression \(t_i\) is A-free, being bound only by the operator in \(A'\) position, who, and A-binding in turn the pronoun him. The parallel between (a) and (b) depends upon the LF rule of QR.

Chomsky considers an alternative explanation of the crossover facts in (9) in which (b) is ruled ungrammatical by Principle C at SS, every everyone being treated as an R-expression. This would resolve the apparent conflict between the facts in (6) and (9). We might then consider whether the principles of the Binding Theory all apply at SS.

However, there is another type of case which seems to require the application of the Binding Theory at LF; this includes examples with preposed NPs, as in (12) and (13):

\[
\begin{align*}
(12) & & \text{which picture of himself does John like} \\
(13) & & \text{that picture of himself John likes}
\end{align*}
\]

In neither of these cases does John c-command himself at SS. It is generally assumed that the proper configuration can be obtained by a rule of reconstruction at LF. After application of such a rule, (13) would presumably look like (14):

\[
\begin{align*}
(14) & & \text{John likes that picture of himself}
\end{align*}
\]
where the proper c-command and governing category requirements of the Binding Theory are met. However, in the usual representation of questioned sentences at LF, the \textit{wh}-phrase is treated as an operator, with the pied-piped material acting like a domain restrictor:

\begin{equation}
\text{For which } x, x \text{ a picture of himself}, \text{John likes } x
\end{equation}

Here, the proper structural relation between anaphor and antecedent is still not obtained.

Van Riemsdijk & Williams (1981), point out that to obtain the required configuration in such a case, either an undesirable structure building rule is required at LF, or we must assume an ill-motivated system of layered traces. They suggest that the answer to the reconstruction problem is to divide Move-\(\alpha\) into two rules: NP-movement applying to Ds to derive representations at a new level, NP-Structure, followed by \textit{wh}-movement, yielding SS. The Binding Theory, then, would apply at NP-Structure. They assume that topological NPs originate in the matrix sentence, are \textit{wh}-moved into COMP and then dislocated and raised into topic position. Hence their proposal is intended to account for cases like (12) and (13). But note that their analysis also encounters problems with anaphora and \textit{wh}-movement. Consider (16) and (17):

\begin{equation}
(16) \begin{align*}
\text{(a)} & \quad \text{That picture of John, he likes} \\
\text{(b)} & \quad \text{Which picture of John, does he like}
\end{align*}
\end{equation}

\begin{equation}
(17) \begin{align*}
\text{(a)} & \quad \text{* He likes that picture of John} \\
\text{(b)} & \quad \text{* He likes which picture of John}
\end{align*}
\end{equation}

The sentences in (16) seem fine after \textit{wh}-movement, while the corresponding unmovable examples in (17) are ungrammatical. If Binding principles apply at NP-Structure, but \textit{wh}-movement occurs in the mapping from NP-Structure to S-Structure, then these facts would not be predicted.\footnote{I ignore here the problem of how the rule would be formulated. Given the special Topic node under S\(^o\) proposed in Chomsky (1977), and subsequent work by Koster on sentential subjects (1978), we might either suppose that topics are first \textit{wh}-moved and then dislocated from COMP into TOP, or that they are base generated in TOP, with a PRO base-generated in the matrix clause and \textit{wh}-moved into COMP, along the lines of Chomsky’s treatment (1981) of purpose clauses. In neither case, then, would “reconstruction” be simply a case of undoing \textit{wh}-movement.}

Finally, we might consider whether we can resolve these problems by making reconstruction of a \textit{wh}-moved constituent optional at LF. David Pesetsky (p.c.) has pointed out examples like (18a):

\begin{equation}
(18) \begin{align*}
\text{(a)} & \quad \text{[Which picture of himself, that Mary likes], will John give her, for}\_k \\
\text{(b)} & \quad \text{* John will give her, which picture of himself, that Mary likes}
\end{align*}
\end{equation}

Here, reconstruction of the \textit{wh}-moved constituent into the position of \(t_k\) as in (18b) is both obligatory, so that \textit{John} may bind \textit{himself}, and impossible, since then \textit{her} would c-command \textit{Mary}. Since the intended reading of (18a) seems felicitous, we may conclude that optional reconstruction is not the answer to our problem.

Hence, we face problems whether applying Binding Theory at LF, SS or NP-Structure. A further possibility is that while Principle C applies at SS, Principles A and B apply at LF. However, in view of the problems just noted with reconstruction, it is not clear how this proposal would be an improvement.

### 2.1.2 Problems of interpretation

It has never been made entirely clear in the Government and Binding framework how indices are to be interpreted, although the terminology used in describing the intuitions on which the theory is based, e.g. “coreference” and “disjoint reference,” strongly suggests that indices are to be interpreted. In fact, as Reinhart (1983) has pointed out, if you treat indices just as uninterpreted syntactic devices, then you still have a problem with sentences like (19), which motivated Lasnik’s (1976) important discussion of disjoint reference conditions, leading to Principles B and C of the Binding Theory:

\begin{equation}
(19) \quad \text{Felix likes him.}
\end{equation}

If indices are uninterpreted, then even though you may rule out a derivation in which \textit{Felix} and \textit{him} are coindexed, there is nothing to prevent \textit{him} from picking up the same referent as \textit{Felix} from some previous mention of that individual in discourse. Hence, we must seek an adequate interpretation.

In the LGB system, single indices are assigned at random somewhere before SS. Then the relation of binding is defined over pairs of coindexed NPs where one Structure is fairly abstract, and is not transformationally related to D-Structure or S-Structure. This permits him to account for bound anaphora, including reflexives, in pseudo-clefts and other constructions where there are clear arguments against a transformational relation to S-Structure (see Higgins (1972) for decisive arguments against such a transformational relation).
c-commands the other. If not bound in this fashion, an NP is free. It seems that the usual, though tacit interpretation of the LGB system is simple: if two NPs are coindexed, whether bound or free, they are coreferential; if they are not coindexed, they are not coreferential. The bound/free distinction only serves to identify the environments in which A-pronouns must be bound and in which disjoint reference holds, via the independent Principles A, B and C of the Binding Theory.\(^5\)

Note that “coreferential” here is used in a fairly loose sense. Chomsky is careful to note (LGB, p.314):

Recall that we are not considering the problems of the theory of reference... but are concerned rather with properties of LF-representation that enter into interpretations of sentences in terms of intended coreference and intended distinct (disjoint) reference, where the “reference” in question does not carry ontological commitment.

The consequences of this view, especially for disjoint reference, have not been made sufficiently clear in the framework. Chomsky (p.315, fn.3) briefly considers the following example.\(^6\)

\[(20)\] I dreamed I was Jesus and I forgave me for my sins.

He points out that “a different ‘referent’ will presumably be assigned to I and me in the ‘constructed world’ of the dream,” though there is another sense in which one might claim that the two terms corefer.

From this point of view, consider (21):

\[(21)\] I\(_1\) like me\(_2\)

Lasnik (1980) argues that (20), like the variation where the pronouns are coindexed (ruled out by Principle B), should be ungrammatical because of the lexical meaning of the pronouns. But the apparent awkwardness of this sentence can be overcome in the proper context:

\[\text{We see, then, that like example (20), (21) is not ungrammatical, but only pragmatically odd, in the sense that we seem to be referring to the speaker with two terms which are not coreferential.} \]

A similar problem underlies Lasnik’s (23):

\[(23)\] We\(_1\) like me\(_2\)

The coindexed version of this sentence would be ungrammatical, since we and me are only overlapping in reference, not coreferential.\(^7\) But to interpret non-coindexed as disjoint reference in the strictest sense seems incorrect, given the meanings of the pronouns.

Now consider the following:

\[(24)\]

\begin{align*}
(a) & \quad \text{He}_1 \text{ bought himself}_1 \text{ a dog (bought a dog for himself}_1). \\
(b) & \quad \ast \text{He}_1 \text{ bought him}_1 \text{ a dog (bought a dog for him}_1). \\
(c) & \quad \ast \text{We}_1 \text{ bought myself}_1 \text{ a dog (bought a dog for myself}_1). \\
(d) & \quad \ast \text{We}_1 \text{ bought myself}_2 \text{ a dog (bought a dog for myself}_2). \\
(e) & \quad \text{We}_1 \text{ bought me}_2 \text{ a dog (bought a dog for me}_2). \\
(f) & \quad \text{We}_1 \text{ heard me}_2 \text{ on the radio.} \\
\end{align*}

In (24a – b), we see that a reflexive is required for coreference in this configuration.\(^8\) But in (c) and (d), reflexive myself is not grammatical; this is because of lack of agreement in (c), violation of Principle A in (d). However, me is acceptable in (e), though presumably it could not be coindexed with we. (f) simply provides another grammatical example of this relation. So what seems to be at issue, again, is coreference-under-a-description or some related notion, rather than coreference in the strictest sense.\(^9\)

\(^5\)Reinhart (1983) seems to interpret the system in LGB somewhat differently; cf. especially pp. 52 – 53.

\(^6\)Compare (i), from Lakoff (1972):

\[(i)\] I dreamed that I was Brigitte Bardot and that I kissed me.

\(^7\)Lasnik argues convincingly against allowing coindexation to be interpreted as overlapping reference. I will not repeat his arguments here.

\(^8\)Except in certain hillbilly dialects, where him is acceptable in (b).

\(^9\)Consider Chomsky's related example, (LGB, 5.1, 1.ii)):
In this light, consider some of the problems raised by Evans (1980). One type of example shows that non-coreference must be defined with respect to speakers' intentions. Among these are cases of mistaken identity, as in the utterance of (25) where the speaker does not recognize that the man in the sunglasses is in fact Lou:

(25) That man in the sunglasses resembles Lou.

This seems to have the same flavor as the example in (20); for the speaker, the two terms are not coreferential.

Another type of apparent counterexample to Principle C is exemplified in the sentences in (3):

(3) (a) Only Reagan voted for himself.
    (b) Only Reagan voted for Reagan.

Here it is less plausible to claim that there is some sense in which the speaker intends two different referents for the two tokens of the name Reagan in (3b). Chomsky (LGB p.227, fn.27) responds to this second type of case by claiming that “Principle (C) may be overridden by some condition on discourse, not a very startling fact,” and suggest the following principles:

(26) Avoid repetition of R-expressions, except when conditions warrant.
(27) When conditions warrant, repeat.

Note that (26) is independently required to account for the oddness of (28), an example pointed out by Reinhart (1983):

(28) The flowers that we bought for Zelda pleased Zelda.

Since the first instance of Zelda does not c-command the second, then even if they are coindexed, the second is free. (Note that either instance may be replaced by her, but that herself is infelicitous.) Hence, Principle C does not rule out this type of example, as it would (29):

(29) Zelda liked the flowers that we bought for Zelda.

(26) seems in fact to cover all the cases ruled out by Principle C, with the exception of the crossover cases. This makes the principle look suspiciously redundant.

It has been suggested by some (see, for example, discussion in Lasnik (1980)), that this and other problem cases point to the superiority of the “On Binding” (Chomsky (1980a)) system of binding, where a given NP has both a referential index and one or more anaphoric indices. Lasnik’s (1980) example in (30) shows that plural pronouns with split antecedents pose a problem for the LGB Binding Theory, since coindexing they with either John or Bill would seem to exclude the other as a “partial” antecedent, yet coindexing it with neither would seem to exclude both:

(30) John₁ told Bill₂ that they₇ should leave.

Lasnik points out that in the “On Binding” framework, if we consider two NPs which have neither referential nor anaphoric indices in common to be referentially free, so that they may either corefer or not, then assigning a distinct index to they, say $\theta$, leaves it free to be coreferential with the two proper names as a set. But similar problems occur with conjoined NP antecedents, as illustrated in Seely’s (1), and here Lasnik’s solution seems inadequate:

(1) Alan and Margaret ate their dinner.

In (1), not only is there the issue of how they can at once refer to both Alan and Margaret, with their different indices, but of how to avoid claiming that Alan and Margaret are disjoint in reference from the NP which includes them. Surely we would not want to claim simply that the conjoined NP is free in reference and just happens to be coreferential with the set of NPs which it contains.

A different sort of problem is raised by what Saxon (1984) calls disjoint anaphors. She describes such pronouns in Dgo, an Athapaskan language of northern Canada. These apparently have the same distribution as reflexives, and hence would seem to fall under Principle A of the Binding Theory, but a disjoint anaphor is to be interpreted as having any referent other than the NP to which it is bound. If English was like Dgo, we might have such a pronoun, call it herother. Then Mary likes herother would be grammatical, with herother bound by Mary and meaning ‘Mary likes some other person we’ve already mentioned.’ The sentence * Herother likes Mary would be ungrammatical because the pro-
noun is not bound in its governing category. This is obviously a problem for the assumption that bound NPs are coreferential.

I note that Higginbotham's (1983) linking framework provides a solution for the problems with split antecedents and conjoined NP antecedents, since a pronoun (or the conjoined NP) may be referentially dependent on (linked to) more than one NP. However, Higginbotham still has problems with the McCawley/Evans cases, with the disjoint anaphors, and with the sloppy identity cases; we will discuss in the following section, as well as leaving unresolved the question of the relationship between referential dependence and coreference.

### 2.1.3 Reinhardt’s proposal

Here we will briefly consider Reinhardt’s paper “Coreference and bound anaphora” (1983). She proposes a radical revision of the Binding Theory. Instead of freely indexing structures which are then filtered by Principles A, B and C, she restricts the assignment of indices to pairs of an NP and a pronoun which satisfy the requirements on bound anaphora. She defines this term as follows: “I will use the term bound anaphora for all and only the cases where the pronoun is interpreted as a bound variable.” Bound variables are interpreted in her system via lambda abstraction, so that a coindexed NP and pronoun will end up as variables bound by the same lambda operator. Her indexing procedure is as follows:

\[(31) \text{(Optional)}\]

Coindex a pronoun P with a c-commanding NP \(\alpha\) (\(\alpha\) not immediately dominated by COMP or \(S'\)).

**Conditions:**

(a) If P is an R-pronoun, \(\alpha\) must be in its minimal governing category.

(b) If P is a non-R-pronoun, \(\alpha\) must be outside its minimal governing category.

Since coindexation is restricted here to only one kind of anaphora, it is not a precondition for coreference. The procedure permits a derivation of examples where antecedents c-command anaphors, but rules out weak crossover cases like (32) automatically because the c-command condition is not met:

\[(32) \quad * \text{His mother loves everyone.}\]

Since the only way a quantifier and a pronoun can be “coreferential” is via bound anaphora, there is no derivation for the intended reading.

One of the primary advantages of this proposal is that it can handle sloppy identity examples like (2):

\[(2) \quad \text{Mary thought she had the mumps and Alice did too.}\]

On the sloppy, or bound variable reading, she is coindexed with Mary. The translated first conjunct then contains the lambda expression shown in (33). The lambda predicate is copied onto the second conjunct, and there Alice provides the value of the variable after lambda-conversion:

\[(33) \quad \text{Mary} (\lambda x (x \text{ thinks that } x \text{ has the mumps})) \land \text{Alice} (\lambda x (x \text{ thinks that } x \text{ has the mumps}))\]

Reinhart claims that disjoint reference has not a syntactic but a pragmatic basis. She points out (following others, e.g. Postal (1969))) that the disjoint reference environments, both for pronouns and for full NPs, closely mirror the environments where bound anaphora is possible. So, where a speaker may use an A-pronoun as a bound anaphor, she may not in general use a B-pronoun to refer to the same entity. And more generally, where she may use a pronoun to indicate bound anaphora, she may not use a full NP, which does not permit that interpretation, unless she has a reason to avoid a pronoun. This account of disjoint reference based on pragmatic principles has a major advantage over the syntactic accounts, in that cases such as Evans’ (3), repeated here, fit in naturally:

\[(3) \quad (a) \quad \text{Only Reagan voted for himself.} \]

\[(b) \quad \text{Only Reagan voted for Reagan.} \]

In (3b) there is reason to avoid bound anaphora, since the semantics of expressions containing only give different truth conditions in the bound and non-bound versions (a) and (b). Thus, the pragmatic disjoint reference strategy is overridden.

Reinhart’s important contribution towards our understanding of anaphora lies in clearly distinguishing bound anaphora from other kinds of coreference. However, there are problems which remain unresolved here.

---

\(^{10}\) The requirement that \(\alpha\) not be immediately dominated by COMP or \(S'\) is intended as equivalent to the requirement that it be in an A-position. Minimal governing category essentially equals governing category ((5) above). She uses the term R-Pronoun where I use A-Pronoun to refer to reflexives and reciprocals.
The first is that Reinhart does not address the "reconstruction" problems, i.e., the cases involving wh-movement and topicalization discussed in 2.1.1. Neither does she address the question of how to treat cases like I like me, where two NPs which are apparently coreferential can yet not be coindexed. As we saw in Section 2.1.2, a wide range of interpretative problems seem to be related to this question. She claims that the cases of mistaken identity fall together with the McCawley/Evans cases as exceptions to the pragmatic disjoint reference strategy, but there seems to be a missing generalization here.

Another problem for Reinhart relates to the pragmatic disjoint reference strategy, as Edwin Williams (p.c.) pointed out to me. This is that it fails to predict a difference in the weak and strong crossover examples in (34a,b):

\[(34) \quad \begin{align*}
(a) & \text{ His mother loves Steve.} \\
(b) & \ast \text{ He loves Steve.}
\end{align*}\]

Both constructions permit bound anaphora, as in (35) and (36):

\[(35) \quad \text{Everyone's mother loves him.}\]

\[(36) \quad \text{Everyone loves himself.}\]

and hence under Reinhart's proposal we would expect that (34a) would be as unacceptable as (34b). I will discuss these examples further in Section 2.2.3 below.

(35) illustrates another important problem with this approach, based on c-command at SS, the treatment of constructions where the relation between a subject-internal possessive NP and a direct object is at issue. In this example, the direct object pronoun may be interpreted by most speakers as a variable bound by the quantifier everyone, but there is no c-command relation between them, and hence they cannot be coindexed by (31). This type of case and related examples involving inverse linking have been taken as evidence that the Binding Theory applies at LF (cf. especially May (1977,1985)). I will consider the problem of inverse linking and possessive NPs at some length in Chapter 4.

2.2 An alternative proposal

The point of departure for the proposal I will present here is the idea that pronominal elements do not have the same referential potential as full NPs. Pronominals are always referentially dependent. In this respect they are analogous to variables in the predicate calculus: they must be bound by some antecedent term in order for the utterance in which they occur to be interpreted.

But binding in natural language is a richer system than that in the predicate calculus. First, there are two kinds of structures which permit binding between an NP and a pronoun. One is established solely intrasententially, making reference in a configurational language like English to the relation of c-command. I will call binding licensed by these structural relations c-command binding. The other kind of structure is the hierarchical discourse structure we considered in Chapter 1, where the possibility of anaphoric relations is dependent upon the relation of accessibility. Such structures license discourse binding. It is important to note that c-command binding and discourse binding are not in complementary distribution: discourse binding may take place in structures where c-command binding is also possible. We will see this in the treatment of the Reinhart examples where both sloppy and non-sloppy readings are possible.

The other central difference between binding in the predicate calculus and in natural language is that in the latter we have more than one kind of individual variable, each with different requirements on the structures in which it can be bound. English A-pronouns, like the disjoint anaphors of Dogrib, must be c-command bound in their governing category, while B-pronouns may either be c-command bound outside their governing category or discourse bound. English thus seems to display a clear complementary distribution. Generally, in a given structure either an A or a B pronoun may be bound, but not both. Again, I follow Heim (1982) in considering deictic pronouns to be discourse bound.

Actually, even the English data are not so clearcut as the Binding Theory would lead us to believe. First, there are a few constructions where the complementarity does not hold. Thus, in (i), the choice of pronoun doesn't make any apparent difference in grammaticality or sense:

\[(i) \quad \text{John pulled the blanket over him/himself.}\]

In other cases, however, although either an A or a B-pronoun is grammatical, the choice makes a subtle difference in meaning. Consider:

\[(ii) \quad \text{John saw those pictures of him yesterday.}\]

\[(iii) \quad \text{John saw those pictures of himself yesterday.}\]

(ii) seems to reflect the speaker's point of view of the pictures, while (iii) reflects John's. This observation seems to be verified by the contrast between (ii) – (iii) and (iv) – (v), where only the verb has been changed:

\[(iv) \quad ? \text{John took those pictures of him yesterday.}\]

\[(v) \quad \text{John took those pictures of himself yesterday.}\]

Here, it seems that the speaker's point of view is somehow less felicitous. Further, consider the case of (vi) – (vii), where the Binding Theory would also lead us to expect only an A-pronoun:
2.2. AN ALTERNATIVE PROPOSAL

(37) Coindex a pronoun \( \alpha \) with an NP \( \beta \) such that \( \beta \) c-commands \( \alpha \).

Conditions:

(a) If \( \alpha \) is an A-pronoun, \( \beta \) must be in its governing category.
(b) If \( \alpha \) is a B-pronoun, \( \beta \) must be outside its governing category.

(Optional)

Unlike Reinhart’s proposal, coindexed NPs are not necessarily treated as variables bound by the same operator. In the mapping from SS to DR, an NP indexed \( i \) will generally be assigned a discourse referent \( x \) with the same index \( i \). The exception is the case where a pronoun is coindexed with a c-commanding NP. Then, whether the two receive the same discourse referent, as in the case of English A- or B-pronouns and their c-commanding antecedents, or a different one, as would be the case with Dogrib disjoint anaphors, depends on the lexical content of the coindexed NPs. Any pronoun which is not c-command bound, and hence coindexed with an antecedent, must find a discourse antecedent, that is, a discourse referent which is accessible to it in the DR. We’ll see examples of this below.

We may now reexamine some of the interpretative problems we discussed in Section 2.1.2 in the light of this new proposal.

As already suggested, the problem with Dogrib is no longer a problem when coindexation merely serves as a diacritic for binding at SS. The disjoint anaphors must be coindexed with a c-commanding NP in their governing category, just like English A-pronouns. But in the mapping from SS to DR their lexical content, perhaps in the form of a feature [+disj] or the like, induces us to choose any accessible antecedent except the NP with which they are coindexed.

In the sloppy identity cases, we may achieve the same results that Reinhart achieved with her distinction between the bound anaphora (sloppy) reading and the stipulated coreference (non-sloppy) reading with the distinction between a reading where the pronoun she in the first conjunct of (2) is c-command bound by Mary and one where she is discourse bound by Mary. We construct DRs for the two readings as follows:

The mapping to a DR is a top-down procedure reducing the original sentence to a structure with a discourse referent for each term and predicates over the discourse referents. We begin with the indexed SS in (38) and develop the DR in (39) in steps.

(38) Mary\(_1\) thinks she\(_1\) has the mumps and Alice\(_2\) does too.

(sloppy reading)
First, in step (a), we introduce a discourse referent in the DR for the subject of the first conjunct, placing on it the condition that it be Mary in any possible interpretation (b). The original sentence then becomes a condition on that discourse referent, (c). Then we reduce the condition further by finding a discourse referent for the pronoun she. But she is already indexed 1, and hence already has a discourse referent, x₁. We reduce the condition accordingly in (d). We now reduce the second conjunct, first introducing the discourse referent x₂ for the subject Alice in (e) and putting the condition (f) on x₂. Since there is no VP, or predicate, on that subject, in (g) we borrow one from the preceding conjunct, replacing all instances of x₁, the discourse referent for Mary, by x₂, the discourse referent for Alice. This DR will have an interpretation along the lines of ‘Mary is an x such that x thinks x has the mumps, and Alice is an x such that x thinks x has the mumps.’

In constructing a DR for the nonsloppy reading in (41), however, we begin from the SS in (40), where Mary and she are not coindexed.

\[
\begin{array}{ll}
(39) & \begin{array}{ll}
(a) & x_1 \\
(b) & \text{Mary}(x_1) \\
(c) & x_1 \text{ thinks she}_1 \text{ has the mumps} \\
(d) & x_1 \text{ thinks } x_1 \text{ has the mumps} \\
(f) & \text{Alice}(x_2) \\
(g) & x_2 \text{ thinks } x_2 \text{ has the mumps}
\end{array}
\end{array}
\]

\[
\begin{array}{ll}
(41) & \begin{array}{ll}
(a) & x_1 \\
(b) & \text{Mary}(x_1) \\
(c) & x_1 \text{ thinks she}_2 \text{ has the mumps} \\
(e) & x_1 \text{ thinks } x_2 \text{ has the mumps} \\
(f) & x_2 = x_1 \\
(h) & \text{Alice}(x_3) \\
(i) & x_3 \text{ thinks } x_2 \text{ has the mumps}
\end{array}
\end{array}
\]

When we reach the stage in the construction where we need to reduce the VP containing she, after step (c), we introduce a new discourse referent for the pronoun (d), since there is no discourse referent x₂ already in the discourse. But since she is a pronoun, it must be bound. We do this by equating the discourse referent x₂ with another, accessible discourse referent. The discourse referent for Mary, x₁, is one possible antecedent, so we equate them (f) (though, of course, there might be other readings where she refers to an even earlier discourse referent). In a model theoretic interpretation, the equation assures that the two terms will be coreferential. In steps (g) – (i) we treat the second conjunct as before, replacing all instances of x₁ by x₃. But the result here differs from that in (39) because x₂ is already bound to the discourse referent for Mary, and so we get the reading where Alice thinks that Mary has the mumps.

I said above that coindexation would not be uniformly interpreted in this approach. It now no longer means coreference, but only indicates that a pronoun is c-command bound by an NP. Notice that non-coindexation does not mean disjoint reference, either. In the case of plural pronouns with split NP antecedents, we may assign distinct indices to the individual NPs involved, then permit the discourse referent for a plural pronoun to be equated with a set of indices of other accessible discourse referents, in the spirit of Lasnik’s proposal. (42) is an indexed SS for (30) above, while a partial DR is shown in (43):

\[
\begin{array}{ll}
(40) & \text{Mary}_1 \text{ thinks she}_2 \text{ has the mumps and Alice}_3 \text{ does too.}
\end{array}
\]

\[
\begin{array}{ll}
(42) & \text{John}_1 \text{ told Bill}_2 \text{ that they}_3 \text{ should leave.}
\end{array}
\]
2.2. AN ALTERNATIVE PROPOSAL

This approach to anaphora provides a means of expressing more precisely the fact that in interpreting anaphoric coreference what seems to be at issue is not coreference in the strictest sense, but something vaguely like coreference-under-a-description. It is not having the same or a different referent in the world which is at issue. Instead, the indices on two given NPs indirectly indicate whether they have the same or different discourse referents. Much here depends on the speaker's intentions. In explaining the Jesus example, (20), and example (23), we need only note that there is no requirement that distinct or non-overlapping discourse referents map onto distinct or nonoverlapping referents in a model. Thus I and me, we and me can have distinct indices at SS, and be mapped onto distinct discourse referents, but still be mapped onto the same referent or overlapping referents in the model into which the DR is embedded (in this case the mapping onto referents will be in virtue of the lexical content of these indexical pronouns). Similarly, in mistaken identity cases such as (25), The man in the sunglasses resembles Lou, two NPs may be indexed differently and have different discourse referents because the speaker does not know they actually refer to the same individual in the world.

Having sketched the basic properties of my proposal, now I will turn to consider how it might be extended to deal with three of the major issues in the theory of anaphora, reconstruction, disjoint reference, and crossover.

2.2.1 Reconstruction

We saw in Section 2.1.1 how the problem of where to apply the Binding Theory revolved around examples which seemed to require reconstruction or something like it. To summarize the problem, the example in (12), repeated below, shows that in some cases reconstruction of a moved element into its base-generated position seems necessary to account for the use of an A-pronoun. But (16b) shows that in other instances reconstruction would give infelicitous results from the point of view of Principle C. And Pesetsky's example (18a) is a case where reconstruction is both obligatory, for the binding of himself by John, and infelicitous, since then she would c-command Mary. We need to have our cake and eat it.

(12) Which picture of himself does John like

(16) (b) Which picture of John does he like

(18) (a) [Which picture of himself that Mary likes] will John give her,

The crucial factor in these cases is the relation between the preposed element and its trace in the matrix clause. I propose that we use this relation without

(43)

\[
\begin{array}{c}
\begin{align*}
& x_1 \ x_2 \ x_3 \\
& \text{John}(x_1) \\
& \text{Bill}(x_2) \\
& x_3 = \{x_1, x_2\}
\end{align*}
\end{array}
\]

Consider again example (1), repeated here with indices:

(1) [Alan\textsubscript{1} and Margaret\textsubscript{2}]\textsubscript{3} ate [their\textsubscript{3} dinner]\textsubscript{4}.

Apart from the question of how to distinguish between the so-called group and distributive readings of such examples, which I will consider in detail in Chapter 3, on the approach I am sketching here there is no problem with the anaphoric relations indicated. Even though Alan and Margaret are not coindexed with the NP which contains them, this does not mean they are disjoint from it in reference. Non-coindexation does not imply non-coreference. Rather, the discourse referents for the proper names will stand in a constitution relation to the discourse referent for the whole subject NP, as shown in the DR in (44):\textsuperscript{14}

(44)

\[
\begin{array}{c}
\begin{align*}
& x_1 \ x_2 \ x_3 \ x_4 \\
& \text{Alan}(x_1) \\
& \text{Margaret}(x_2) \\
& x_3 = \{x_1, x_2\} \\
& x_3\text{\ 's dinner}(x_4) \\
& x_3 \text{ ate } x_4
\end{align*}
\end{array}
\]

Each of the proper names, as well as the full subject, induces the introduction of a discourse referent, and the condition on the discourse referent for the full subject \(x_3\) specifies that its reference is the set of the entities referred to by the conjuncts. The plural pronoun their, coindexed with the full subject, then automatically receives the same discourse referent, and the truth conditions amount to 'the set consisting of Alan and Margaret ate its dinner.'

\textsuperscript{14}See Chapter 5 for a different way of representing conjoined subjects in DRs.
2.2. An Alternative Proposal

An alternative proposal to the definition of coindexing in (37) is to redefine the procedure to optionally treat preposed NPs as if they were reconstructed. In (12), we see that the VP node of the matrix contains *himself* via the trace of the preposed *wh*-element. Note that the definition of governing category given by Chomsky in LGB ((5), above) already includes the term contain, so in each of the cases under consideration, the governing category for *himself* is the matrix sentence. Hence, condition (a) of the coindexing procedure (45) is satisfied.

Now consider (16b). If we attempt to coindex *John* with *he*, we will fail because although he c-commands *John*, *John* is not a pronoun, and *John* does not c-command *he*. How then can the two NPs be coreferential? Via discourse binding. In order to see how this works, let us construct DRs for the topicalized counterparts of (12) and (16b).\(^{16}\)

A preposed element is the first entered in a DR. In the case of the topicalized counterpart of (12), (47), this element contains an A-pronoun, and the mapping procedure recognizes that this must be coindexed with another NP in its governing category. Let us assume as a matter of technical implementation that the discourse referent of an A-pronoun is initially marked with a *. Then, when and if

---

\(^{15}\)In this definition, the specification that the trace be case-marked is designed to avoid reconstructing NP-movement cases.

\(^{16}\)The *wh*-examples are somewhat more complex for reasons that are irrelevant here, but would work along the same lines.
2.2. AN ALTERNATIVE PROPOSAL

I will adopt Reinhart's general pragmatic approach to the disjoint reference problem, but note that with the distinctions I have drawn among types of binding and types of pronouns, we can develop a finer scale of binding strength than was possible in her treatment. In (54) is a list of the three pragmatic grades of binding, distinguished by the degree of ambiguity they permit:

(54) Three Pragmatic Grades of Binding:

(a) c-command binding of A-pronouns
   ex: Zelda saw herself.
(b) c-command binding of B-pronouns
   ex: Zelda thought she saw a mouse.
(c) discourse binding (of B-pronouns only)
   ex: Annie told us about the surprise party.
   The flowers in Zelda’s room pleased her.

C-command binding of A-pronouns is the strongest, since it is least likely to be ambiguous. With c-command binding of B-pronouns we are less sure of what binds the pronoun; consider the example in (b) in a context following The cat crept silently toward the barn, where she might be taken to be discourse bound by the cat instead of c-command bound by Zelda. And discourse binding is notoriously full of potential for ambiguity, as in the illustration in (c), where her can easily refer to either Annie or Zelda.

One can define the binding potential of two positions in a given syntactic structure as the strongest kind of binding permitted there. Observance of the Gricean cooperative principle in conversation (Grice (1967)) leads one to use the strongest means he has to make the identity of referents unambiguous. I offer, then, a slight revision of Reinhart’s pragmatic approach to disjoint reference:

(55) Pragmatic Disjoint Reference Strategy:

(a) Speaker’s Strategy: Use the strongest binding potential of the structure you are using, unless you have reason to avoid binding.
(b) Hearer’s Strategy: If the speaker doesn’t take advantage of the strongest binding potential of the structure she is using, then, unless she has reasons to avoid binding, she doesn’t intend her expressions to corefer.

As in Reinhart’s original proposal, these pragmatic strategies contain an unless clause which accounts for the acceptability of examples such as (3b), despite their apparent violation of disjoint reference. Further, in this version of the prag-
matic account, we might expect that the stronger the binding potential of a given structure, i.e. the greater the opportunity to avoid ambiguity, the more difficult it is to avoid binding without leading the hearer to assume disjoint reference, in line with the Strategy. We see this in the differential acceptability of (53a) and (53c). We may also use this to provide an account of the differential acceptability of (34a) and (b), an account which we noted that Reinhart’s original proposal did not offer:

\[(34)\]
\[
\begin{align*}
(a) & \quad \text{His mother loves Steve.} \\
(b) & \quad \text{* He loves Steve.}
\end{align*}
\]

The structural relation between the pronoun and NP in (b) has the strongest binding potential, since c-command binding of an A-pronoun would be possible here, as in *Steve loves himself*. Hence it is a strong violation of (55). But, however we account for the possibility of anaphora in (a), c-command binding of an A-pronoun in this example is not possible (*John’s mother loves himself*), so no violation of (55) is apparent. However, note that the indexing procedure in (45) would not coindex his and Steve in this example. This leads us to discussion of the crossover problem.

First, however, I want to point out a problem with this approach to disjoint reference. In the preposed NP examples, we encounter cases like:

\[(56)\]
\[
\text{Which picture of John does he like?}
\]

\[(57)\]
\[
\text{Which picture of himself does John like?}
\]

The problem here is that we know that c-command binding of A-pronouns is possible between the positions indicated, as in (57). But given that the strongest binding potential of this structure is realized in (57), why isn’t that sentence preferable to the one in (56), where we see only discourse binding? I believe that there may be discourse constraints on the types of binding available in such examples. This may be related to Lakoff’s examples where it is claimed that for some speakers “backwards” anaphora is preferable to “forwards”:\(^{17}\)

\[^{17}\text{I am not able to find a reference for this contrast, although Stockwell et al. (1973, p.196) cite Lakoff (1968) on the related problem of the differential acceptability of (i) and (ii):}

\[
\begin{align*}
(i) & \quad \text{Near him, John saw a snake.} \\
(ii) & \quad \text{Near John, he saw a snake.}
\end{align*}
\]

I am also able to imagine appropriate discourse contexts for (ii), but as with (58), it seems far less felicitous out-of-the-blue than (i) and seems to require a contrast between

\[(58)\]
\[
\text{? In John’s room he smokes pot.}
\]

\[(59)\]
\[
\text{In his room John smokes pot.}
\]

However, it seems to me that (58) is perfectly felicitous in an appropriate context. Consider the following, for example:

\[(60)\]
\[
\text{I spoke yesterday with several students who have been flagrantly violating the dormitory regulations. In Mary’s room there is a large refrigerator and a hot plate. In John’s room he smokes pot. Steven...}
\]

It would seem inappropriate in light of (60) to characterize (58) as ungrammatical. Rather, it seems that the problem here is one of how to characterize the contexts where such a sentence is felicitous and to explain why one is less likely to utter it out of the blue than (59).

2.2.3 Crossover

Consider the classic crossover cases in (61) and (62):

\[(61)\]
\[
\begin{align*}
(a) & \quad \text{* who, does he like John?} \\
(b) & \quad \text{he likes everyone} \\
(c) & \quad \text{he likes John} \\
(d) & \quad \text{he likes John} \\
(e) & \quad \text{he likes JOHN}
\end{align*}
\]

\[(62)\]
\[
\begin{align*}
(a) & \quad \text{* who, does his mother like John?} \\
(b) & \quad \text{his mother likes everyone} \\
(c) & \quad \text{his mother likes John} \\
(d) & \quad \text{his mother likes John} \\
(e) & \quad \text{his mother likes JOHN}
\end{align*}
\]

The (a) and (b) examples for both structures are considered totally unacceptable. The reason for this in this framework is twofold. First, neither the trace \(t_i\) nor the wh-element which binds it, nor the quantifier everyone is a pronoun. Hence

John and other salient referents.
they cannot be coindexed with the c-commanding NP by the indexing procedure (45). Second, discourse anaphora is impossible because who and everyone do not actually refer, as does, e.g. a proper name, and hence it is not possible that they refer to something already salient in the discussion which might then be accessible to he.

Now consider the cases with proper names in (c), (d) and (e). (61c), with non-contrastive stress on John is just plain bad. He and John cannot be coindexed in this or any of the remaining cases, since John is not a pronoun. We get the same unacceptable result in the weak crossover case in (62c) as noted by Chomsky. But the differential binding potential of the two structures explains the difference between the unacceptable (61d) and the acceptable (62d). In these cases, it is not that likes is focused. Rather, John being already salient in the discourse, the proper name John is unstressed, like a pronoun, and likes receives default stress. Hence, though he and John can’t be coindexed, the discourse referent for he in (62d) can be equated with that of the previous occurrence of John. Though this suffices to overcome the disjoint reference strategy in the weak crossover construction, it does not in the stronger case of (61d). It is only contrastive stress, requiring a particular type of preceding discourse, which overcomes the disjoint reference strategy in the strong crossover case (61e), as well as in the weak (62e). These are the types of cases Evans discussed. An appropriate discourse might be as in (63):

(63) **Speaker A:** John doesn’t like anyone.
    He doesn’t like Al.
    He doesn’t like Sam.
    He doesn’t like Ginger.

**Speaker B:** But he likes JOHN.

In the DR constructed for this discourse, successive occurrences of he are given individual discourse referents, each of which is equated with that for the first occurrence of John. The final, contrastive occurrence of John receives a different index from that of the first occurrence and the pronouns. The two tokens of

18Chomsky (1976,1980b) discusses this fact about the weak crossover cases and compares it with the acceptability of examples like (62d). His analysis, however, is very different.

19See Ladd (1980) for discussion of this phenomenon, wherein a pronoun or NP which would normally receive nuclear stress is unstressed (Ladd’s “destressed”) because it is old information, and the nuclear stress moves onto the constituent immediately to its left.

20By ‘contrastive stress,’ I mean functionally contrastive, in the spirit of Culicover & Rochemont (1983). The phonological distinction between (c) and (e) is one of pitch accents (see Pierrehumbert (1980)), not one of degree of stress. (c), for example, might end with a falling boundary tone, (e) with a rising.

21She adds (p.198), “In general only A-positions are visible for scope indexing and binding ... Thus, the rule of Scope Indexing takes into account not the wh-word, but its trace.”
(67) CONDITION ON VARIABLES

(a) \( \text{Pro}_i \) must be c-commanded by \( \text{NP}_i \), if \( \text{NP}_i \) is an inherent quantifier.

(b) \( \text{Pro}_{ij} \) must be c-commanded either by \( \text{NP}_{ij} \) or by \( \text{NP}_j(i) \) (Indirect Binding).  

By (64), the scope of an NP relative to other NPs is given by the use of slash indices, subject to the constraints in (65). Consider one of her examples:

(68) (a) Two men, wrote to a woman_{ij}.

(b) She didn’t reply.

Here, a woman in (a) has narrow scope with respect to two men. The pronoun she in (b) cannot be coindexed with a woman because the condition on variables, (67b) requires that a pronoun with slash indices be c-commanded by its antecedent, that is, in our terms, that it be c-command bound.

Haik accounts for universal donkey sentences by indexing them as follows:

(69) \([\text{NP}_{-1(2)} \text{every farmer} \ [\text{who}_{2} \text{t}_{1(2)} \text{owns} \ [\text{NP}_{-2(1)} \text{a donkey}]]) \text{beats } it_{2/1}\)

The NP a donkey is first indexed 2, the subject trace in the relative clause 1. The trace takes wide scope over a donkey, so that the index of the latter becomes 2/1 by (64a), and the index of the former becomes 1(2) by (66). She assumes that the subject NP automatically receives the same index as the trace of its relative clause, so that every farmer who owns a donkey is indexed 1(2). The subject c-commands the pronoun it, and hence, by (67b), the pronoun may receive the index of a donkey, 2/1.

Haik accounts for weak crossover by requiring that it is not wh-moved elements.

22This version of the Condition on Variables, her (61), p.203, is later superseded by her (84), p.211, given in (67) below:

(67') CONDITION ON VARIABLES

(a) \( X \) binds \( Y \) (directly or indirectly) only if \( X \) c-commands \( Y \) at S-Structure.

(b) \( \text{Pro}_{ij} \) is licit if c-commanded by \( \text{NP}_{ij} \).

In this version, (67a) and the indirect binding disjunct of (67b) are combined into (67'a), and the remainder of (67b) becomes (67'b). I have used the earlier version (67b) for expositional clarity.

23This version of the Condition on Variables, her (61), p.203, is later superseded by her (84), p.211, given in (67) below:

24Compare infinitival Rationale Clauses (in order to ...), which seem to play the same adverbal role semantically. See Jones (1985) for extensive discussion of these. Among other things, these are also sentential adverbs and require control by an agentic subject.
These grammaticality judgments are predicted by Haïk's system, but it is not clear that the motivation for the ungrammaticality of (72) is correct. Notice that in similar structures, anaphoric reference to the matrix subject out of this type of adverbial clause is not in general possible where that subject is under the scope of another argument in the matrix clause:

(73) * Four people$_{1/2}$ made a tape in every language$_{2}$ so I could study their$_{1/2}$ dialects.

(73) is unacceptable with every language taking wide scope over four people and their bound by four people. The descriptive generalization seems to be that NPs in the matrix clause are only accessible as antecedents to pronouns in the adverbial clause when the (agentive) matrix subject has wide scope. Thus, it is not clear that these examples argue for Haïk's proposals.

Haïk points out other cases of asymmetry which involve the subject and object of the relative clause in donkey sentences:

(74) [Some people who $t_1$ kicked a donkey$_{2/1}$]$_{1(2)}$ hated it$_{2/1}$

(75) ?? [Some people who a donkey$_{2/1}$ kicked $t_1$]$_{1(2)}$ hated it$_{2/1}$

She claims that her theory predicts the difference in acceptability between (74) and (75); however, I cannot see how. In both examples, the trace in the relative clause has wide scope over a donkey. The 'backwards scope' we see in (75) should certainly be licit in wh-structures generally. Consider (76):

(76) Who$_1$ does every dog$_{2/1}$ like $t_1$

The reading indicated, where who has wide scope over every man and an appropriate answer might be "our mailman," is felicitous. Here, the indexed structure is parallel to that of the relative clause in (75). Also, in both (74) and (75), the subject, indexed 1(2) by (66), c-commands the object, indexed 2/1, and hence, by (67), the subject should be able to indirectly bind the pronoun.

Haïk points out that the explanation for the unacceptability of (75) cannot be found in the ECP (Empty Category Principle) of Kayne (1981) (presumably under the assumption that the indefinite a donkey would undergo QR, leaving a trace in subject position). She offers the following examples as evidence:

(77) Some people who believed that a donkey was waiting in the courtyard called to it through the window.

(78) ?? Two donkeys that Mary gave a carrot to ate it.

In (77) a donkey is in subject position, as in (75), so if the ECP was responsible for the unacceptability of the latter, we would expect (77), on the reading with a donkey narrow under the subject of the relative clause, to be unacceptable as well. However, (77) seems fine on this reading. In (78), the narrow scope indefinite, a carrot in this example, is not in subject position, yet the unacceptability seems similar to that of (75). Haïk argues that this is due to lack of c-command of the indefinite by the trace. Note, however, that the indirect object-marking preposition to is the sort most plausibly treated as a case marking, rather than a true proposition, so that we might consider the trace to c-command the direct object a carrot after all. Thus, although (77) shows that the ECP does not offer an explanation of the problem, (78) shows that c-command may not be the answer either.

Steve Berman and Karina Wilkinson (in work reported in Berman (1985)) have suggested a different view of the unacceptability of (75). They propose an account which retains the core of Heim’s (1982) account in FME Change Semantics, but adds to it an S-Structure constraint in terms of Kayne’s (1983) Connectedness. Briefly, this constraint rules out coindexing the indefinite NP and the pronoun unless their g-projection sets form a subtree (in Kayne’s (1983) sense of g-projection sets). They show how this accounts for (75) and a variety of other structures. Since something like Connectedness (or Koster’s (1985) closely related Global Harmony requirement) does seem to play a role in a variety of types of structures, including also parasitic gaps, this proposal seems plausible.

A final range of examples which Haïk’s proposal treats involve backwards bound anaphora. These involve “Jacobson’s Sentences” (referring to Jacobson (1979)), as in (79) and (80), and the Bach-Peters (see Bach (1968), (1971)), or Crossing Coreference Sentences, as in (81) and (82):25

(79) Everyone$_1$ told [her$_{2/1}$ mother$_{3/2/1}$ that [his$_1$ wife$_{2/1}$] should get a job.

(80) * [Her$_{2/1}$ mother$_{3/2/1}$ told everyone$_1$ that [his$_1$ wife$_{2/1}$] should get a job.

24 See Bach & Partee (1980) for discussion of how such “empty prepositions” argue for a function-argument approach to binding over structural conditions such as c-command.

25 See Jacobson (1979) for extended discussion of these.
CHAPTER 2. ANAPHORA, COREFEERENCE, AND BINDING

2.3 INDIRECT BINDING

(84) EXTENDED NAME CONSTRAINT

If NP, is not within the scope of NPj (that is, NP1 ≠ NPj), then NP1 must be closed with respect to NPj; that is, NP1 may not contain a free variable indexed j.

This condition is intended to account for a variety of facts. Among them are that in cases such as (85), the quantified subject NP must have wider scope than an object which contains a pronoun bound by it, and that wh-extraction is not possible from a definite NP, as in (86):

(85) Everyone, likes some film he; saw.

(86) * Who; did you like that picture of t;

This constraint is plausible; however, note that it is not necessary under a Tarskian semantic interpretation. Under such an interpretation, the pronoun in the wide scope object in (85) would denote whatever was assigned to the i-th variable by the assignment function used in its interpretation. Since the interpretation of the subject ultimately does not rely on the values assigned by that assignment function, but ranges over all assignment functions just like it except for the value assigned to the i-th variable, then the interpretation of the pronoun and that of the quantified subject would be independent. In a discourse theory such as Kamp's or Heim's, one requirement on definite NPs, including pronouns, is that they have an accessible antecedent. Since the subject of (85) would not be accessible to the pronoun in a wide-scope object, the pronoun would be without an antecedent, and the example would be felicitous.

Polly Jacobson (p.c.) points out an important problem with Haïk's account of crossing coreference examples. Haïk predicts that (87) should be fully acceptable with the coreference indicated, by the same mechanism which coindexes it and the direct object in (82):

(87) * Every farmer that owns it beats a donkey.

However, this example cannot be interpreted in this way. (87) shows that the binding of a subject internal pronoun by an object is far more restricted than the binding of an object internal pronoun by the subject. Although Jacobson's (1979) account is, as Haïk claims, inadequate to deal with crossing coreference under quantification, it does treat the two pronouns differently, and thus foreshadows the asymmetry which we see between (69) and (87).

I will not attempt to develop here an account of bound backwards anaphora, or of backwards anaphora more generally. I think the phenomena are complex.

In (79) everyone c-commands, as well as taking wide scope over, both his wife and her mother. Thus, although his wife does not c-command her in her mother, everyone may indirectly bind her, so that his wife and her may be coindexed. But in (80) everyone does not c-command her, and hence cannot indirectly bind it.

Haïk shows that previous attempts to account for the Crossing Coreference sentences have been inadequate. Jacobson (1979) had argued that the pronoun in the relative clause of the subject of sentences such as (81) is derived by replacement of a full NP, the woman who wrote to him, identical to the object which seems to bind the pronoun. However, this approach could not explain examples such as (82), where the pronoun in the first relative clause is coindexed with a quantified direct object; simply replacing the pronoun with a quantified NP identical to the object would not give the same truth conditions as the bound reading. Higginbotham & May (1981) proposed to account for examples such as (82) by means of an operation called Absorption, where two quantifiers which have been raised at LF become one binary quantifier, with mutual c-command between the two original quantified NPs. This mutual c-command then permits each NP to bind a pronoun in the relative clause of the other. Haïk argues that Absorption is too powerful an apparatus. She points out that in the Crossing Coreference sentences, the NP which is an "inherent quantifier" generally has to c-command the other NP, as the NP with the determiner every c-commands the indefinite in (82). When this is not the case, as in her (83), the example is unacceptable, even though absorption would predict grammaticality:

(83) * [Some child who had heard it2/1]1/2 believed [every story that was told to him1]2/1

Haïk's account of the Crossing Coreference Sentences is based on Indirect Binding. In (82), it in the relative clause of the subject is indirectly bound by the subject wh-trace and him is directly bound by NP1, which c-commands it. She can also account for the impossibility of the binding relations shown when the object has wide scope under the subject, by the use of a principle which she calls The Extended Name Constraint, extending Gueron's (1981) Name Constraint:
CHAPTER 2. ANAPHORA, COREFERENCE, AND BINDING

Peter Sells (p.c.) has noticed examples such as (88), where a wide scope object binds a pronoun inside the subject:

\[ (88) \quad \text{A plaque indicating the date of its incorporation may be found in every American city.} \]

Since this example seems acceptable, Haik's account of the asymmetry between (83) and the acceptable Crossing Coreference sentences (81) and (82) in terms of the c-command required for indirect binding seems questionable, although I agree with her that Absorption is too strong a mechanism. Also, although many people seem to find the crossing coreference reading of (82) acceptable, none of the informants I have discussed the example with found this was the first reading (preferring one where the first pronoun is interpreted as discourse bound by some unspecified antecedent — “it, whatever it is”), and many, including the author, find the crossing coreference reading unacceptable here. When judgments are so fuzzy, it seems unlikely to me that we should use the same mechanism for binding which gives us the readily interpretable bound readings of the donkey sentences.

To summarize the discussion so far, Haik claims that her proposal can handle donkey sentences at S-Structure without an additional level of representation or interpretation, that her account in terms of c-command predicts certain subject-object asymmetries, and that it can handle Jacobson's and Crossing Coreference sentences. We have already seen that the asymmetries and the backwards bound anaphora in the last examples seem more complex than Haik has considered, and that although configurational characteristics of S-Structure may be involved, they seem to involve more “global” structural characteristics, in terms of Connectedness or Global Harmony, and not simply c-command. And although no further levels beyond S-Structure are required to handle the range of data Haik considers, we saw in Chapter 1 that modal subordination in general is crucially a discourse phenomenon, involving accommodation and other pragmatic factors, and cannot be reduced to characteristics of the S-Structure of individual sentences. Thus, her approach would ultimately prove inadequate to handle modal subordination, of which the conditional donkey sentences seem to be special cases.

Other problems with Haik's approach arise from her use of indices on NPs to indicate relative scope. How are her S-Structures, with slash indices to indicate relative scope of NPs, to be interpreted? If NPs are interpreted as generalized quantifiers, then they must take as arguments property-denoting elements. One way to do this in Haik's framework would be to treat the complexly indexed S-Structures as indications of (possibly iterated) predication involving lambda abstractions. Where an NP a has wide scope over another NP b, then a would be taken as the subject of a lambda abstract which includes b. Thus, a simple sentence such as (89) would have an interpretation corresponding to the more abstract (90):

\[ (89) \quad \text{Everyone}_1 \text{ likes someone}_2/1 \]
\[ (90) \quad \lambda x[x \text{ likes someone}] \text{ (everyone)} \]

In a limited range of cases, such as the simple (89), the interpretations which are derived in this way correspond to the readings predicted by May (1977), or by the system of van Riemsdijk & Williams (1981) or Williams (1986), where quantifier scope is indicated at S-Structure by coin-indexing an NP with the higher constituent over which it has scope. However, more complex examples pose problems for this way of interpreting Haik's system: what determines the limit of the range over which an abstraction may be formed on such an approach? In the extensive literature on quantifier scope, various arguments have been offered for quantifying in, or quantifier raising, or quantifier indexing not only at S, but at VP and NP, and perhaps even CN. (See Chapter 4 for extended discussion.) Nothing in Haik's system gives us the means to distinguish between the various scope possibilities of a given NP in this way. Further, her generalization (65), that operator scope is clause-bounded, admits of many counterexamples, including those involving NPs in the sentential complements of verbs.

But the crucial problem with her proposal for the representation of quantifier scope is that the relative scope of two NPs is not a question of a direct relation between the NPs themselves, as it appears in Haik's system, but is indirect: if an NP a occurs in a property denoting constituent (syntactic or abstracted) which is predicated of another NP b, then we say that a is 'in the scope of' b. But anaphoric relations between NPs appear to be direct: 'is the antecedent of' is a direct relation between two NPs without the mediation of predication. Thus, the mechanism by which Haik proposes to collapse the representation of scope with the representation of bound anaphora suffers from a fundamental conceptual confusion about the different kinds of relations involved.
Chapter 3

Distributivity

There are a variety of issues which will concern us in this chapter, but they all revolve around the phenomenon of distributivity in English, with special attention to how it affects the anaphoric potential of NPs. This phenomenon is illustrated in the following examples:

(1) Four men lifted a piano.
(2) Bill, Pete, Hank, and Dan lifted a piano.
(3) Each man lifted a piano.
(4) Bill, Pete, Hank, and Dan each lifted a piano.

(1) and (2) are each ambiguous. For example, on the so-called 'group reading' of (1), the men in question lifted a single piano together, but it may be that none of them lifted it alone. The group reading of (2) is similar. On the distributive reading of (1) or (2) each of the men has the property of having singlehandedly lifted a piano. If we fix the scopes of the NPs in (3) so that the subject has wide scope over the object, then it is unambiguously distributive — there is no reading where the group composed of all the men together lifted a piano. Likewise, (4), with adverbial each, is unambiguously distributive. If the context in which (3) is uttered is such that we know that there were just four men, Bill, Pete, Hank, and Dan, then the readings of (3) and (4) with the subject taking wide scope are true in just the same situations.

Now compare the distributive readings of (1) and (2), and the readings of (3) and (4) where the subject NP has widest scope. All share two properties: The first is that, despite the fact that the indefinite direct object is singular, a piano, there may be as many as four pianos involved in any situation in which one of these examples is taken to be true—it may be that each man lifted a different
piano. Of course, even on the group readings of (1) and (2), the meaning of the indefinite NP is generally taken only to put a lower bound on the number of pianos involved; but the use of the singular on this reading is usually taken, via the Gricean maxim of quantity, to imply that so far as the speaker knows only one piano was involved (see, e.g. Horn (1972), Levinson (1983), Hirschberg (1985)). Hence, the distributive/group ambiguity is associated with a possible difference in the cardinality of the set of pianos involved.

The cardinality effect is closely related to the second characteristic which these examples share: a piano may not serve as an antecedent for subsequent anaphors in discourse. This is shown by the infelicity of (5) following (1), (2), (3), or (4) (on the relevant distributive readings), with it intended as anaphoric to a piano:

\[(5)\]

It was very heavy.

But there is a difference in the anaphoric potential of the subjects of (1)–(4) on the relevant distributive readings. The subject of (3) may not license anaphora in discourse, so that (3) followed by (6), with he anaphoric to each man, is infelicitous:

\[(6)\]

He developed a crick in his back later.

This reflects a well-known property of quantificational NPs, that they only license what I have called c-command anaphora. However, the subjects of (1), (2) and even the unambiguously distributive (4) may license discourse anaphora. Consider any of these examples followed by (7):

\[(7)\]

They then gathered to receive the promised award.

They refer to four men or Bill, Pete, Hank and Dan. (3) may also be followed by (7), but notice the lack of number agreement between the subject of (3) and the pronoun in (7). We most often use a NP such as each man where there is a group of men already salient in the discourse, and it would then be this group which served as antecedent for they, and not each man. In any case, I will argue in Chapter 5 that the use of a universally quantified NP may in general license the accommodation of a plural discourse referent. It is thus a plural discourse referent, already salient or accommodated, which would serve as discourse antecedent of they in this case.

Notice that the verb gather in (7) is generally regarded as a group predicate; that is, it is only felicitously predicated of a subject NP with a group reading. They, then, has a group interpretation, even when its intended antecedent, such as the subject of (1), (2), or (4), occurs in a sentence with a distributive interpretation.

There are other examples which show that the subjects of (1) and (2), under either the group or the distributive interpretation, or the subject of (4) may serve as antecedent to a plural pronoun interpreted distributively. This is the case in (8):

\[(8)\]

They developed a crick in their backs later.

In general, developing a crick in one’s back is something only an individual can do. The singular indefinite direct object in (8), then, leads us to the distributive interpretation: we are concerned with a property of each of the three men denoted by they, not of the group as a whole. Another point of interest in this example is the plurality of their backs. This seems to be an example of the dependent plural phenomenon, wherein an NP with plural morphology seems to receive the interpretation of its corresponding singular form. Note that although this complement to the head crick is plural, we do not understand the direct object as being about a single crick which is in all three backs. Rather, on the usual interpretation there is a different crick in each back. The dependent plural here is optional for many speakers, for whom the singular their back is equally acceptable.

Summarizing, it appears that (1) and (2) are ambiguous with respect to the group/distributive distinction, while (3) and (4) are not. The distributive readings of these examples share two properties: the cardinality effect and the inaccessibility of the indefinite direct object to serve as a discourse antecedent. However, the subjects of (1) and (2), under either the group or the distributive reading, and the subject of (4) may serve as discourse antecedents, while the subject of (3) may not.

How are we to explain these facts? In what follows, I will argue for a very simple theory of distributivity and the distinction between distributive and group readings. Distributivity is a property of predications, combinations of a subject and a predicate. Note that by "a predication" I do not mean to restrict attention to the relation between a syntactic subject and its VP. The NP in a predication may be a non-subject, and the predicate may not be the syntactic VP of the sentence uttered, but may be derived instead via lambda abstraction. The distributive reading may be triggered either by a quantificational determiner in the subject NP or by the presence of an explicit or implicit adverbial distributivity operator on the predicate. If a plural subject is nonquantificational (doesn’t contain a quantificational determiner), we say that it is group denoting. A group reading arises when neither a group-denoting subject nor an adverbial element of its predicate contributes the quantificational force underlying distributivity. In cases where the subject of a predication is group-denoting, the corresponding discourse referent may serve as a discourse antecedent even if the predicate is modified...
by an adverbal distributivity operator. But adverbal distributivity introduces a universal quantifier; any indefinite which occurs within the modified predicate will be under the scope of this universal quantifier, and hence the discourse referent associated with the indefinite will be anaphorically inaccessible to subsequent discourse. In Chapter 4, I implement the theory of distributivity argued for here, to show how the facts about anaphora and distributivity fall out of these simple assumptions in conjunction with the assumptions about anaphora in discourse discussed in the preceding chapters and with a theory of the representation of quantifier scope.

In developing this theory, I will begin with a detailed review of the three principal approaches to distributivity in the literature on plurality and quantification. Then I will turn to the exploration of the main aspects of my proposal, the semantics of groups and of plurality more generally, adverbal distributivity, and the distinction between group denoting and distributive, or quantificational, NPs. Then I will address two kinds of issues which an adequate theory of distributivity must address, one pertaining to nonconditional donkey sentences and the other the question of the analysis of dependent plurals. Finally, I will draw a few conclusions about the general character of distributivity.

A note on my goals in this chapter: The data pertaining to the analysis of distributivity are quite complex, and judgments about interpretation are sometimes subtle and slippery. Further, as a prerequisite to a fully adequate theory of the group/distributive distinction, we require adequate theories of phenomena such as reciprocals, partitives, events, the generic bare plural, and dependent plurals, as well as analyses of the lexical content of a variety of determiners, all of which are quite complex and difficult in themselves. Linguistic theory at present cannot provide us with all of these requirements, and, as a consequence, the theory of distributivity which I offer here cannot be the final word on the subject. My goal is to discover criteria of adequacy for a theory of distributivity, and then to develop a theory which meets as many of these criteria as possible. In order to do this, I will have to make certain decisions about the treatment of dependent plurals, and the like, and I will try to indicate where there is not yet firm evidence to support the particular choices I have made.

3.1 Previous approaches to the group-distributive distinction.

3.1.1 Distributivity as quantifier lowering/raising

Lakoff (1970, Section III) is an early discussion in the linguistic literature of the distributive/group distinction, a distinction which Lakoff discusses in terms of a group reading vs. a 'quantifier-reading' in examples such as the following:

(9) That archaeologist discovered nine tablets.

(10) All the boys carried the couch upstairs.

(11) Every boy carried the couch upstairs.

(12) That archaeologist discovered few tablets.

Lakoff claims that (9) and (10) are ambiguous: On the group reading, the archaeologist discovered nine tablets as a group, and the boys carried the couch upstairs as a group. On the quantifier reading (what I have called the distributive reading), there are nine distinct tablets that the archaeologist discovered, and each of the boys carried the couch individually.

Lakoff's discussion of the ambiguity of (9) vaguely suggests that he views distributivity as a determinant of the number of events denoted by such sentences: "It can mean either that the archaeologist discovered a group of nine tablets or that the number of tablets that he discovered altogether totaled nine, though they may not have been in a group."

I think that the question of distributivity is logically distinct from that of the number of events denoted by a sentence. Consider the following example:

(5) John gave a whole pumpkin pie to two girls.

The sentence is ambiguous: did John give a total of one pie to a group of two girls, or a whole pie to each of two girls? On the first reading, with the group-denoting indirect object, neither of the girls by herself has the property of having been given a whole pumpkin pie by John, while they each have this property on the second, distributive reading. On the group reading of the indirect object, it seems clear that only one event is denoted, but what about the second, distributed reading? Perhaps the giving was simultaneous, in the same location. Is this one event or two? One might argue that spatio-temporal continuity is neither sufficient nor necessary to individuate events. But it is just such fuzziness in our sense of what constitutes an event, or, perhaps more to the point, of what denotes an event, that makes it difficult to find any direct correlation between the truth conditions of such sentences and the number of events involved.

In sum, I think that the semantic theory of events is not yet sufficiently well developed to permit clear claims about the number of events denoted by a given sentence. A sentence such as (5) may just be vague or noncommittal about whether or not the tablets were discovered at one time. The distinct question of which entities have a particular property, as in (9), (10) and (12), appears to have a much clearer answer for a given reading of a particular sentence.

Schein (1986) has developed a theory which is aimed at accounting for a restricted class of examples involving distributivity in terms of quantification over events. In Section 4.2.1 I will discuss the relevance of this proposal for examples involving what Lakoff has called "plural quantification"; however, because the terms in which Schein couches his theory are generally incommensurable with those of the theories under
the boys carried the couch upstairs alone. (11) and (12), on the other hand, are unambiguous, with only a quantifier reading. Lakoff (1985) had proposed that the scope of quantifiers should be indicated in deep structure by generating them as higher predicates which are then lowered transformationally in the derivation of surface structure. In his discussion of (9)-(12), he proposes that only the quantifier reading be derived by ‘quantifier lowering’ or, alternatively, ‘quantifier raising’. He has no theory of the representation of the group-reading; however, from the examples and his discussion it seems that he intends in situ interpretation as a necessary, if not sufficient requirement. Also, although he is not explicit on this point either, presumably the rule of Quantifier Lowering is obligatory for quantifiers such as every and few, in order to account for the non-ambiguity of (11) and (12). Then either numerals and all (the) are ambiguous between quantifier and non-quantifier interpretations, or the rule of Quantifier Lowering only applies optionally to them. In either case, it seems that the two readings of (9) and (10) are to be distinguished by whether or not Quantifier Lowering has applied.

As further evidence for this proposal, Lakoff presents examples (13) and (14):

(13) Sam believed that that archaeologist discovered nine tablets.

(14) Sam believed that all the boys carried the table upstairs.

He claims that (13) and (14) have three readings each: both a group reading and a quantifier reading which have narrow scope with respect to the opaque verb believe, and a quantifier reading with wide scope. They seem to lack the wide scope group reading, where there is some group of tablets or of boys of which Sam holds some belief, although he might not describe them as such. These judgments are not easy. Note with respect to (13) that if there is a group of tablets of which Sam holds the belief that it was discovered, then it would be true of each tablet in the group that he holds this belief about it. Hence, the truth conditions for the wide scope group reading do not differ clearly from those for the wide scope quantifier reading. (15) is a better example, in that the truth conditions for wide scope group and quantifier readings differ considerably:

(15) Sam believed that the archaeologist gave nine tablets to a museum.

On the wide scope quantifier reading, each of nine tablets is such that Sam believed that the archaeologist gave it to a museum (possibly different museums),

discussion here, I will not attempt to review his proposal in detail.

3Unless the wide scope quantifiational reading is taken to indicate that there were nine distinct events of tablet discovery, as opposed to a single discovery of a group of nine tablets. See footnote 1 above for discussion of this possibility.

3.1. PREVIOUS APPROACHES

while on the wide scope group reading, the whole group is such that Sam believed that it was given to a (single) museum. In (16), I find the wide scope group reading easier to get than the wide scope quantifier reading (as is also the case with the narrow scope group reading vs. the narrow scope quantifier reading, in this particular example). In (14), I do find the wide scope group reading difficult, if not impossible, but this reading is much easier to get for the structurally similar (16):

(16) Sam believed that six boys formed a club.

Consider a situation in which the six boys in question are the Jones brothers: Sam may believe that the Jones brothers formed a club without knowing how many of them there are. If the speaker knows that there are six Jones brothers, then I think he can truly assert (16). Compare (15) also to the similar (17), where neither the wide scope nor the narrow scope quantifier reading seems available:

(17) Sam believed that all the boys formed a club.

I can’t get a reading where it’s true of each of the boys that Sam believed he formed a club. I don’t get a wide scope group reading of (17), either. Whatever the source of the variation from example to example, Lakoff’s argument from examples (13) and (14) does not seem conclusive.

Lakoff presents one further argument for the Quantifier Lowering analysis of distributivity. Since Quantifier Lowering is a movement rule, one would expect it to obey Ross’ (1967) constraints on movement. In particular, it should obey the Coordinate Structure constraint. Consider Lakoff’s examples (18) and (19):

3David Pesetsky (p.c.) suggests that the wide scope quantifier reading is possible with the similar (i):

(i) Sam has at one time or another held the belief that all the boys formed a club.

I still don’t get the intended reading, though I do with (ii):

(ii) Sam has at one time or another held the belief that each of the boys formed a club.

4See also Hütink (1973) critique of Lakoff’s use of examples such as (6) to argue for the ambiguity of nine tablets. As Hütink points out, “from the fact that an expression exhibits an ambiguity when embedded in a certain kind of [e.g. opaque] context, it does not follow that it is ambiguous when considered alone”.

4
(18) (a) John and nine boys are similar. (Unambiguous)
(b) John and all the girls are similar. (Unambiguous)
(c) *John and every linguist are similar.
(d) *Few philosophers and John are similar.

(19) (a) John is similar to nine boys. (Ambiguous)
(b) John is similar to all the boys. (Ambiguous)
(c) John is similar to every linguist. (Unambiguous)
(d) Few philosophers are similar to John. (Unambiguous)

The examples in (18) involve intransitive, symmetric similar. If we assume with Lakoff that Quantifier Lowering obeys the Coordinate Structure constraint, then we expect that the quantifier reading is not possible for constituents of a conjoined NP. When an optionally quantificational NP such as numeral CN or of all the CN (cf. the ambiguous (9) and (10)) is conjoined with another NP, as in (18a) or (18b), the result is unambiguously interpreted as a group, and there is one property which makes all members of the group similar. However, quantified expressions such as every CN undergo obligatory Quantifier Lowering, as shown in the unambiguous example (11) above with only the quantifier reading. Such NPs are not felicitous when conjoined with another, non-quantificational NP, as in the ungrammatical (18c). The same quantificational analysis of few would account for (18d). In (19), with transitive similar, the (a) and (b) examples are ambiguous: there may be a single, shared property by virtue of which John is similar to the group of boys, or there may be a different property shared by John with each of the boys, with the direct object receiving the quantifier reading. But (19c) and (19d) are unambiguous, with only the quantifier reading of every linguist and few philosophers.

Pesetsky (1982), in a discussion of Russian constituents of the form numeral CN, notes that within the Government and Binding framework, James Higginbotham has suggested that the distinction between group and distributive readings of numeral phrases “may be traced to the optionality of QR”, or Quantifier Raising. In a complex analysis, Pesetsky shows that one type of Russian numeral CN displays a range of characteristics typical of NPs which have undergone QR, including the ECP, or Empty Category Principle, of Kayne (1980). He calls this the

4I borrow the expression CN, mnemonic for “common noun phrase”, from Montague grammar.

In general I use CN instead of N” or N’ in order to avoid questions of the number of bars involved in this immediate daughter constituent of NP (or QP). A CN may include prehead modifiers, such as adjective phrases, and posthead complements, adjuncts and, according to some analyses, relative clauses.

4Pesetsky argues that these constituents in Russian are not NPs but Quantifier Phrases, or QPs.

3.1. PREVIOUS APPROACHES

‘no agreement’ numeral CN, because it fails to trigger agreement with a predicate. It also displays idiosyncratic case behavior, appearing in the genitive case regardless of context. This type of numeral CN is felicitous with group predicates, such as the Russian counterpart of the English disperse, so that it seems to have only a distributive interpretation. It may well be correct that these Russian constituents are quantificational, thus explaining their felicity with group predicates in general and the evidence that their behavior with respect to the ECP parallels that of other, clearly quantificational NPs. However, although this is compatible with the Lakoff/Higginbotham view of the nature of the group/distributive distinction, it does not in and of itself argue that the group reading of NPs (with or without numeral determiners) arises due to the absence of QR. I believe that the objections which I will raise to Lakoff’s proposal will also cause problems for a similar GB approach, at least for English.

Although I believe that Lakoff’s proposal is correct in recognizing the quantificational character of distributivity, there are two crucial types of example which show that it is inadequate. First, there are distributive readings of sentences with nonquantificational subjects. Consider the following:

(20) Mary, Susan, and Kathy have broken their leg at one time or another while learning to ski.

The only pragmatically plausible reading here is where each of the individuals in the group denoted by the subject has the property denoted by the predicate. Although many similar examples are not as acceptable to those (like me) who prefer the plural form of direct objects in such cases (the "dependent plural"), the existence of such examples shows that conjoined proper names, which are not regarded as quantificational in any theory I am aware of, can have a distributive interpretation. This interpretation has very similar truth conditions to those of closely related examples with clearly quantificational subjects, as in (21):

(21) Each girl has broken her leg while learning to ski.

The behavior of the two types of subjects in (20) and (21) differs in two ways which demonstrate the “referential” character of the conjoined proper names and the quantificational character of each girl. The first is the possibility of discourse anaphora involving the two examples. Consider:

(22) Mary, Susan, and Kathy have broken their leg at one time or another while learning to ski. They had to wear a cast for a long time.
Each girl has broken her leg while learning to ski.
# She should have stayed on the practice slopes longer.

While the anaphoric link between the pronominal subject of the second sentence in to the subject of the first sentence in (22) in felicitous, this is not the case in (23). This is exactly what we would expect if Mary, Susan, and Kathy is nonquantificational, and hence available as a discourse antecedent, while each girl is quantificational, and hence only able to serve as antecedent in c-command anaphora.

The other test for the distinction is their behavior in sloppy identity constructions such as (24) and (25):

(24) Mary, Susan and Kathy love their mother and Bob does too.

(25) Each girl loves her mother and Bob does too.

(24) is ambiguous between the sloppy reading, where Bob loves his own mother, and the strict reading where he loves the mother of Mary, Susan and Kathy. But (25) has only the sloppy reading. As discussed in Chapter 2, Reinhart (1983) argues that the sloppy reading requires bound anaphora, and I argued there that the nonsloppy reading arises due to discourse binding. Thus, the two NPs again appear to have a different binding potential. (Of course, in certain contexts both examples have a strict reading where Bob and the girls all love the mother of some other female person already under discussion. But that reading is not relevant here, where the examples are to be considered as if uttered out of the blue.)

Lakoff’s account, distinguishing between the distributive and group readings according to whether an NP has undergone a rule of Quantifier Lowering or Quantifier Raising, cannot capture the truth conditional parallels between the clearly quantificational cases such as (21) and cases like (20), while at the same time accounting for the anaphoric facts.

The second type of problem is more serious, since it shows a contradiction in Lakoff’s basic assumption about the relation between distributivity and quantifier scope. Example (26) shows that the function of Quantifier Lowering/Raising as an indication of the scope of an NP is incompatible with the proposal that it is restricted to distributively interpreted NPs:

(26) Five insurance associates gave a $25 donation to several charities.

(26) is multiply ambiguous. The reading which interests us is where a single group of five insurance associates gave a donation of $25 to each of several charities. On this reading, five insurance associates has a group reading, while several charities is distributive over a $25 donation, so that it has a quantifier reading. Five insurance associates has wider scope than several charities; i.e. the relative scopes of the NPs in (26) are as in (27):

(27) $\delta_{\text{group}} - \text{several}_{\text{distr}} - $25

If the distributivity of several charities and its scope over a $25 donation are to be represented in a logical form by Quantifier Lowering or Raising, then five insurance associates would have to undergo that rule as well, despite its group reading.

Note that one could not maintain that all group denoting NPs interpreted in situ automatically have wide scope. On this view, the wide scope of the subject of (26) would not entail that it had been moved. But this approach would also predict that the subject of (26), on its group interpretation, could not have narrow scope with respect to the quantificational (and hence raised) several charities. However, there is such a reading of (26), where each charity was given $25 by a possibly different group of five insurance associates. Therefore, the scope of the group denoting subject with respect to the quantificational indirect object is free, demonstrating that the assumption that distributive NPs undergo quantifier movement while group denoting NPs are interpreted in situ is inadequate.

Finally, even if the Quantifier Lowering/Raising approach were adequate descriptively, we would still need an account of how the different readings were derived from the raised and non-raised structures. Would the interpretive differences arise entirely from the structural differences, or would lexical differences between NPs make a contribution as well? The different potential for distributivity of different types of NPs suggests the latter; however, if lexical factors are involved in interpretation, it is unclear why the movement rule would be a necessary component in the explanation of distributivity (though, of course, it might play another role — the indication of the scope of NPs).

---

4 In a similar move, Aoun, Hornstein, & Sportiche (1981) argue that any is interpreted in situ but has a special logical translation which always, in effect, gives it wide scope. This is based on the claim that any always has a wide scope universal interpretation, though it does not display the properties associated with Q/R NPs. However, they do not note that the examples they consider all permit a different logical form — where any is interpreted as an existential which has narrow scope with respect to a downward entailing operator, such as negation (cf. Ladasaw (1979)). Their argument is seriously undermined by their failure to address this possibility.
3.1.2 The distributive-group distinction as a function of predicate type

Some lexical items have been held to permit only a group or a distributive reading of sentences in which they occur. The examples in (28) – (31) illustrate this:

(28) (a) * John disperses.
(b) The committee disperses.

(29) (a) * John walks together.
(b) The men walk together.

(30) (a) * Mary is among John.
(b) Mary is among the unicorns.

(31) (a) John, Paul, George, and Ringo are pop stars.
(b) Paul is a pop star.

In (28)–(30), from Bennett (1974), are a number of lexical items which require a group denoting subject or complement. Disperse seems to require such a subject; hence (28a) is unacceptable, since John does not denote a group. But (28b) is fine; it seems that the committee denotes a group. The fact that it is syntactically singular shows that the group/nongroup distinction is not directly tied to syntactic number. As (29a,b) shows, adverbial together seems to turn the predicate which it modifies into one which takes only a group denoting subject. And (30a,b) show that among seems to require a group denoting complement.

In (31), from Link (1983), there doesn’t seem to be a group reading as distinguished from a distributive reading — the predicate be a pop star seems to be true only of individuals. So, if (a) is true, then it entails (b) automatically. Thus, be a pop star is said to be a distributive predicate. In this class are often included verbs such as walk, eat, talk, and others which seem to be related to personal identity or individual will.

It is these types of examples which inspire the second general approach to the distributive/group distinction, which emphasizes the contribution of particular lexical predicates. Following Montague (1973), both CNs and VPs are predicates, so on this approach the lexical entries of such elements might contain information about which class (or classes) they belong to. This type of approach was originally investigated by Bennett (1974), who distinguished two classes of predicates, those which contain individuals in their extension, and those which contain groups. More recently, the distinction between distributive predicates such as be a pop star and other, non-distributive predicates has provided the foundation of Link’s (1983) theory of distributivity. In this section, I will review in detail Bennett’s theory, and compare it briefly with Link’s. I will postpone a more detailed consideration of Link’s theory until Section 3.2, where I discuss his theory of the semantics of groups.

Bennett (1974) proposes a treatment of plurality, including the distributive-group distinction, by an extension of the fragment of English treated by Montague (1973). The central feature of his analysis is the use of two different classificatory distinctions. The first is a syntactic number distinction: CNs may be either singular or plural. Pluralization is accomplished by a rule which does not affect the semantic type of the CN involved. Besides affecting its morphology, the syntactic number of a CN restricts the determiners with which it can combine and the number of the resulting NP, and the syntactic number of a subject NP triggers number agreement in the verb of its predicate. NPs themselves may be plural as well without having plural CN heads. For example, John and Mary is plural. Agreement in number of pronouns with their antecedents is also syntactic.

The other distinction is manifested in both the syntax and the semantics: there are two kinds of CNs, individual level and group level. Bennett distinguishes these in the categorial syntax by the use of ‘CN’ as the category for the individual level CNs and the same symbol with a bar over it for the group level. For typographical convenience, I will use ‘CN$_g$’ for the group level CNs. The categorial difference between CN and CN$_g$ is reflected semantically in different types: the denotation of individual level CNs is a function from individuals in the universe of discourse to truth functions, type $(e, t)$ but the denotation of group level CNs is a function from sets of individuals (the groups) to truth functions, type $(\{e, t\}, t)$. There is also a new basic category for group level verbs such as gather, /em riot, and disperse: IV$_g$, interpreted by the same type as CN$_g$. This distinction in basic categories and types effects a split in derived categories as well: for example, at every level where a CN$_g$ may enter into a larger constituent, we must lift the categories of any element we combine it with and of the resulting constituent, and we must also lift the semantic type of their interpretations. The inventory of basic categories must be expanded to include a group level NP, T$_g$ (= t/IV$_g$), and group level pronouns it$_g$ and they$_g$, to permit quantifying in group level NPs; Bennett also needs various verbal categories such as TV$_g$ (= IV$_g$/T$_g$), which takes a group level object NP to form a group level predicate, IV/T$_g$, which takes a group level object NP to form an individual level predicate, and IV$_g$/T, which takes an individual level object NP to form a group level predicate. There is also a category for adverbial together, IV$_g$/IV, which takes an individual level predicate to form a group level predicate; and a category for prepositions such as among which take group level complements, A/T$_g$. Thus, because of the individual-group type distinction in Bennett’s system, NPs may only combine with VPs on

Footnote 3: Bennett departed from Montague in this regard; CN in PTQ (Montague 1973) was intensional, of type $(e, t)$. Bennett’s treatment of CN has been widely adopted among Montague grammarians; see Dowty, Wall & Peters (1981) for discussion, Janssen (1984) for arguments that Montague’s CN type is preferable in some cases.
CHAPTER 3. DISTRIBUTIVITY

the same level, whether group or individual, and the same constraint is found in other rules of function-argument application as well.

The two kinds of distinctions cross-cut each other: there are both singular and plural individual level CNs, for example man vs. men, as well as singular and plural group level CNs, for example committee or mob vs. committees, mobs. There are two main classes of determiners, one class combining with singular CNs or CNs, the other with plural CNs and CNGs. Some of the most common determiners and their classification by Bennett are shown in (32):

(32) 

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
</tr>
<tr>
<td>every</td>
<td></td>
</tr>
<tr>
<td>one</td>
<td></td>
</tr>
<tr>
<td>two, three...</td>
<td></td>
</tr>
<tr>
<td>the</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td></td>
</tr>
<tr>
<td>few</td>
<td></td>
</tr>
<tr>
<td>some</td>
<td></td>
</tr>
<tr>
<td>many</td>
<td></td>
</tr>
<tr>
<td>most</td>
<td></td>
</tr>
<tr>
<td>all</td>
<td></td>
</tr>
</tbody>
</table>

There are a variety of rules (S12–S17, S21, S22) which form NPs from these determiners in combination with CNs and CNGs. Singular CNs on the individual level, may only feed individual level NPs, of category T. Plural CNs may combine to form either an individual level NP, T, or a group level Tg. Singular CN, may either form part of singular Tg or plural T; the latter possibility permits the derivation of sentences with semantic, but not syntactic, number agreement between subject and predicate, such as The committee walk. Plural CNg always combines with a determiner to form a Tg; however S2 (optionally) changes a plural Tg to a plural individual level T; so that, in effect, a plural CN of either type may enter into a NP of either type.

Because there is no direct mapping from singular T to Tg and only singular CNg and plural CN or CNG can feed rules forming Tg, certain readings are ruled out. These involve singular, individual level NPs of category T in combination with group level predicates or prepositions, as in (28)–(30), repeated from above:

(28)  
(a) * John disperses.
(b) The committee disperses.

(29)  
(a) * John walks together.
(b) The men walk together.

(30)  
(a) * Mary is among John.
(b) Mary is among the unicorns.

John, of type T, can never be the subject of a group level predicate such as disperse, of type IVg, or of any predicate modified by the group predicate forming adverb together, nor can it be the complement of a group level preposition such as among. A singular CN such as man can never form a group level NP either. Since committee is of type CNg, it forms an NP of type Tg, which can combine with such predicates. Similarly, the plural CN men can combine with the to form a group level NP of type Tg.

Because Bennett’s theory locates distributivity in properties of predicates, rather than in a scope rule such as Quantifying In, he predicts some readings which Lakoff cannot derive in his theory, where quantifier movement determiners both distributivity and scope. This is illustrated in the readings which each theory predicts for Bennett’s example (33):

(33) Ten men applauded two women.

Bennett permits a large number of different analyses of (33), depending on the group vs. individual level of the NPs, whether or not they are quantified in, and the relative order of their scopes. However, several of these analyses yield logically equivalent interpretations, and the total number of distinct readings he derives is seven. These are shown informally in (34), where only the relative scope of the two NPs and their types is given. Subscript ‘d’ indicates an individual level, or distributive reading. Subscript ‘g’ indicates a group reading:

(29)  
(a) * John disperses.
(b) The committee disperses.
Bennett's readings of (33):  

(i) \[10 \text{ men}_d\] \[2 \text{ women}_d\]  
(ii) \[2 \text{ women}_d\] \[10 \text{ men}_d\]  
(iii) \[10 \text{ men}_d\] \[2 \text{ women}_d\]  
(iv) \[2 \text{ women}_d\] \[10 \text{ men}_d\]  
(v) \[10 \text{ men}_d\] \[2 \text{ women}_d\]  
(vi) \[2 \text{ women}_d\] \[10 \text{ men}_d\]  
(vii) \[10 \text{ men}_g\] \[2 \text{ women}_g\]  

The readings in (ii), (iv) and (vi), where \(2\text{ women}\) has wide scope over \(10\text{ men}\), must involve Quantifying In because of the inverse scope; however, all of the other readings can correspond to analysis trees which do not utilize Quantifying In. 

Lakoff only predicts five readings of (33), as shown in (35). The readings are numbered as they are in (34), in order to make clear how the two theories compare. Here, they are followed by the logical form which derives the reading for Lakoff:

Lakoff’s readings of (33):  

(i) \[10 \text{ men}_d\] \[2 \text{ women}_d\]  
(ii) \[2 \text{ women}_d\] \[10 \text{ men}_d\]  
(iii) \[10 \text{ men}_d\] \[2 \text{ women}_d\]  
(iv) \[2 \text{ women}_d\] \[10 \text{ men}_d\]  
(v) \[10 \text{ men}_d\] \[2 \text{ women}_d\]  
(vi) \[10 \text{ men}_g\] \[2 \text{ women}_g\]  

The two readings which are not available in Lakoff’s theory are those in which an NP interpreted as a group has wide scope over an NP interpreted distributively. These readings don’t make a great deal of difference for the truth conditions of (33), but recall that in Section 3.1.1 we found that the lack of such a reading was a problem in representing the reading of (26) indicated in (27):

Five insurance associates gave a $25 donation to several charities.

Bennett’s proposal would permit us to derive the reading in (23), and is thus an improvement over Lakoff’s. However, there are several problems with Bennett’s approach. One kind of problem arises with coordinate VP structures. Consider (36), after an example from Karina Wilkinson (p.c.):  

John and Mary won a lottery drawing and then developed insomnia worrying about the money.

The most prominent reading of (36) is one in which John and Mary jointly won a lottery drawing, then each developed insomnia. Bennett’s approach to what are often considered the inherently distributive predicates, such as eat and walk, treating them as predicates on the individual level only, suggests that he intended to derive the distributive entailments of such verbs from their type, so we might assume that he would have regarded develop insomnia as a predicate on the individual level only. We thus appear to have conjointed VPs of different types — one on the group level, the other on the individual level. Under the usual assumption that only constituents of the same type may be conjoined, this is a problem for Bennett’s approach.

But perhaps the most obvious problem with Bennett’s approach is the proliferation of types which it entails, as Bennett himself points out.\(^6\) Some predicates must be treated as several ways ambiguous. For example, applaud in this fragment is a member of four different categories, with four corresponding translations, \(IV/T\) (John applauded Mary), \(IV/T_g\) (The committee applauded the mob), \(IV/T_g\) (John applauded the mob), and \(IV/T_g\) (The committee applauded Mary).

A related problem involves group level CNs such as committee. Since plurality in this system is distinct from the group vs individual level distinction, there is no way to form a group of groups reading for plural group nouns such as committee, a type of reading necessary for examples such as (37):

The committees met last week.

where the committees may have met all together. Bennett himself noted the existence of CNs such as federation denoting groups of groups, and observed that an adequate treatment of such CNs would involve further type lifting across the board, resulting in an even more embarrassing array of types.

\(^6\) For extensive discussion of the problem of proliferation of types in Montague Grammar, see Parsons (1979).
Link's (1983) general approach to distributivity is that it is a lexical property of one-place predicates, but unlike Bennett, he doesn't reflect this in a type distinction; thus he avoids the problems with VP conjunction and the proliferation of types. As we will consider in more detail in Section 3.2, a group in Link's system is just another kind of individual, so that predicates such as *disperse* are of the same type as Bennett's individual predicates. Link does not consider such group predicates, focusing instead on predicates such as *be a pop star* in (320). He defines a special class of Distributivity Predicates (which I will abbreviate as "DistPs"); these DistPs are those which do not contain groups in their extension, but only single ("atomic") individuals (see his definition D19, p.314). He then defines a meaning postulate for this class of predicates which guarantees that whenever one of them takes a plural subject, it holds of all the individuals which make up the group denoted by the subject. From this, Link is able to guarantee the valid inference from (31a) to (b), repeated below, on the assumption that the predicate *be a pop star* is a DistP:

(31) (a) John, Paul, George, and Ringo are pop stars.
(b) Paul is a pop star.

Though Bennett's and Link's theories differ in many respects, they are alike in locating the distributive-group distinction primarily in lexical characteristics of predicates. But building this view of the distinction into the formal structure of a semantic theory is, on the one hand, unnecessary because it is redundant, and, on the other hand, it fails to capture some important generalizations about the nature of distributivity. First, the fact that a particular lexical item is a group predicate or a distributive predicate doesn't really need to be specified independently: it follows from the sense of the predicate itself. What does it mean to gather or to disperse? By virtue of the meaning of such a predicate, its subject must denote a group of individuals (or a mass of something), performing in a way peculiar to a group (or mass). Viewed in this way, these verbs are no more special than a verb such as *grasp*, which, on one of its senses, can only be true of an individual with a certain type of movable thumb. What is it to be a pop star or to walk or to die? The actions or states denoted by these verbs can generally only be performed or endured by an individual with a single will and consciousness. It is for this reason that we think of them as distributive. Although it may well be that only atomic individuals are in the extension of such distributive verbs in their strict sense, this follows from our knowledge of what is required for them to be true of an individual, that is from our knowledge of selectional restrictions on their use (however selectional restrictions may be encoded in the grammar).

Note also that many of the predicates which might be considered group predicates or distributive predicates are not composed of single lexical items. For example, it seems that only a group can make a good team. Under standard assumptions about the lexicon, there would be no entry for such a predicate, since its meaning could be compositionally determined on the basis of the meaning of its parts. But, given what it means to make a team, we would naturally assume that its extension contained only groups. The same is true of win a relay race but not of win a 100 meter dash, the latter presumably a distributive predicate. These classifications are a question of world knowledge about the denotations of the terms involved.

The other kind of problem with founding a theory of the group/distributive distinction on the properties of predicates alone is that it does not take into account the important and systematic contributions of determiners and adverbial elements like floated quantifiers to distributivity. By restricting our attention to predicates alone, it seems that we are missing an important generalization about the unity of the contribution made by these various elements. This problem leads us naturally to consider the third major approach to the distributive-group distinction, wherein it is viewed as a property of determiners.

### 3.1.3 Determiners and the distributive-collective-cumulative distinction

In this section, we will consider an approach to the distributive/group distinction which focuses on the contribution of determiners. First, we will review an influential proposal in this vein by Remko Scha (1981). Then we will consider Scha's claim that there is a third kind of reading, the cumulative reading; the evaluation of this proposal involves exploring more generally what it means for a relation to hold between two groups, and leads us into a brief consideration of the semantics of reciprocal sentences.

#### 3.1.3.1 Determiners as the locus of the distinction

Remko Scha's "Distributive, Collective and Cumulative Quantification" (1981) locates the kind of distinction we are concerned with in the semantics of determiners, instead of in the distinction between distributive and group (or "collective") verbs and CNs which was the basis of Bennett's approach. In addition, he contends that a two-way distinction between distributive and group, or "collective" readings is not sufficient: there is a third kind of reading, which he calls the "cumulative", and his proposal is aimed at accounting for these as well.

For Scha, all CNs, whether singular or plural, denote sets of singleton sets of individuals, rather than sets of individuals. The advantage of this is that it permits distributive and collective predicates to be of the same type: When a

---

7I have been influenced in this discussion by Dowty (1986). See Section 3.3.1 below.

8Stein (1971) also interprets individuals as singleton sets in a fragment of Thai.
distributive determiner is combined with a CN by function-argument application, the result is a function from one place predicates to truth values, as in Montague’s system. For example, the determiner *all*, on its distributive interpretation, when combined with the plural CN *boys*, yields the interpretation indicated by the logical form in (37):\(^9\)

\[
(37) \quad \lambda P [\forall \{x\} (\text{boys}(\{x\}) \rightarrow P(\{x\}))]
\]

(\{x\} is the set containing x.) This resembles the standard interpretation of such an NP, as in Montague’s treatment; however, here *boy* denotes a set of singleton sets, and quantification binds a special variable over singleton sets. *All* also has a collective interpretation.\(^{10}\) This sense of *all* forms the union of the singleton sets in the extension of the CN, as in the logical form for *all boys* in (38):

\[
(38) \quad \lambda P [P(\cup(\text{boys}))]
\]

A one-place predicate \(P\) in such a system need only be a function from sets to truth values. It will then combine with either (37) or (38), since in each case its argument is a set.

This predicts that any subject-predicate combination should be grammatical, regardless of whether the verb would be individual-level or group-level on Bennett’s approach. Scha, indeed, says that his grammar is “more tolerant”, accepting sentences such (39) as grammatical, although the quantifier *each* is unambiguously distributive:

\[
(39) \quad \text{Each boy gathers.}
\]

Further, since he regards *all* as ambiguously either a group or a distributive quantifier, (40) is ambiguous, yielding both the expected collective reading and an implausible distributive reading:

\[
(40) \quad \text{All boys gather.}
\]

Scha regards (39) and the distributive reading of (40) as “semantic anomalies”, of the same kind as the famous *colorless green ideas sleep furiously*, that is, presumably, as cases which violate selectional restrictions.

The table in (41) presents Scha’s classification of the determiners he treats:

\[
(41) \quad \text{Scha’s classification of determiners}
\]

<table>
<thead>
<tr>
<th>Distributive</th>
<th>Collective</th>
</tr>
</thead>
<tbody>
<tr>
<td>each</td>
<td></td>
</tr>
<tr>
<td>every</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td></td>
</tr>
<tr>
<td>both</td>
<td></td>
</tr>
<tr>
<td>(\phi)</td>
<td>(\phi)</td>
</tr>
<tr>
<td>all</td>
<td>all</td>
</tr>
<tr>
<td>some\text{sing/pl}</td>
<td>some\text{pl}</td>
</tr>
<tr>
<td>no\text{sing/pl}</td>
<td>no\text{pl}</td>
</tr>
<tr>
<td>2, 3, 4 \ldots</td>
<td>2, 3, 4 \ldots</td>
</tr>
<tr>
<td>the\text{sing}</td>
<td>the\text{pl}</td>
</tr>
</tbody>
</table>

A number of determiners in this taxonomy are considered unambiguously distributive — all the singular determiners and *both*. The null determiner, *all*, plural *some* and the numerals are ambiguous; only the plural definite article is unambiguously a group determiner.

Scha notes that Bennett (1974) and Hauser (1974) both treat plural *the* as ambiguous, with a distributive reading as well. He presents examples such as the following as evidence that the only correct reading of plural *the* is the group reading:
(42) The squares contain the circles

\[ \text{circle}(y) \rightarrow \exists x (\text{square}(x) \& \text{contain}(x, y)) \]

None of Bennett’s types for contain, IV/T, IV/T_y, IV_y/T_y, or IV_y/T, yields (43). Anytime the subject is treated as distributive, this has the effect of universal quantification; but in (43) it is as if existentially quantified. One might try treating the subject NP as a group but the object as distributive, with the V type IV_y/T. This gets the correct, universal quantification over the circles and gives a sense ‘each of the circles is contained in the group of squares’, very close to (43). But other examples are even more complex, as in (44), with truth conditions something like (45), so that it may be true in the situation shown in the accompanying figure:

(44) The sides of rectangle 1 cross the sides of rectangle 2

\[ \exists x \exists y [\text{side-of } R1](x) \& \text{side-of } R2(y) \& \text{cross}(x, y)] \]

Scha then proposes to treat (42) and (44) by treating all the definite NPs as group-denoting, then applying meaning postulates on the verbs contain and cross, as in (46) and (47):\(^{11}\)

\[ \begin{align*}
(46) & \quad [\text{contain}(u, v)] \rightarrow [\forall y ([y]ev \rightarrow \exists z ([z]eu \\
& \quad \& \text{contain}([z], [y])))] \\
(47) & \quad [\text{cross}(u, v)] \rightarrow [\exists z \exists y ([z]eu \& [y]ev \& \text{cross}([z], [y]))]
\end{align*} \]

I agree with this approach to the truth conditions of sentences with plural definite articles. However, I don’t feel that these particular examples constitute by themselves a sound argument against treating the definite article as ambiguously distributive or group denoting. Note that Bennett could also derive the meanings in (43) and (45) by treating all the definite NPs in these particular examples as group denoting and the verbs of type IV_y/T, then adding the meaning postulates (46) and (49). Of course, he would also predict other readings with the distributive sense, but even if this seems infelicitous in these examples, is it in general impossible for plural definites to have a distributive interpretation, or is this interpretation only marked?

Under the assumption that definite plurals are unambiguously collective, Scha follows Bartsch (1973) in using meaning postulates on verbs such as walk and eat.

\(^{11}\)Again, Scha’s logical forms are somewhat different; however, I believe (46) preserves the main intended truth conditional effects of his (5) in Section 6. His logical form also includes a requirement that the set denoted by the subject be non-empty, but I have omitted that as inessential here. \(\varepsilon\) stands for ‘is an element of’.
to derive distributive readings when these combine with definite plural subjects. For example, in order to derive the appropriate truth conditions for (48), Schä proposes a meaning postulate such as (49):

\[(48) \quad \text{The boys walk.}\]

\[(49) \quad ([\text{walk}(x)] \rightarrow [\forall y (\{y\} \cap x \rightarrow \text{walk}(\{y\})])]\]

One final aspect of Schä's analysis which is of interest here is his treatment of what he calls 'cumulative quantification', describing the most plausible reading of examples such as (50):\(^{12}\)

\[(50) \quad 600 \text{ Dutch firms have 5000 American computers.}\]

He claims that this example has the reading in (51), a reading which could not be obtained from any combination of group and distributive readings of the two NPs:

\[(51) \quad [\text{Cardinality } (\{x\}) [\text{DuF}(\{x\}) \& \exists y [\text{AmC}([y]) \& \text{have}(\{x\}, \{y\})]) = 600] \& [\text{Cardinality } (\{y\}) [\text{AmC}([y]) \& \exists x [\text{DuF}(\{x\}) \& \text{have}(\{x\}, \{y\})]) = 5000]\]

It is in order to obtain this reading for (50) and similar examples that Schä invents a new mechanism for combining verbs with their arguments. Instead of taking \( n \) arguments one at a time, his \( n \)-place predicates take a single argument: an \( n \)-tuple consisting of \( n \) NPs. In addition, for examples with cumulative quantification, pairs of NPs of the form \( \text{numeral CN} \) may be generated by a special rule as a single "compound noun", with a meaning which is the cartesian product of the meanings of the two NPs individually:

\(^{12}\)The following example, from Partee (1975), illustrates the cumulative reading more vividly:

\[(i) \quad \text{(A total of) three women gave birth to (a total of) five children.}\]

She calls the intended reading of (i) the 'total-total' reading. The example improves on Schä's because although a conglomerate or group of companies could jointly own one or more computers, no two women can give birth to the same child.

The type of reading Schä has in mind here also seems to be what Lauri Carlson (1988) and Schein (1986) call the "sum of the plurals" reading.

---

3.1. PREVIOUS APPROACHES

\[(52) \quad \text{DuFs} \times \text{AmCs}\]

There are also compound numerals, and one of these may combine with a compound noun of the same arity \( n \) to form an NP which is a function from \( n \)-place relations to truth values. When such a compound NP combines with an \( n \)-place verb, a sentence results. In the case of (50), the truth conditions which result are equivalent to those in (51).

The Schä/Bartsch approach to the treatment of distributive verbs via meaning postulates retains some of the good features of Bennett's theory, while improving on it in other respects. The main advantage of the approach Schä advocates over earlier approaches is that he is able to predict readings of examples such as (26), \textit{five insurance associates gave \$25 to several charities}, where a group NP has wide scope over a distributive NP, without the type proliferation of Bennett's fragment. He achieves this by regarding distributivity as a lexical property of determiners, rather than as a function of scope or of the type of CNs and VPs.

But this same aspect of Schä's theory has some disadvantages, as well. He restricts the contribution of predicates to the distributive/collective/cumulative distinction to meaning postulates and selectional restrictions. One problem arises with the class of mixed verbs whose subjects sometimes appear to be distributive, at other times collective. \textit{Lifted}, as in examples (1)–(4) above, is one example of this very common class, which also includes \textit{bring, carry, give, take, own}, and many others. The group reading of definite plural NPs is strongly preferred with mixed predicates. For example, (53), with the mixed predicate \textit{bring (something)} is generally considered quite ambiguous. As we might expect, (54), with a distributive-only subject, has only a distributive reading, but (55), with a definite plural subject, strongly suggests a group reading:

\[(53) \quad \text{Four women brought a salad to the potluck.}\]

\[(54) \quad \text{Every woman brought a salad to the potluck.}\]

\[(55) \quad \text{The women brought a salad to the potluck.}\]

Yet in some contexts (55) can have a distributive reading. Consider the following discourse:

\[(56) \quad \text{The women are going to bring a salad to the potluck.}\]

\[(57) \quad \text{The women are bringing a salad to the potluck.}\]
Every woman brought a dish to the potluck.
The hostess asked those from Acton to bring a casserole.
The women from Boxborough brought a salad, and those from Littleton a dessert.

The subject of the underlined sentence is interpreted distributively.

The problem here is that the markedness of the distributive reading of the definite plural subject in (55), as opposed to the distributive potential of the subjects of (55) and (54), seems to support Schä's classification of plural the as unambiguously collective, as does the accessibility of this subject of the underlined sentence in (56) to serve as a discourse antecedent, as in (57):

(55) The women from Boxborough brought a salad, and they came early to help set up.

If they is not the source of the distributivity here, then it must lie elsewhere. It is only other source of distributivity were meaning postulates on predicates, like examples such as (55) would force one to claim that meaning postulates could be optional. But this seems inadvisable.

Alternatively, one might claim that the verbs under consideration are ambiguous, with both a distributive and a collective sense which are otherwise identical in their entailments. This is undesirable, since it would require a proliferation of ambiguity which would parallel Bennett's proliferation of types for verbs like appear. And in any case, it would still leave the markedness of the distributive reading of (55) unexplained. Although I think Schä is correct in claiming that the distributive/group distinction rests in part on the nature of the determiner involved, there is a broader characterization of distributivity which will permit us to avoid this type of verbal ambiguity.

The other aspect of Schä's theory which bears close examination is his claim that there is a third kind of reading, the cumulative, and the way in which he proposes to account for such readings. We will consider the cumulative reading in more detail in Section 3.1.3.2. Here, however, I want to point out problems with Schä's formal account of how these readings are derived. Recall that Schä treats n-place verbs as functions taking a single argument, an n-tuple of NPs, and that the principal motivation for this unusual feature of his grammar appeared to be the treatment of cumulative readings of sentences such as (50). This approach is not well motivated syntactically, and in a more complex fragment I believe it would encounter serious problems with island constraints, conditions on bound anaphora, control, and other complex syntactic issues. Here, I will only illustrate my contention with another kind of problem: VP ellipsis. In Schä's fragment, there is no VP constituent syntactically or semantically. Hence, some of the standard approaches to the problem of VP ellipsis, such as Sag (1976) and Williams (1977), which state constraints on ellipsis in terms of VP constituents could not be translated into this framework. A phonological copying rule would not suffice, since issues of scope and anaphora (including sloppy identity) are involved. There are a number of other syntactic constructions which involve constituents such as VP or the notion of subject (whether basic or configurationally defined), and these present problems for an approach which simultaneously inserts all the arguments of the verb.

3.1.3.2 Cumulative readings

Here we will consider Schä's claim that there is a third type of reading, the cumulative. I will first review earlier work by Langendoen on reciprocals and their relation to other sentences with multiple group denoting plural NPs. Langendoen's insights will then prove useful in developing an analysis of Schä's collective and cumulative readings. I will adopt a suggestion of Barbara Partee's, and treat sentences with cumulative readings as a subclass of the class of sentences with two or more collectively interpreted NPs. I will argue that such sentences merely denote relations between as many groups, and that, as with reciprocals, any further details about the nature of the involvement of individual members of the groups in that relationship, which might be suggested by the particular lexical items involved, ought not to play a part in the truth conditional interpretation of the sentence. Finally, I will briefly review work by Gil (1982) which tends to support this view.

In "The Logic of Reciprocity" (1978), Langendoen attempted to discover an analysis of reciprocals such as each other and one another in which they make a unified, compositional contribution to the logical form of sentences in which they occur. In addition, he related the interpretation of sentences with reciprocals to that of a more general type of sentence including relations between plural NPs. He calls sentences of the form subject-verb-reciprocal object "Elementary Reciprocal Sentences" (ERSs), and those of the form plural N-verb-plural NP "Elementary Plural Relational Sentences" (EPRs). While he does not specify exactly the class of NPs he permits in EPRs, all of his examples use plural definite descriptions. Thus, I take it that he is interested in nonquantification, group denoting NPs.

Langendoen considers in detail six possible schemata for the truth conditions of ERSs: the strongest, Strong Reciprocity (SR), is given in (58), while the weakest, Weak Reciprocity (WR), is given in (59); in each, A stands for the set denoted by the subject, R for the relation denoted by the verb:

\[
(58) \text{Strong Reciprocity:}
\]

\[
(Vx, yx)(x \neq y \rightarrow xRy)
\]
(59) Weak Reciprocity:

\[(\forall x \in A)(\exists y, z \in A)(x \neq y \& x \neq z \& xRy \& zRx)\]

In SR, a given member of \(A\) must stand in the relation specified by the verb with each other member of \(A\); while in WR, a given member of \(A\) need only play each of the two roles specified by the arguments of the relation with some other member of \(A\). Langendoen argues in detail that any schema stronger than WR is too strong for the general case. The crucial kind of example, illustrated by (60), involves an asymmetric, disconnected relation on an unfounded set, since only WR, and none of the stronger schemas which Langendoen considers, can assign the value ‘true’ to such an example:

(60) The integers succeed one another.

Langendoen then turns to consider EPRSs such as (61):

(61) The women released the prisoners.

He considers the schematic representation of the truth conditions for such sentences given in (62), where \(A\) represents the set denoted by the subject NP, \(B\) the set denoted by the object, \(R\) the relation denoted by the verb:

(62) \((\forall x \in A)(\exists y \in B)(xRy) \& (\forall w \in B)(\exists z \in A)(zRw)\)

He notes that we can deduce WR, (59), from (62) if we i) set \(A = B\) in (62), and ii) add the condition that \(R\) not be reflexive \((x \neq y, x \neq z\) in (59)). “That is, WR follows as the truth conditional-schema for ERSs from the truth conditional-schema for EPRSs by substitution of the interpretation of the reciprocal element for the object phrase.” (p.187) He takes this as additional evidence for the claim that reciprocals make a uniform, compositional contribution to the sentences in which they occur, and for the correctness of WR, and therefore as evidence against the contention of Fiengo & Lasnik (1973) that the logical form of ERSs depends on properties of the relation \(R\) denoted by the verb.

In fact, as Langendoen points out, even the truth conditions for (61) which would be derived from the schema (62) may be too strong. “There remains ... the possibility that two or more of the women may have acted in the release of one or more of the prisoners in such a way that none of those women can be said to have individually released any of the prisoners.” (p.186) Langendoen claims that the same kind of problem arises in deriving a logical form for the ERS The

3.1 Previous Approaches

women released one another from WR, (59). He then proposes to weaken WR and the corresponding schema (62) for EPRSs to (63), Weak Reciprocity for Subsets (WRS), and (64), respectively, where \(\subseteq\) stands for the ‘subset of’ relation:

(63) Weak Reciprocity for Subsets (WRS):

\[(\forall x \in A)(\exists x_1, x_2, Y = \phi, Z = \phi \subseteq A)(x_1R_1x \& x_2R_2x \& \neg(x_1Y) \& \neg(x_2Y) \& R(x_1, Y) \& R(Z, X_2))\]

(64) \((\forall x \in A)(\exists x \subseteq A, Y \neq \phi \subseteq B)(xRx \& R(X, Y)) \& (\forall w \in B)(\exists w \subseteq B, Z \neq \phi \subseteq A)(wRw \& R(Z, W))\)

(64) specifies that each individual in the set \(A\) denoted by the subject must be a member of a subset of \(A\) which fulfills the role denoted by the first argument of the verb, and each individual in the set \(B\) denoted by the object must be a member of a subset of \(B\) which fulfills the role denoted by the second argument. (63) again follows from (64) on setting \(A = B\) and requiring non-reflexivity of the relation.

But even these schemata are too strong for some ERSs and EPRSs. The range of types of relations between individual members of the groups involved has been discussed at some length in the literature on reciprocals. Consider the following:

(65) The plates were stacked on top of one another.

(66) The leaves touched each other.

(67) The men knew each other.

(68) The men touched each other.

(69) The men killed each other.

(70) John’s grandparents hate one another.

Langendoen’s (65) is one of a group of exceptions to his WR/WRS. The problem with such examples, in contrast to (60), is the fact that the sets of individuals involved are well-founded with respect to the asymmetric relation—there is a first plate. He shows that such exceptions are systematic, in that they all involve restrictions on the spatial or temporal relations that order the elements of the set.
denoted by the subject, so that the order must go from top to bottom, outside to inside, front to back, left to right or right to left, or from earlier to later. Thus, while (68) is acceptable, (71) is not:

(71) the plates are stacked underneath one another.

Lauri Carlson’s (1980) (66) can be true in a situation where the leaves are all on a tree and most of them touch one or more other leaves. Thus, WR/WRS is too strong for it. But Fiengo & Lasnik’s (1973) (67) seems to require that all the men knew all the others, i.e. Strong Reciprocity, (58). They show that weak reciprocity alone would permit (67) to be consistent with both (72) (where a cyclic or infinite asymmetric relation held between the men) and (73):

(72) No two men knew each other.

(73) Many of the men did not know each other.

This seems undesirable, and thus suggests that WR is too weak. However, Colin Gordon’s (p.c.) very similar (68) is weaker than (67), in that not all the men have to touch all the others for the sentence to be true, and stronger than (66). And the counterparts of (72) and (73) do seem to be consistent with (68). Finally, in (69), also due to Gordon, it seems that only some of the men (but not just one or two) must have killed some of the others (but not just one or two) for the sentence to be true, and here again the sentence is consistent with counterparts of (72)–(73). These examples show that neither the semantic characteristics of the predicate alone, nor of the arguments alone, suffice to determine the strength of the reciprocal relation denoted. Chomsky’s (1975) (70) is a case where several different interpretations are possible: perhaps each couple is involved in mutual hatred, or perhaps both members of each of the two couples hate both of the members of the other couple, or perhaps there is just a lot of hatred between the four people. These are the kinds of examples which have been used to claim, for example by Fiengo and Lasnik (1973), that a uniform, compositional treatment of sentences containing reciprocals is not possible.

There is another approach to the problem. Emmon Bach (UMass. colloquium, 1980) has suggested that the reciprocal can contribute uniformly to a compositional derivation if we include in its translation a context sensitive quantifier ENOUGH. Like few, many, and most, as well as adjectives like big, etc., what is enough depends on lexical properties of the relation involved, as well as characteristics of the individuals involved and elements of the context. Bach proposes that reciprocals be treated along the same lines as reflexives are treated in Bach & Partee (1980), that is, as causing the introduction of an argument-reducing function over predicates. For example, if a two-place predicate takes a reflexive object in Bach & Partee (1980), it becomes a one-place predicate, but its single argument, the subject, fills both places of the original two-place predicate. In sentences with transitive verbs, this is accomplished formally by introducing a variable meaning into the translation in the object position of the reflexive and simultaneously introducing an operator SELF 1 into a Cooper Storage device (see Cooper (1975)). The operator is taken out of store at the VP level and takes the two-place relation as its argument; Bach & Partee (1980) describe SELF 1 as “a function which turns a two place relation R into a one-place predicate SELF1(R), which is true of an individual x just in case R holds between x and x”. Thus, there is no direct translation for himself in Bach & Partee (1980), since the effect of a reflexive is partly syncategorematic. Since Bach (p.c.) envisions a similar treatment of reflexives, I will not offer a schema for their direct translation into intensional logic. However, incorporating the use of Bach’s ENOUGH, we might develop a Langendoen-type schema for the logical form of reciprocal sentences in line with Bach’s suggestion, as in (74):

(74) \exists A(ENOUGH x \in A)(ENOUGH y, x \in A)(x \neq y & x \neq z & R(x, y) & R(z, x))

In other words, ‘there is a group A enough of whose members stand in the roles specified by the relation with enough other members of the group’.

Note that, if we add existential quantification over the set A, both (58), Langendoen’s SR, and (59), Langendoen’s WR, are subcases of (74), Bach’s reciprocity:

(58) Strong Reciprocity:

(\forall x, y \in A)(x \neq y \to x R y)
3.1. Previous Approaches

A version of Bach’s proposal which uses subsets, paralleling WRS, could be formulated to account for the relevant readings. Thus, the proposal provides flexibility in the strength of reciprocals, at the same time permitting a compositional treatment in which the limited distributivity involved in ERSs derives from the lexical contribution of the reciprocal itself, constrained by other pragmatic and lexical factors in the context.

Another way of expressing Bach’s insight might be even less explicit about the precise nature of the reciprocal relation: the truth conditions for an ERS would be something like ‘the group A (denoted by the subject) has the property of reciprocally R-ing itself (where R is the relation denoted by the verb)’. What it means to bear some relation reciprocally to oneself could vary quite a lot, depending on the particular lexical items and contextual factors involved. The one factor which all such reciprocal relations would have in common would be the anti-reflexive condition which we see in both Langendoen’s (59) and Bach’s (74), which does somehow seem central to the idea of reciprocity. Reciprocity, then, reduces to a restricted type of relation between a group and itself.

Although Langendoen is not entirely explicit on this point, I think that his central insight is that EPRSs display the same wide range of relations between two (or more) groups as ERSs display between a group and itself. I think this is essentially correct, and that it is noncoincidental, but I think his suggested logical form for such sentences, (62) (and the subset version (64)), fails for two reasons: first that there is no way to compositionally derive instances of this schema from actual EPRSs and, second, as we saw with ERSs, the schema is too strong to capture all of the types of relations which may be involved in our understanding of such sentences. With respect to the first problem, in order to compositionally derive a logical form for a sentence such as (61) which instantiates the general schema (62), one of two possibilities must be the case: one of these is that one or both of the group denoting NPs contributes some quantificational force to the logical form. But neither the universal quantification nor the existential quantification in (62) can be lexically inherent in one or the other of the two NPs in a sentence such as (61) — if they are switched a very similar meaning results. Alternatively, the quantificational force would have to arise syncategorematically, but it is very difficult to see how this could be handled in a compositional fashion. The rule would have to have access to information about the two NPs to the effect that both were group denoting, and thus simultaneously to information about both

14One of these is that one or both of the group denoting NPs contributes some quantificational force to the logical form. But neither the universal quantification nor the existential quantification in (62) can be lexically inherent in one or the other of the two NPs in a sentence such as (61) — if they are switched a very similar meaning results. Alternatively, the quantificational force would have to arise syncategorematically, but it is very difficult to see how this could be handled in a compositional fashion. The rule would have to have access to information about the two NPs to the effect that both were group denoting, and thus simultaneously to information about two.

15Actually, (77) is just the simplest logical form for the reading I intend. Either NP may be quantified in or quantifier raised, though in cases where both NPs are group denoting, this will not make a difference in truth conditions.
(77) \( R(A, B) \)

Here, the set denoted by \( A \) and the set denoted by \( B \) stand in the relation \( R \). For (61), the group denoted by the \textit{women} and that denoted by the \textit{prisoners} stand in the \textit{release} relation. But of course, \textit{release} has certain presuppositions, such as that the entity released was formerly in captivity; also, given our real world knowledge about such things, there are certain subtasks related to releasing someone.\(^{16}\) If the agent in this relation is a single individual, we generally assume that this individual performed all the subtasks (though even this is not necessarily so; compare the governor released the prisoners, which is similar in this respect to (76)).

On the other hand, when we assign agentivity to a group, as in (61), we may make certain assumptions about the roles of the individuals in the group. G. Carlson (1977, p.61ff.) discusses the relation between a collective entity and its members, in connection with the following examples with the singular definite subject \textit{the battalion}:

(78) (a) The battalion was wiped out.
(b) The battalion is quite tired now.
(c) The first battalion handles ammunition.
(d) The battalion shifted its position slightly.
(e) The first battalion has served its country for 200 years.
(f) The battalion has been dismantled.

As he points out, in these sentences there is a wide range of types of entailments concerning the individual members of the battalion. He writes (p.62),

In place of searching for some quantifier that ranges over members of the battalion, the question becomes one of discovering how it is that we recognize as true or false certain things said about the battalion. In this process of recognition, we bring with us a whole set of assumptions about the world and how real armies work in the world. In many cases, we can infer what sort of quantification would be appropriate. But this process of inference, I believe, should be viewed as an extra-grammatical process, and hence beyond the scope of semantics (as the area of semantics is conceived here).

How do we individuate groups? Consider L. Carlson’s (1980) discussion of what it is to be a “set theoretical individual”:

\(^{16}\)Again, I have been influenced in this discussion by recent work by Dowty (1986) on the relation between distributivity and what he calls the lexical ‘subentailments’ of particular verbs.

\(3.1. \text{PREVIOUS APPROACHES} \)

A plurality can always be identified by identifying each of its members individually. But significantly, there are other means in use too. One can identify a plurality with the same or similar means with which one identifies singular objects: by space-time localization, and by functional considerations (describing the internal organization and external function of the plurality).

In some cases, the functional cohesion of a group may be a question of socially defined roles. Consider (79), similar to G. Carlson’s examples, but with a plural definite subject:

(79) The Marines invaded Grenada.

(79) is true, although not all members of the U.S. Marine Corps went to Grenada. We think of the Marines as an organized body by virtue of their official function as a military body, waging war. The entire body cooperates in a venture such as an invasion, but this may entail subtasks which are not directly involved, including maintaining proper functioning of bases in home territory, getting money from Congress, and the like. (75) is another case where we are likely to view the subject as such an organized body.

In (61), however, we have no basis for considering the group denoted by the subject to be cohesive, apart from their sex, except on the assumption that they all participated directly in the release. This combination of lexical information and world knowledge thus suggests inferences about the members of the group which are not licensed by the truth conditions of (61) alone.

Now let us return to reconsider Scha’s claim that there is a cumulative reading which is distinct from the collective. Note that both Scha’s examples of collective quantification, as in (42) and (44), and his example of cumulative quantification, (50), have truth conditions which are subsumed under Langendoen’s schema for the truth conditions of EPRs, (64), as well as my schema (77):

(42) The squares contain the circles.
(44) The sides of rectangle 1 cross the sides of rectangle 2.
(50) 600 Dutch firms have 5000 American computers.
(64) \[(\forall x \in A)(\exists y \in B)(y \neq y \in B)(x \neq x \in B)(\forall x \in A)(x \neq x \in B)(\forall x \in A)(x \neq x \in B)(\forall x \in A)(x \neq x \in B)] \]
Further, Scha’s cumulative readings arise with precisely the same determiners which give rise to the collective readings, so there seem to be only two classes of determiners, the distributive and the collective-cumulative.

Barbara Partee (1985) has proposed that there is no fundamental semantic distinction between Scha’s collective and cumulative readings. In each case, the group denoted by the subject, whether the NP is definite or indefinite, simply bears the relation denoted by the verb to the group denoted by the object. It is this generalized reading, subsuming the cumulative readings of Scha as well as those where two plural arguments each receive a collective interpretation, which I will refer to as the group-group reading. Depending on the lexical characteristics of the particular verb and NPs involved, our real world knowledge about their denotations, and other contextual factors, we may draw further implications from such a sentence about the nature of the involvement of individual members of the groups denoted, including (64) as one subcase, but these implications are not themselves a part of the semantic representation of the sentence. Link (to appear) makes a similar proposal about cumulative readings, assigning them a group-group interpretation (which he terms ‘CC’, for ‘collective-collective’) in the model. He notes that such an approach raises the question: “where does the line of demarcation run between proper readings and mere models realizing a reading?”

David Gil (1982) argues (on the basis of questionnaires administered in several languages) that what I have called the group-group reading is very common. He distinguishes four types of readings of sentences which contain two plural NPs: two readings where the NPs are interpreted distributively (with earlier relative scope), a reading which he calls the “strong symmetrical”, which corresponds to Langendoen’s SR, and the “weak symmetrical”, which corresponds to Langendoen’s WR. He doesn’t discuss the clear group-group readings, but he notes (p.453) that all the readings, including both kinds of symmetrical readings, can be easily represented by quantification over sets.

Gil’s findings indicate that the symmetrical readings are strongly preferred over the asymmetrical readings by those who filled out his questionnaire, and, in addition, that the strong symmetrical reading is preferred over the weak. I think this result is interesting in that it shows clearly the availability of symmetrical, or group-group readings. However, I have some reservations about accepting his conclusions. Unfortunately, Gil includes only a few of the examples which he used in his questionnaire. I note that those he does mention are all of the form numeral $CN_{pl}$-verb-numeral $CN_{pl}$. Thus, although he wants to use his results to make claims about general features of the representation of quantified NPs, he seems to have restricted himself to a class of NPs which may all be group denoting.

If his examples were, in fact, all of this form, the results favoring the symmetric interpretations would provide further evidence that NPs with numeral determiners are unambiguously group denoting, since the distributive reading appears to be highly marked, but would say nothing about the readings available for distributive NPs. Further, the only verbs which he uses in his examples are read, see, and run, verbs which Bennett and others have called distributive-only. Thus, the preference for the strong symmetrical reading over the weak may have to do with lexical entailments of these particular verbs, rather than a global preference for strong over weak symmetry.

### 3.1.4 Summary

In this section, we have considered three kinds of approaches to the distributive/group distinction. On type of theory, represented by Lakoff’s proposal, bases the distinction on whether or not an NP has undergone quantifier movement. This approach correctly identifies distributivity in the quantificational force of certain NPs, but confuses distributivity with the NP’s scope. Another approach, represented by Bennett’s theory, sees the distinction principally in the lexical characteristics of predicates; but this tends to encode too much lexical and pragmatic information in the syntactically-driven aspects of truth conditional interpretation, and, in addition, fails to recognize the systematic contribution of determiners and adverbal elements to distributive readings. The third approach, represented by Scha’s theory with its emphasis on the contribution of determiners to distributivity, also fails to account for the full range of distributive readings, since it seems that even those plural NPs which are most clearly group denoting, those with a definite determiner, may sometimes occur with a distributive reading.

Finally, I reviewed a range of examples involving group denoting NPs, including Langendoen’s reciprocal sentences, those with multiple definite plural NPs, and Scha’s cumulative examples, and argued that they all receive a group-group reading. Except for the possible exception of the reciprocal sentences, where the reciprocal itself may contribute a distributive element to the truth conditions of the sentences, the apparent quantificational element in our understanding of many of these examples is not a part of their truth conditions, but is only implied on the basis of various types of pragmatic factors, including lexical and contextual elements and world knowledge.

### 3.2 Groups as individuals

One of the central requirements for an adequate theory of distributivity is a theory of the semantics of group denoting expressions. In Bennett’s (1974) theory of plurality, groups in the model were sets of individuals. This led to problems.

---

This has been argued by various authors, including Hoeksema (1983), Kadmon (1988) and Link (1986).
because of the difference in the types of individual and group level NPs. Schu (1981) overcame this problem by making both singular and plural NPs denote sets, the former singleton sets. In a series of recent papers (1983, 1984, 1988, to appear), Godehard Link has developed a semantics of plurality which treats groups as distinct individuals in a model, instead of sets. I believe Link makes two important contributions to our understanding of plurality. One is his emphasis on the interpretation of groups as individuals in their own right, and the other is his exploration of the relation between the count and mass domains by the use of a lattice structured domain. In this section I will explore some of the consequences of Link’s view of groups for a theory of distributivity.18 First I will consider Link’s proposal in detail. Then I will turn to a more general discussion in light of Link’s theory of three topics central to the semantics of plurality: the interpretation of conjoined NPs, the relationship between syntactic and semantic plurality, and the phenomenon which Link has called “plural quantification.” The view of the semantics of groups and plurality which emerges will lay the foundation for much of the discussion in subsequent sections.

### 3.2.1 Link’s use of lattice structures

The central hypothesis of Link’s (1983) “The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach” is that plural NPs denote individuals, just like singular NPs, rather than sets of individuals. He is able to implement this idea while obtaining the proper truth conditions for sentences in which plurals occur by giving the domain of individuals in his model the organization of a complete join semi-lattice. Before I consider his proposal in detail, I will briefly describe the mathematical structures called lattices, for those who are not familiar with them.

A lattice is a partially ordered set in which each two-element subset has both a supremum and an infimum.19 A partially ordered set is an ordered pair consisting of a set and a relation over its members. Typically, this relation is symbolized by ‘≤’; this translates informally as ‘is less than (or equal to)’. In a partially ordered set \((X, ≤)\), the relation has three properties:

1. It is reflexive, so that for any member \(x\) of \(X\), \(x ≤ x\),
2. It is antisymmetric, so that for any two members of \(X\), \(x\) and \(y\), if \(x ≤ y\) and \(y ≤ x\), then \(x = y\), and
3. It is transitive, so that for any members \(x\), \(y\), and \(z\), if \(x ≤ y\) and \(y ≤ z\), then \(x ≤ z\).20

### 3.2. Groups as individuals

The definition of a lattice also involves the terms supremum and infimum. These may be defined as follows: If \(A\) is a subset of \(X\), an upper bound for \(A\) is an element of \(X\) such that for any element of \(A\), say \(a\), \(a ≤ x\). Similarly, a lower bound is any \(x \in X\) such that for all \(a \in A\), \(x ≤ a\). The supremum, or least upper bound, of some subset \(A\) of a partially ordered set \(X\) is an upper bound of \(A\) which is a lower bound for the set of all upper bounds for \(A\) in \(X\). Similarly, the infimum of \(A\) is the greatest lower bound of \(A\) in \(X\), that is, a lower bound of \(A\) which is an upper bound for the set of all lower bounds for \(A\) in \(X\). Intuitively, the supremum of some subset \(A\) of \(X\) is the smallest thing in \(X\) which is greater than (or equal to) all the elements of \(A\). It can be shown that any finite non-empty subset of a lattice, and not just those with two members, has both a supremum and an infimum. Furthermore, we say that a lattice is complete if any (possibly infinite) subset of the lattice has both a supremum and an infimum.

A number of structures with which most of us are familiar are instances of lattices. For example, if \(X\) is any set, \(\mathcal{P}(X)\) denotes the power set of \(X\), and \(\subseteq\) is the relation of set inclusion, then \((\mathcal{P}(X), \subseteq)\) is a lattice. For any two sets \(A\) and \(B\) in \(\mathcal{P}(X)\), the supremum of \(\{A, B\}\) is \(A \cup B\) and its infimum is \(A \cap B\). A Boolean algebra is also a kind of lattice.

The structure which Link uses to organize \(E\), the domain of individuals in his model, is a complete join semilattice. This is a partially ordered set in which any (potentially infinite) subset has a supremum, but there is no guarantee that each subset has an infimum. The semilattice on \(E\) is determined by the i-sum operator, \(\vee_i\), which induces a part-whole ordering, \(\subset_i\), on the individuals in the domain. \(\vee_i\) has a syntactic counterpart, \(\oplus\). The i-sum of two individuals is their supremum in the lattice. Informally, what this means is that for any atomic individuals in the domain, say Annie, \(a\), and Bernard \(b\), there is another individual, \(c = a \vee_i b\), which is the i-sum of \(a\) and \(b\). That is, which is the denotation of \(\text{Annie} \oplus \text{Bernard}\) (Annie and Bernard). \(a\) and \(b\) both then stand in the i-part relation, \(\subset_i\), to \(c\). Furthermore, if there is another individual Danny, \(d\), in the model, then there is a further individual \(e = d \oplus \{a \oplus b\}\), which is the i-sum of \(d\) and \(c\) and is the denotation of \(\text{Danny} \oplus \{\text{Annie} \oplus \text{Bernard}\}\) (Danny and Annie and Bernard). For convenience, I will sometimes call an individual such as \(c\) an i-sum and talk about its i-parts \(a\) and \(b\). Because it is a distinct individual, \(c\) may have properties which neither \(a\) nor \(b\) have; for example, \(c\) may have the property of being a couple, though neither \(a\) nor \(b\) does.

Since the semi-lattice is complete, Link can also define an abstraction operator \(\sigma\), which forms an individual term of the form \(\sigma x \mathcal{P} x\), where \(P\) is a one-place predicate. Such a term denotes the i-sum of all individuals that are \(P\)s, which is an individual itself. In Link (1988), he points out that \(\sigma\) is actually a generalization of the iota operator: “If the extension of \(P\) is a singleton set then \(\sigma x \mathcal{P} x\) and \(\iota x \mathcal{P} x\) denote the same thing, viz. the unique element in this set.”

---

18See also Landman (1989a, b) for further discussion of the semantics of groups.
19This definition is drawn from the first chapter of Bell & Slomson (1969), on which this discussion of lattices is based.
20If the set were totally ordered, then the relation would also have the property of dichotomy, that is for any two members \(x\) and \(y\), either \(x ≤ y\) or \(y ≤ x\).
CHAPTER 3. DISTRIBUTIVITY

$E$ has a subset $A$ which contains all the atoms of $E$, where an atom is a singular object. In $A$, in addition to individuals which correspond to objects in the ordinary sense of that word, Link includes as atomic individuals all "individual portions of matter". These form a subset $D$ of $A$. The elements of $D$ are ordered by another relation, the material part-whole relationship denoted by $\leq$, which induces another complete join semilattice, though this one is not necessarily atomic. There is also an order-preserving semilattice homomorphism from $EA(0)$ onto $D$; essentially, this maps any (atomic or non-atomic) individual in the domain onto its corresponding portion of matter. Hence, we can talk about the material correlate of a particular individual without identifying the two. Link claims that this offers some interesting insights into longstanding puzzles, such as how (80) can be given a consistent interpretation:\footnote{But see Bach (1986) for discussion of a problem with Link's treatment of this example.}

(80) This ring is new, but the gold which constitutes it is old.

As I discussed briefly in Section 3.1.2, Link's general approach to distributivity is that it is a lexical property of one-place predicates. He defines a special class of Distributive Predicates (which I will abbreviate as DistrPs), which includes any predicate whose extension contains only atomic individuals (see his definition D19, p.314). His syntax guarantees that DistrPs are combined with the special "*" operator when they are predicated (via his rule T4) of a plural subject. This operator works on one-place predicates to generate all the i-sums of members of the extension of $P$, $*P$. Formally, the extension of $*P$ is the complete join semilattice generated by the extension of $P$, i.e. all the individuals in the extension of $P$ and their closure under the i-sum operator (all the groups which they form). He then offers special axioms and meaning postulates for DistrPs. His theorem (T.10), p.316, guarantees that a "*d DistrP will be distributed over the i-parts of the extension of any individual it is predicated of:

T.10 Distr(P) $\rightarrow \forall x \forall y[(*P(y) \& \text{-i-part-of}(x, y) \rightarrow *P(x)]\footnote{a \& b is the material fusion of a and b, a single, atomic individual.}

From all this, Link is able to guarantee the valid inference from (31a) to (31b), on the assumption that the predicate 'be a pop star' is a DistrP:

(31) (a) John, Paul, George, and Ringo are pop stars.
    (b) Paul is a pop star.

The denotation of Paul is an i-part of the denotation of John, Paul, George, and Ringo; the DistrP in (31a) carries the "*" and is subject to the distributivity theorem T10 on DistrPs; and hence what is predicated of the i-sum is predicated of the i-part.

Predicates other than DistrPs may have mixed extensions, including both atomic individuals and i-sums, but they cannot be starred when predicated of a plural subject, since the "*" only applies to DistrPs. Some nonDistrPs may occasionally have a distributive meaning, but this requires the use of another operator, to be discussed below.

Among the non-distributive predicates are the mass terms (MTs). The denotation of a MT is a complete subsemilattice of the subdomain $D$, that is, it includes some set of atomic portions of matter and all their i-sums.

Both plural count nouns and mass nouns have what Link calls the cumulative reference property:

(81) If a is water and b is water then the sum of a and b is water.

(82) If the animals in this camp are horses and the animals in that camp are horses, then the animals in both camps are horses.

This property is encoded in theorems (T.11) and (T.12):

T.11 $\forall x \forall y(*P(x) \& *P(y) \rightarrow *P(x \oplus y))$

T.12 $\forall x \forall y(*P(x \oplus y) \rightarrow *P(y)) \text{ for } P\in\text{MT}\footnote{In Link (to appear), which was written before Link (1986), D seems to play a slightly different role than the simple, general operator I describe in the text. For example, it is defined only over DistrPs; and transitive verbs may be marked with a double D preceding, indicating that they are distributive over both arguments. The way I use the operator here, it may apply to any predicate whatsoever, DistrP or not, syntactic or}$

As Link notes, "The set approach to plural objects does not carry over to the case of mass terms, thus missing the structural analogy between the two cases. Inherent in the notion of a set is atomicity which is not present in the linguistic behaviour of mass terms." (1983,p.305)

There is another group of predicates which are neither DistrP nor MP. Among these are what he calls mixed predicates — those which sometimes receive a distributive interpretation, other times collective. As with other nonDistrPs, these may have mixed, atomic and nonatomic denotations. In Link (to appear) and Link (1986), he introduces an adverbial operator, $D$, which serves to introduce distributivity in the cases where such predicates receive the distributive interpretation:\footnote{In Link (to appear), which was written before Link (1986), D seems to play a slightly different role than the simple, general operator I describe in the text. For example, it is defined only over DistrPs; and transitive verbs may be marked with a double D preceding, indicating that they are distributive over both arguments. The way I use the operator here, it may apply to any predicate whatsoever, DistrP or not, syntactic or}
(83) \[ DVP := \lambda x \forall y \{ \text{atomic-i-part-of}(y, x) \rightarrow \text{VP}(y) \} \]

\( D \) serves the same function in two other kinds of cases. First, Link adopts Dowty & Brodie's (1984) thesis, which we will consider in Section 3.3.1, that floated quantifiers such as \textit{each} are adverbials which induce distributivity. He uses \( D \) to translate these adverbials. (84a), with a mixed predicate, has a collective interpretation which does not imply (84b), but if an adverbial applies to the predicate, implicit \( D \) as in (a'), or the explicit \textit{each} as in (a''), the result does imply (b):

(84) (a) Bill, Pete, Hank, and Dan lifted a piano. 
(a') Bill, Pete, Hank, and Dan \( D \{ \text{lifted a piano} \} \). 
(a'') Bill, Pete, Hank, and Dan each lifted a piano. 
(b) Pete lifted a piano.

Also, \( D \) may apply to a VP derived by lambda abstraction, so that, for example, the indirect object in (85) below may be interpreted distributively. A (simplified) logical form for (85) which incorporates \( D \) is shown in (86):

(85) John gave a pumpkin pie to two girls.

(86) \( \text{(two girls) } D \{ \lambda x (\text{gave}(j, a, x, z)) \} \)

Here, the property of having been given a pie by John is predicted distributively of two girls. When we apply the translation of \( D \) in (83) to the predicate it operates on in (86), the result is as in (87a), which reduces by lambda conversion to (87b):

(87) (a) \[ \lambda x \forall y \{ \text{atomic-i-part-of}(y, x) \rightarrow (\lambda z (\text{gave}(j, a, x, z))(y))(\text{two girls}) \} \]
(b) \[ \forall y \{ \text{atomic-i-part-of}(y, \text{two girls}) \rightarrow (\text{gave}(j, a, x, y)) \} \]

3.2. GROUPS AS INDIVIDUALS

so that each of the two members of the group denoted by \textit{two girls} (atomic i-parts of the plural individual) has the property of having been given a pie by John.

In Link (1986), he extends and revises his (1983) theory of the logic of plural and mass terms to make it compatible with the generalized quantifier framework. Recall that in this framework, an NP does not denote an individual (whether atomic or plural), but a set of properties; for example, \textit{a woman} denotes the set of all properties which some woman has, while \textit{John} denotes the set of all the properties which John has. The conjoined NP \textit{John and Mary}, i.e. the term \textit{John@Mary}, denotes a third set of properties, those of the individual \textit{john'\&mary'}. Since this individual may have properties which neither John nor Mary has, e.g. that of being a couple, then this set of properties is not just the intersection of the properties which John has with those which Mary has. This translation of \textit{Mary and John} into a generalized quantifier assumes that \textit{and} translates as \( \odot \). There is another type of \textit{and}, the distributive or Boolean \textit{and}, and we will briefly consider this and related issues pertaining to the semantics of conjunction in Section 3.2.2.

Finally, Link (1984, to appear) introduces a function \( h \) from (nonatomic) i-sums in the domain \( E \) onto what he calls impure atoms. Thus, every i-sum has an impure atomic correlate. Though the impure atoms are a subset of the atomic subdomain \( A \), they are distinct from the "pure atoms", such as the denotation of \textit{John}. When a syntactically plural term, such as \textit{the Beatles}, is enclosed in angled brackets, as in \textit{(the Beatles)}, the result denotes an impure atom. Link's motivation for this device seems to lie principally in the analysis of examples involving NP conjunction and distributivity, and we will review these in the following section.

3.2.2 NP conjunction and distributivity

In this section, I will review some existing analyses of NP conjunction in its relation to distributivity. This is not intended to be a complete theory of conjunction, which involves a number of questions which are beyond the purview of this work. Rather, I will focus on three aspects of conjunction which are directly relevant to the question of distributivity and the semantics of groups. The first is the existence of (at least) two types of NP conjunction, collective vs. intersective, or Boolean, conjunction; I will give a preliminary characterization of the two types of conjunction and point out the types of examples which argue for the distinction. The second issue which I will briefly consider involves the conjunction of quantificational with nonquantificational NPs, and in particular whether some unacceptable examples of this may be taken to argue for a movement analysis of distributivity. Finally, I will look at examples which Link (1984) takes to motivate the mapping from i-sums onto impure atoms.

See, for example, Gazdar (1980), Rooth & Partee (1982), and Partee & Rooth (1983) for more general discussion of conjunction. See Link (1984) for discussion of a particular type of conjoined NP, the hydas.
Several authors have recognized that there are two basic types of conjunction, though terminology and formal analyses differ. Hoeksema (1983) calls the two kinds of conjunction, “collective” and “intersective”; these are exemplified in (88) and (89), along with Hoeksema’s proposed syntactic and semantic schemas for each type:  

(88) Joe and Ellen have a new dog.  

**COLLECTIVE CONJUNCTION:**  
NP[+pl] & NP[+pl] → NP[+pl]  
\[[a \& b\text{-}\text{coll}}]VP\] = VP’(a’ ⊕ b’)

(89) Each man and each woman has a new dog.  

**INTERSECTIVE CONJUNCTION:**  
NP[-pl] & NP[-pl] → NP[-pl]  
\[[a \& b\text{-}\text{int}}]VP\] = VP’(a’) ∩ VP’(b’)

Van Eijck (1983) discusses the same two types of conjunction in terms of what he calls “the scope of and”. For him, in the collective interpretation and has narrow scope, while in the intersective interpretation, it has wide scope. Van Eijck treats these cases in the context of Discourse Representation Theory, though I won’t go into the details of his proposal here. He also points out that there is another type of conjunction, “appositional coordination”, illustrated by (92):

(92) His aged servant and the subsequent editor of his collected papers was with him at his death-bed.

The two conjuncts of the subject of (92) are just two different descriptions of the same person, with the result that the conjunct is singular in number, as is the verb. When was is replaced by were, a different, collective interpretation results. Van Eijck treats appositive conjunction with “wide scope and” as well. This amounts to the truth conditions of ‘his aged servant was with him at his death-bed and the subsequent editor of his collected papers was with him at his death-bed’. But this conjunction reduction paraphrase may be true where the servant is not the same person as the editor, whereas (92) does not seem felicitous in that case. The singular verb seems to introduce a presupposition in this case that the two NPs denote the same individual, i.e. that the conjoined subject is pragmatically singular. This would account for the apparent difference between (92) and the conjunction reduction paraphrase.

Massey (1976) gives the following contrasting examples (slightly modified by me):

(93) Gerald Ford and Pierre Trudeau weigh over 270 lbs.  

(94) Joe Green and The Refrigerator weigh over 270 lbs.

---

26. The syntactic and semantic schemas in these examples are not Hoeksema’s, but I think they preserve the spirit of his proposal.

27. The subjects denote an ex-president of the United States and an ex-prime minister of Canada, respectively.

28. American football players, with the Pittsburgh Steelers and the Chicago Bears, respectively.
The preferred interpretation in (93) is one where the i-sum of Gerald Ford and Pierre Trudeau weighs over 270 lbs. (94) would generally be taken to mean that each of the two individuals denoted weighs over 270 lbs. These readings are strongly preferred on pragmatic grounds alone. Note that, though (94) on this interpretation has the same truth conditions as the conjunction reduction paraphrase, ‘Joe Green weighs over 270 lbs. and The Refrigerator weighs over 270 lbs,’ the number of the verb in this example doesn’t warrant the intersection interpretation of and. But, as Link (1986) points out in conjunction with a similar pair of readings, the interactive interpretation isn’t necessary in (94) in order to derive the correct truth conditions. Rather, the conjunction may be interpreted collectively, but the predicate is interpreted distributively using the $D$ operator. The truth conditional result is the same as that of interactive conjunction. Thus, it seems that the crucial examples which argue for interactive conjunction are those which involve (a) nonplural agreement on the verb, as in (89) and (92), and/or (b) distributive (non-group denoting) NPs as conjuncts, as in (89) and (90).

Consider again (90). Collective conjunction is not possible in this example because it involves forming i-sums of individuals in the domain, and distributive NPs do not denote such individuals. Recall that Lakoff (1972) offers examples which purport to show that only individual or group denoting NPs may coordinate with proper names; compare (95) and his (18d):

(95) The philosophers and John are similar.
(18d) * Few philosophers and John are similar.

As discussed in Section 3.1.1, Lakoff attributes the unacceptability of (18d) to the inability of Quantifier Lowering to operate into coordinate structures, since this would violate the Coordinate Structure Constraint. We have seen that a descriptively adequate theory of quantifier movement must, like Bennett’s (1974) theory, permit the Lowering, Raising, or Quantifying In of group denoting, as well as quantificational NPs, since the former may take scope over the latter. One might, however, claim, that the movement of group denoting NPs is optional, whereas quantificational NPs have to move, as in May (1977). This difference, then, would be taken to account for the difference in acceptability between (95) and (18d), since few philosophers would have to be lowered, raised or quantified in, violating the Coordinate Structure Constraint, while the philosophers may be interpreted in situ. However, such a theory would predict that quantificational NPs may never occur in coordinate structures; but we have examples where they do, as in (89) and (90) above, and in (96) and (97) as well:

(96) Both apartments and most of the mobile homes have a fire extinguisher in the kitchen.
(97) Few tv comedies and even fewer children’s shows have any redeeming content.

(96) may only mean that both apartments have a fire extinguisher in the kitchen and most of the mobile homes have a fire extinguisher in the kitchen, an interpretive conjunct. (97) has a similar interpretive interpretation. It is generally conceded that both is strictly distributive, and most and few seem to be as well. Hence, it seems that distributive NPs may occur in coordinate structures so long as the conjunction is interpreted intersectively. The unacceptability of (18d), then, would arise from the fact that one of the conjuncts in the paraphrase of its conjunction reduction interpretation, ‘John is similar’, is uninterpretable, since similar requires a group subject.

What then of the conjunction of quantificational and nonquantificational NPs? Many examples of this sort seem odd to me, but they may be grammatical. Most of them seem to have only intersective readings, which is what one would expect given the above discussion. Consider the following:

(98) John and every kid on the block ate an ice cream cone last night.
(99) The members of the drama club and every visiting director sent a press release to the local paper.
(100) The members of the drama club and every visiting director will attend a banquet Thursday evening.

All these seem a bit awkward to me, but I think they may be grammatical. (98) has only an intersective reading, where John ate an ice cream cone and each kid ate an ice cream cone. (If this example is acceptable, it is a direct counterexample to Lakoff’s claim that proper names and quantificational NPs may not be conjoined.) Likewise, I think that (99) has only the intersective reading, where the directors, for whatever reason, sent different press releases from the one sent by the members of the club. (100) may have two readings. One is an intersective reading, where there will be one banquet attended by the members of the club and a possibly different banquet attended by each of the directors. But there may also be a reading where all of the members and the visiting directors will attend the same banquet; I suspect that this reading is forced by the pragmatics of the predicate — i.e., the likelihood that these people might have something in common to celebrate and the unlikely nature of a situation where each had different banquets to attend on the same Thursday night.
CHAPTER 3. DISTRIBUTIVITY

Mats Rooth (p.c.) has pointed out another sort of example where the quantificational conjuncts do not have a straightforward interactive reading:

\( (101) \quad \text{Each professor and some student of his are writing a paper together.} \)

The most likely interpretation of (101) in a null context might be paraphrased, "for each professor \( x \), \( x \) and some student of \( x \) are writing a paper together". Thus, the first conjunct appears to have wide scope over the entire sentence. Under the scope of this operator, a group formed by collective conjunction, and consisting of \( x \) and \( x \)'s student, has a property which may only be true of groups. Prima facia, this sort of example appears to be a violation of the coordinate structure constraint of Ross (1967). However, an adequate assessment of its import would require a deeper investigation into the semantics of conjunction and, in particular, its relation to quantifier scope, than I am prepared to undertake here.

Finally, let us turn to the example which motivates Link's (1984) mapping from i-sums onto impure atoms:

\( (102) \quad \text{The Leitches and the Latches like each other.} \)

On the assumption that there are two Leitches and two Latches, Link claims that there are three ways to interpret this example. One reading results from treating each of the conjoined NPs as denoting a nonatomic i-sum, so that the whole subject denotes a single i-sum with four atomic i-parts. As we discussed in Section 3.1.3.2, a reciprocal predicate is collective; the reciprocal itself each other then suggests or entails that the relation of liking hold between atomic i-parts of the group denoted by the subject. This is the reading which is prominent if we introduce floated all:

\( (103) \quad \text{The Leitches and the Latches all like each other.} \)

However, there are two other readings of (102) which cannot be obtained if each of the conjuncts of the subject denotes an i-sum. Link uses these readings to argue that the conjuncts may each denote an impure atomic individual, or couple.

When the plural conjuncts of the subject of (102) are interpreted as impure atoms, the logical form of the entire subject is as in (104):

\( (104) \quad \lambda pp\langle (\sigma x \text{Leitch}'(x)) \otimes (\sigma x \text{Latch}'(x)) \rangle \)

3.2. GROUPS AS INDIVIDUALS

Now the predicate may be interpreted in two ways: If it is collective, the group denoted by (104), which contains just two atomic i-parts, has the property of liking each other. On this interpretation, each couple likes the other couple. If the predicate itself is interpreted distributively, then the like each other property will be predicated of each couple. The further lexical distributive effect of the reciprocal then appears to be able to act on the individual members of each couple, even though the couples are themselves atomic individuals. Think of this reading as a retort to someone else's utterance of I don't know of any couples where the husband and wife like each other. In this respect, the distributivity of the reciprocal seems stronger than that of the \( D \) operator, in that it can entail relations among subparts of an atomic individual.

Given this approach to conjunction, we can analyze another type of example:

\( (105) \quad \text{The pitchers and catchers were practicing their signals.} \)

Suppose we regard the plurality which is morphologically realized on both of the two CN heads of the subject as a single operator having wide scope over the whole conjunct, and treat the denotation of pitcher and catcher as an impure atomic individual. Then by interpreting the predicate distributively, we are able to derive a reading of (105) where each pitcher/catcher pair has the property of practicing their own signals (i.e., the signals are not common to the entire set consisting of all of the pitchers and all of the catchers).

There are other CN conjuncts, such as husband and wife, which also seem to have this atomic pair interpretation, perhaps even as a convention on their use. Note the fixed order of the conjuncts. Also, there is a possible interpretation of examples such as (106) where normally distributive predicates may be said to hold only of the ordered pair individual:

\( (106) \quad \text{The husband and wife ate an ice cream cone/smoked a cigarette.} \)

It seems that the couple may have shared an ice cream cone or cigarette, though it is also possible to get a distributive interpretation, where each ate a cone or smoked a cigarette.

3.2.3 SYNTACTIC VS. SEMANTIC PLURALITY

There are a variety of ways in which the general notion of number is reflected in the syntax of natural language. In English and related languages, two of the most important are morphologically realized: the number of nouns and that of verbs. Here I want to discuss the relation between these syntactic reflexes of number and the semantics of plurality. First I will discuss verb agreement, beginning with a review of Hoeksema (1983), where an intimate relation is posited between
CHAPTER 3. DISTRIBUTIVITY

the number of a verb and semantic interpretation. I will argue that this view is incorrect, and that instead we should view verbal number as a purely syntactic agreement phenomenon. Then I will turn to consider plurality in CNs. There, in contrast to verb agreement, I believe that syntactic number is directly reflected in semantic interpretation, and I will show how this might be reflected in a Link-type semantics of plurality.

Like Link, Hoeksema (1983) uses a structured domain and views distributivity as a property of predicates. However, his structured domain contains two kinds of entities, individuals and groups. In this theory, singular nouns and VPs denote sets of individuals, while plural nouns denote sets of groups, and plural VPs denote mixed sets containing both individuals and groups. Syntactic number is taken to be directly related to semantic number ("number concord is analyzed as a semantic phenomenon", p. 71), and thus the requirement of subject-verb agreement acts as a semantic filter as well. Distributivity is regarded as a lexical property of one-place predicates, introduced by means of meaning postulates.

There are several problems with Hoeksema’s approach which make Link’s preferable. Hoeksema himself points out that there is often a disagreement in number between subject and predicate. Some examples involve there insertion constructions, as in his (107); others, such as (108), involve nouns which are syntactically singular in number, but plural in the sense that their singular extension usually contains groups in some pretheoretic sense:

(107) There’s all sorts of explanation.

(108) The committee gather. (British English)

Jespersen (1911, Vol.II, Chap.6) offers a large number of examples which display a lack of number concord, from which the following are drawn:

(109) (p.170) from Defoe: Not one in ten of them write it so bad.

(110) (p.172) from Wells: Do you mean to say, neither of you know your own numbers? [cf: neither of them know(s)]

(111) (p.174) from Shaw: Public and private life become daily more theatrical.

(112) (p.180) from Norris: incoherences, to which nobody, not even themselves, were listening

Also, Barbara Partee (p.c.) has pointed out that VP ellipsis presents a problem for this approach when the subject of one conjunct is singular, the other plural, since for Hoeksema, singular and plural VPs have different types of denotations.

(113) John bought a house, and Bill and Mary did, too.

These problems with Hoeksema’s approach suggest that it is a mistake to posit a direct relation between the syntactic number of a verb and its semantic interpretation. Rather, the morphological number on a verb is the outcome of an agreement phenomenon, with all the conflicting demands of syntax and pragmatics which so often plague agreement (cf. problems with gender).

Notice that Link (1983) also posits a relationship between syntactic and semantic number in his use of the * operator. This is true for both CNs and VPs, but the motivation for its use differs in the two cases. In VPs, * marks DistFrPs (and only DistFrPs) with plural subjects. Unlike Hoeksema’s use of the number of the verb to trigger semantic plurality, Link’s use of * on VPs does not depend solely on verbal number, for not all verbs of plural subjects receive the *. But still, on this proposal a semantic characteristic of a particular kind of verb or predicate is determined by the syntactic number of its subject. I think this is questionable, and that verbal number is solely a question of agreement, with no direct consequence for semantic interpretation apart from its role in indicating intersective conjunction. In addition, I argued in Section 3.1.2 that it is not necessary to build aspects of the lexical content of predicates into our semantics, so that * is not necessary or desirable for the characterization of distributivity in predicates.

However, Link (1983) also uses * on morphologically plural CNs, and I think that this use may be desirable, although its applicability needs to be clarified. Recall that he defines * as an operator over one-place predicates, which would automatically suggest its use with CNs. As with its use on VPs, a *d CN denotes the elements of the semilattice generated by its singular counterpart. However, unlike the use of * with VPs, Link offers no rules for its use with CNs, nor any discussion of the relationship between syntactic (or morphological) and semantic plurality.

29Recall the similar problem which faced Bennett’s theory when, as in examples such as (36), two VPs of different types were conjoined. There, however, the types were not defined in terms of syntactic number, but of the individual/group-level distinction. A VP on the individual level could have either a singular or plural subject. For Hoeksema, the problem is both syntactic and semantic, since the two are interdependent.

30I prefer to use the term ‘syntactic plurality’ for two reasons. First, although plurality in nouns is marked morphologically in English, this is not the case universally. In Haitian Creole, for example, plurality is marked with an enclitic particle, yo, which follows the entire NP, including postnominal adjectives and deictic particles. This raises a number of questions about the locus of plurality and its compositional treatment. And, second,
Let us simply correlate syntactic plurality in CNs with the application of the $^*$ operator. We then posit a direct relation between syntactic and semantic plurality, and define the semantic plural of a CN denotation in terms of the $^*$ operator, as the semiflatic generated by its singular denotation. Note that under this definition the extension of a syntactically (and, hence, semantically) singular CN will be a subset of the extension of its plural counterpart.

Most approaches to semantic plurality seem to focus on the NP level, instead of the CN. One such approach to nominal plurality is that of van Eijck (1983), who offers the following formal definition of semantic number, where $[a]$ means the denotation of $a$:

- Call a NP-denotation $[a]$ proper iff $[a]$ is defined and $[a] \neq \emptyset$ and $[a] \neq p(U)$. [Then]
  - (i) semantic number($[a]$) = 1 iff in every model where $[a]$ is proper, $\neg(\emptyset \in [a])$ and $[a]$ contains at least one singleton set.
  - (ii) semantic number($[a]$) = 2+ iff in every model where $[a]$ is proper, $[a]$ has only sets as its elements that contain at least two members.
  - (iii) semantic number($[a]$) is undefined in all other cases.

First, note that this definition has nothing to do with CN number; for example, it makes more than one man plural. Further, in a Link-type framework, this definition would fail to make any distinction between semantically singular and plural definites. Since the denotation of a plural CN contains all the elements in the semiflatic generated by its singular denotation, in a model where there is only one atomic individual in the denotation of the singular CN, that same individual will be the sole member of the plural denotation. The girls in such a model might have the same denotation as the girl, rather than being undefined, as in the type of system van Eijck seems to have in mind. Thus, the semantic number of the followed by a plural common noun would always be 1 by van Eijck's definition. I take these problems as evidence that the approach to nominal plurality in terms of properties of CNs is more useful than an approach to nominal plurality in terms of properties of NPs.

Even the question of the syntactic number of an NP is a complex one. An NP might be considered syntactically plural when its head (or N or N* constituent) is syntactically plural. But there is often some question about what the head is in a particular NP. Barbara Partee (p.c.) points out examples like the following it reminds us that although the plural morpheme in English almost always occurs on the head noun, the CN so marked may be a larger constituent, including adjectives and/or complements to the head so that we have a "bracketing paradox". We will consider such examples in conjunction with the discussion of Rooth (1986a,b) in Section 3.5.2.

In cases such as (114) where the conjuncts themselves differ in number, the number of the whole conjoined NP is in question, and resolution often depends on proximity of the verb to one or the other conjunct. In the cases in (115) and (116), there is some question whether the construction is a partitive or a pseudo-partitive, and the number of the verb vacillates accordingly. In addition, there is the question of the number of disjuncts such as John or Mary, compared with that of John and Mary. These problems with the definition of syntactic and semantic plurality of NPs suggest that the approach I have taken to nominal plurality in terms of CNs is preferable.

Semantic plurality as I have defined it sometimes makes predictions which do not correlate with our initial judgments in particular examples. For example, note the contrast between (117) and (118):

- (117) The boys write girls.
- (118) John writes girls.

We may readily interpret (117) as being true in a situation where each boy only writes one girl, but nonetheless, we expect that more than one girl is involved. Girls here is an instance of the dependent plural, and we will discuss these in some detail in Section 3.6 below. (118), on the other hand, seems to require that John write more than one girl. Yet, on the approach to plurality I have sketched, the extension of girls contains atomic individuals as well as i-sums of girls. This predicts that the utterance John has girls is true if John has written only one girl.

I think this prediction is correct, but that, although the sentence is true in this situation, it is misleading by virtue of Gricean conversational principles. This is so because a plural CN is not only semantically plural, but in cases such as (118) implies that John is involved with more than one girl. The singular NP a girl in (118) would be less likely to mislead here, since it has only atomic individuals in its extension. Then, by the Gricean cooperative principle, if I know that John has written only one girl, I should, in general, use the singular CN.

31 See Selkirk (1977) for discussion of the distinction.
**CHAPTER 3. DISTRIBUTIVITY**

However, consider a situation where a bunch of boys are talking together about the shame of condescending to write to girls:

(119) A: I can’t believe any of the guys would write girls.
B: Wrong! John writes girls — I saw him mail a letter to Mary.

I think that in this context, where the property of being an *x* such that *x* writes elements of *girl* is already salient, B’s utterance can be true and felicitous when John has only written to Mary.

Another type of example, involving VP ellipsis, also argues that the relationship between semantically singular and plural CNs which I have argued for is appropriate:

(120) Al and Steve hurt their backs and so did Mary.

Even though the direct object in the first conjunct is plural, the ellipsis here seems perfectly felicitous. This also seems to indicate that their backs must have atomic individuals in its extension, so that Mary need have hurt only one back for (120) to be true.

Because of a limited number of nouns such as *scissors* and *pants* (*plura ta*), which require classifiers such as *pair af* to be countable, and nouns which show no morphological distinction between singular and plural, such as *sheep* and *fish*, syntactic plurality does not always correlate directly with morphological form. However, apart from these exceptions, we recognize a fairly regular correlation between morphological form and syntactic number (and hence between morphological form and semantic interpretation) in CNs. In particular, in English the plural morpheme -s generally correlates with syntactic plurality.

Further, it seems that in the majority of cases the extension of a syntactically singular count CN contains only atomic individuals. The candidates for exceptions to this correlation include nouns such as *committee, police, legislature, nobility, audience*, and the like. (108) above exemplified a tendency to number agreement problems with these nouns, due to a tension between their syntactic singularity and our feeling that they denote groups. This particular example seems to be British in flavor, but there is a similar lack of agreement in American English with *police, nobility*, and other such ‘group’ nouns. Do the denotations of these singular CNs include non-atomic i-sums, or do they contain only atomic individuals, with the pragmatic implication of plurality arising from the general facts about what these things are?

In Link’s (1984) framework, given the mapping from i-sums onto their impure atomic correlates, there is a sense in which any expression whose denotation contains nonatomic i-sums also denotes atomic individuals. We might, then, say that the denotation of the singular noun *committee* contains i-sums, each of which has a number of i-parts, the members of the committee. But (*committee*) could denote a set of impure atoms, the correlates of the i-sums in the denotation of *committee* under the mapping from i-sums to impure atoms. This might be taken to reflect our sense that a committee is more than its individual members, that it has its own being; and in case we assign the impure atomic interpretation, we could still speak of the committee’s members, interpreting this expression in terms of i-parts of the i-sum correlate of the impure atom.

I don’t think there can be a pretheoretic answer to the question of the denotation of group nouns. Bach (1981a, 1985) argues, correctly I believe, for a “natural language metaphysics”; we are not to bring to our semantics preconceptions of the way the world is, but instead to let language direct us in the construction of our models. In arguing for his view of groups as individuals, Link takes a similar position, with a call for “ontological agnosticism”. In my discussion of adverbial distributivity in Section 3.3.1, I will point out certain technical consequences of giving group nouns atomic or nonatomic denotations, and suggest that within the framework I am assuming it is preferable to assume that they denote pure atomic individuals.

### 3.2.4 Plural quantification

Link (1986) claims that there is a phenomenon which he calls plural quantification, in which we quantify over i-sums rather than atomic individuals in the denotation of a CN. He illustrates this by examples such as the following:

(121) (All) competing companies have common interests.

(122) No two competing companies have common interests.

(123) Two’s company, three is a crowd.

There is a reading of (121) in which *all* quantifies not over individual companies, but over groups of competing companies. The interpretation may be paraphrased, ‘Any given i-sum of competing companies has mutual interests’. This is the only type of reading we can get for (122), with the further restriction that the i-sums of competing companies under consideration have only two i-parts each. And (123) seems to mean something like ‘a group of two is company, while a group of three is a crowd’.

As further evidence of this phenomenon, consider NPs containing obligatorily collective predicates, such as the following:
(124) twin babies
feuding neighbors
opposing forces
matching towels

It seems that the extensions of such modified CNs can only contain i-sums. So, in examples such as (125) and (126), the plural quantification reading is strongly preferred:

(125) Most twin babies love each other.

(126) Many feuding neighbors war constantly (on each other).

In both examples, the predicates are collective, so that only groups, or plural individuals, are in their extensions. Since twin babies only come in pairs, as do the two competing companies of example (122), in order to determine the truth of (125) we simply partition all the i-sums consisting of twin babies into two classes, those who love each other and those who don’t; we then compare their cardinalities, and if the former class is larger than the latter, (125) is true. Like twin babies, we generally think of feuding neighbors as coming in pairs — the expression is likely to connote the Masons and the Dixons, for example. In (126), if we only consider all the i-sums of pairs of neighbors who are feuding, and compare the number of pairs which war constantly with the number of pairs that don’t, then I think we capture our intuitive sense of the sentence. Many may

32 Or i-sums (‘sets’) of i-sums: note that these derived CNs may in some sense be semantically singular — for example, opposing forces may only refer to the i-sum of those forces which are opposed on a single issue. But one may want to talk about the characteristics of more than one such group, that is, the semantic plural of this CN. Since syntactic plurality in English only appears morphologically on the head noun, and since even the “singular” sense of this CN requires it, there can be no syntactic distinction between opposing forces as denoting an i-sum or an i-sum of i-sums.

33 Feuding and twin may occasionally be used as attributions of a single individual, meaning ‘one who is having a feud with someone else’ and ‘one who has a twin’, respectively. I don’t think this kind of reading is available for opposing or matching; though there is another intransitive, non-reflexive use of these adjectives, where the opposing or matching party is already contextually salient. In these cases, the adjective is a function, ‘opposed to x’ or ‘matching x’, which is referentially dependent on a contextually salient argument. I will not be concerned with these readings here. I think it is uncontroversial that the adjectives do have the plural individual denoting sense.

34 Barbara Partee (p.c.) points out that the plural forms of all symmetric or non-asymmetric relational nouns seem to have such readings: cf. friends, sisters, relatives.

J.2. GROUPS AS INDIVIDUALS

either have a proportional reading, where we compare the number of i-sums which have the property with the number which do not and decide if the proportion is adequate, or a context dependent cardinality reading, where we have some vague idea of how many we consider ‘many’ and determine whether the cardinality of the class of pairs which are feuding is equal to or greater than this number. But in either case, plural quantification seems to give the right results.

Now consider (127):

(127) Few people agree (with each other) on this issue.

The determiner few is strongly distributive, but agreement can only take place between individuals in a group. It does not make sense to say of an individual that she agrees with each other. So one is tempted to claim that few here quantifies over i-sums, groups of people; then the sentence would have the paraphrase ‘there are few groups of people who agree with each other on this issue’.

But Angelika Kratzer (p.c.) has pointed out that there is a problem with the plural quantification approach to the truth conditions of (127). Consider a situation in which there are four people, call them a, b, c, and d, such that three of them, a, b, and c, agree on the issue at hand, while the fourth, d, doesn’t agree with any of the others. Now consider the sublattice of E which is generated by these four people; in addition to the four atomic individuals, this sublattice contains all possible nonatomic i-sums of those atomic individuals, eleven in number. Then, Kratzer points out, if we partition these eleven i-sums into two classes, those whose members are all in agreement and those whose members are not all in agreement, then we get the following result: the class of i-sums which agree will include the four individuals denoted by a ⊕ b, a ⊕ c, b ⊕ c, and a ⊕ c ⊕ d, but the class of those whose atomic-i-parts are not all in agreement includes the seven individuals denoted by a ⊕ d, b ⊕ d, c ⊕ d, a ⊕ b ⊕ d, a ⊕ c ⊕ d, b ⊕ c ⊕ d, a ⊕ b ⊕ c ⊕ d. Thus, the number of plural individuals who are not in agreement is greater than the number of those who are, despite the fact that three out of four atomic individuals agree. This does not reflect our intuitive understanding of the proportion of agreement among the group of four atomic individuals.

Generalizing, consider any set of m atomic individuals. For any subset of m consisting of n atomic individuals which all agree, there will be a total of 2^n — (n + 1) nonatomic i-sums all of whose i-parts agree (the cardinality of the power set of a set of n individuals is 2^n, from which we subtract 0, which is not in the lattice, and the cardinality of the singleton sets, since these are atomic). In order to find the i-sums whose members are not all in agreement, we take each member of the set of non-agreers, of cardinality m — n, and add it to each element

35 Strictly speaking, ⊕ is an operation over pairs; however, a ⊕ [b ⊕ c], [a ⊕ b] ⊕ c, and [a ⊕ c] ⊕ b denote the same individual in the lattice, and I have used a ⊕ b ⊕ c here to denote this individual.
of the lattice generated by the agreeers. That is, the number of non-agreeers will be 
\((m - n)(2^n - 1)\). No matter what the number of atomic individuals of the relevant 
sort who agree with each other, the number of i-sums whose members are not all 
in agreement will always be equal to or greater than the number of i-sums whose 
members are all in agreement. Kratzer calls this the proportion problem for 
plural quantification. The problem would appear to arise not only with few, 
but with many and most, which also have proportional readings.\(^{36}\)

The proportion problem shows that in order to obtain the correct truth 
conditions in cases such as (127), not all i-sums are relevant. Rather, what seems to 
be at issue is the cardinality of the maximal collection of experts who agree on 
the issue. For example, consider a situation in which there are twenty people and 
some burning political issue. Suppose that two of them agree on one potential 
solution to the problem, while three others agree on a different solution, and none 
of the others agrees with anyone on any solution. I think the relevant question is 
‘what’s the largest number of people who agree?’ Given the answer to this question, 
in this case ‘three’, then we must determine, by some pragmatically given 
measure, whether or not this number is ‘few’. One way to represent this intuition 
might be via a logical form such as (128):

\[
\text{FEW}(\exists x[\exists y(\text{person}(y) \land \text{agree}(y) \land x = |y|) \\
\quad \land \forall z(\text{person}(z) \land \text{agree}(z) \rightarrow x \geq |z|)])
\]

where \(|y|\) means ‘the cardinality of \(y\)’, and agree\((y)\) means that all i-parts of 
\(y\) agree in the same way. (128) means that the cardinality of the largest i-sum 
of people who agree in the same way is FEW. Though (128) does involve plural 
quantification, it is over a restricted type of i-sum, those which are a group of 
people who agree in the same way, and not over all i-sums which are in the lattice 
\text{person}.

A general truth conditional schema for few which would derive (128) for the 
sentence in question is given in 129):

\[
\lambda CN \lambda VP \text{FEW}(\exists x[\exists y(\text{CN}(y) \land \text{VP}(y) \land x = |y|) \\
\quad \land \forall z(\text{CN}(z) \land \text{VP}(z) \rightarrow |x| \geq |z|)])
\]

The schema in (129) is of type \(((e, t), (e, t), t))\) (a relation between one-place 
predicates), the usual type of determiners. There are two ways in which the 
‘fewness’ of the maximal collection may be determined, corresponding to two 
readings of few: on the cardinality reading, we simply have some number in 
mind, which isn’t very large, and we check to see if the number of people who 
agree (the number of atomic i-parts of \(x\)) is less than or equal to this number. 
On the highly context sensitive proportional reading of few, we have some idea 
of what constitutes a small proportion of agreeers relative to some other set, e.g. 
the set of all people in the situation. In either case, the operator FEW may be 
seen as a predicate which takes the cardinality of a maximal collection, or i-sum, 
as argument.

Note that a parallel reading is available for cases with many, as in (130), with 
truth conditions given by the logical form in (131):

\[
\text{MANY}(\exists x[\exists y(\text{person}(y) \land \text{agree}(y) \land x = |y|) \\
\quad \land \forall z(\text{person}(z) \land \text{agree}(z) \rightarrow x \geq |z|)])
\]

This means ‘the maximal collection of experts who agree in some way is many’.

(132) gives the appropriate logical translation for many, parallel to that for few 
in (129):

\[
\lambda CN \lambda VP \text{MANY}(\exists x[\exists y(\text{CN}(y) \land \text{VP}(y) \land x = |y|) \\
\quad \land \forall z(\text{CN}(z) \land \text{VP}(z) \rightarrow |x| \geq |z|)])
\]

I believe that (129) and (132) give the proper truth conditions for a variety of 
examples, and in addition bring out the intuitive parallel between few and many. 
The monotonic characteristics of these determiners (see Barwise & Cooper 1981) 
may then be attributed to the lexical character of the particular operators FEW 
and MANY. However, I am not certain of the status of these logical schemas in 
the grammar. Note that they will not suffice as the translation of few or many 
in all examples. Angelika Kratzer (p.c.) points out that to utter Few people are 
wear matching sweaters in a situation where there are 100 people and 50 pairs 
of matching sweaters would be true on the “predicative” translation of few in 
(129), since the maximal collection of people wearing matching sweaters would be 
100. However, this is unintuitive. In this example, we seem instead to want to
quantify over i-sums which are pairs of matching sweaters, with a cardinal, rather than a proportional interpretation of few. It is possible that few (and many) are ambiguous, or it may become apparent that the truth conditions which these schemas aim to derive are actually subcases of a more general treatment of few and many.

Schein (1986) addresses a group of related examples, (133)–(136):

(133) (a) Few experts agree.
(b) Few experts ever agree.

(134) (a) Few Democrats vote with the President.
(b) Few Democrats ever vote with the President.

(135) (a) Few good students are unprepared.
(b) Few good students are ever unprepared.

(136) (a) Few advanced students collaborated on three problems.
(b) Few advanced students ever collaborated on three problems.

Schein claims that all these examples have what he calls an event dependent reading, indicated for (133a) and (133b) by a paraphrase like ‘whenever there is an event of agreeing, it involves few experts’. He claims that the same reading is available for both versions of (134)–(136), as well: in fact, as I understand his theory, he predicts the availability of such a reading for all sentences with plural NPs. The schema behind these paraphrases might be given as in (137):

(137) \forall x(\text{event}(x) \land \text{VERBing}(x) \rightarrow \text{FEW}(\text{CN} \rightarrow \in \rightarrow \exists(y)))

Schein uses the thematic role of a particular argument of the verb in these cases to further restrict the nature of the involvement of its denotation in the event, so that in (137), we might specify that the elements of the CN denotation are not involved in just any way in the event, but are involved in ways conventionally attached to the subject argument position of agree by the lexical specification of the verb.

(137) seems to have at least one reading with truth conditions in line with Schein’s event dependent paraphrase, meaning something like ‘few people ever agree on this issue’. I also agree that the event dependent paraphrase which Schein suggests comes close to the truth conditions for the examples in (133) and (136) and for one reading of (134b) and (135b), but there are a number of problems with his proposal.

12. Groups as Individuals

First, I question whether such readings are available for (134a) or (135a). Consider (134a). This seems to be a statement about the voting tendencies of individual Democrats. But I do not get a reading of this example which reflects the event dependent paraphrase ‘whenever there is a voting with the President, it involves few Democrats’. This paraphrase is not about tendencies of individuals; it is about tendencies of overall voting patterns in Congress. (The paraphrase does reflect the “event dependent” reading of (134b).) The type of paraphrase to which Schein suggests falls in (135a) for the same type of reason — it fails to reflect the fact that generalizations are being made over individuals of a certain character.

Further, there does not seem to be a reading of (130) which instantiates the schema in (137), i.e. ‘whenever there is an event of agreeing on this issue, it involves many people’. For example, in a situation where one third of the population is in agreement within that situation there is a sub-situation (or subevent) involving the agreement of the third group of agreement that does not hold in number. Schein’s paraphrase would predict, incorrectly, that the sentence was false. Again, the truth conditions for (130) seem to hinge on maximal collections of people who agree, and not on just any collection of agers. Compare this with a sentence which is slightly more generic in feeling, such as (138):

(138) Many people agree on such matters.

Here, there does seem to be a reading which might be paraphrased ‘typically, when there is discussion of such matters, the maximal set of people who agree is many’. However, this is still not Schein’s paraphrase.

Furthermore, note that the adverb of quantification ever occurs in the (b) version of each of the examples (138)–(136). This is surely not a coincidence. In fact, in all these event dependent readings there is a generic feeling which might be traced to either the explicit generic adverb of quantification or an implicit counterpart. Schein assigns universal or existential quantification over events freely, as an “aspectual difference” between sentences, but this fails to address the relation between the presence of adverbs of quantification and the possibility of quantification over events.

Moreover, ever is a negative polarity item, which means, following Ladusaw (1979), that it must be under the scope of a downward entailing operator. Monotone decreasing few is such an operator, and it seems that in these examples it is the only such operator. Comparison of (139)–(136) with examples which differ only in the presence of many instead of few, shows that it is indeed few which licenses ever:

(139) * Many experts ever agree.
(140) * Many Democrats ever vote with the President.

(141) * Many good students are ever unprepared.

Hence, ever must have scope under that of few. In view of this scope relation, a paraphrase of (134b)) which would be more in line with a compositional interpretation of the examples would be 'the number of Democrats who ever (on some occasion) vote with the President is few'. Instead of wide scope universal quantification over events, we have narrow scope existential quantification. This does indeed seem to be one available reading for (134b), but it is not the only one. There also seems to be a reading closer to Schein's event-dependent paraphrases: 'on each occasion, the number of Democrats who vote with the President on that occasion is few'. But it remains to be seen how this reading could be compositionally derived while respecting the negative polarity of ever.

Thus, while I agree that some of the examples Schein treats involve quantification over something like events (perhaps situations along the lines of Kratzer (1989)), his proposal seems both too general, since it predicts event dependent readings where they do not seem to occur, and ill-grounded in a general analysis of adverbs of quantification or the specifics of particular adverbs such as ever.

In summary, neither quantification over i-sums in the denotation of the CN nor Schein's approach in terms of quantification over events appears to provide a general account of examples such as (127). But we may avoid Kratzer's proportion problem for such examples by means of the treatment of few and many suggested in (129) and (132). I believe that these logical forms may themselves ultimately yield to a deeper explanation, which may very well involve quantification over situations, but this will have to await a better understanding of the nature of situations, and how they are individuated and denoted.

Having addressed the proportion problem, let us return to the general question of plural quantification. There are some cases involving few which seem amenable to treatment with simple plural quantification over i-sums in the denotation of the CN:

(142) Few twin girls have the same taste in clothing.

Notice, however, that (142) involves the type of CN which may only denote an individual with exactly two i-parts. The proportion problem does not arise because in the extension of twin girls none of these dual individuals may combine to form more inclusive i-sums; such combinations would no longer be twins. Also, note that we tend to think of these individuals as units — a unit of twins. This is also the case with the CNs in our original examples, units of competing companies or of two competing companies in (121) and (122), of couples or trios in (123). What

3.3. Determiners and distributivity

In the preceding sections, I have frequently assumed a classification of NPs into two types, according to properties of their determiners, the individual (or group) denoting NPs and the distributive, or quantificational NPs. Here I will address the question of the classification of determiners into the two classes.

Kamp (1981) and Heim (1982) both claim that there are two kinds of NPs, those, such as singular indefinites, which are interpreted as variables and those, such as universally quantified NPs, whose determiner sets up a relationship between the denotation of the CN and that of the predicate of which the NP is subject. The first type of NP I will call an individual denoting NP. The group denoting NPs are a subset of the individual denoting NPs, those whose denotations include non-atomic elements of the lattice-structured domain. The second type of NP I will call the quantificational NPs. These are the NPs which introduce a distributive element into the interpretation of a sentence in which they occur.

By the criteria I will discuss below, the individual denoting NPs include proper names and pronouns, as well as those with indefinite or definite determiners. Among the indefinite determiners are a, singular and plural some, and the numerals. The definite include singular and plural the and the demonstratives (this and that, these and those). There are a number of modified versions of these which are also individual denoting, such as a definite or indefinite determiner followed by a numeral, few, or many. I will not attempt to argue for the status of
bare plural NPs, since this would involve considerations beyond the scope of this work.\footnote{See G. Carlson (1977,1986), Wilkinson (1986) for discussion.}

Quantificational determiners include the universals each and every, both singular and plural no, and the plurals most, few, many, both, and neither. There are a number of other quantificational determiners, simple or complex, which I have not considered here. See Bennett (1974) for an extensive list. As in Bennett's treatment, most of these can be handled as synonyms of the determiners I treat here.

The proposed taxonomy is summarized in the following chart:

(144) \textbf{CLASSIFICATION OF DETERMINERS}

<table>
<thead>
<tr>
<th>Individual Denoting</th>
<th>Quantificational</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>each</td>
</tr>
<tr>
<td>some$_{sg/pl}$</td>
<td>every</td>
</tr>
<tr>
<td>1,2,3 \ldots</td>
<td>no$_{sg/pl}$</td>
</tr>
<tr>
<td>the$_{sg/pl}$</td>
<td>most</td>
</tr>
<tr>
<td>this,that</td>
<td>few</td>
</tr>
<tr>
<td>these,those</td>
<td>many</td>
</tr>
<tr>
<td></td>
<td>both</td>
</tr>
<tr>
<td></td>
<td>neither</td>
</tr>
</tbody>
</table>

Implicit in this taxonomy is the hypothesis that determiners are unambiguous. There are a few cases, notably the numerals and many, where it is not yet clear if this hypothesis can be maintained. More investigation into the behavior of these determiners will be required before this matter can be settled. In the meantime, I have adopted the stronger hypothesis in the interest of making clear predictions about the nature of distributivity.

One problem which complicates the issue here to some extent is that many of the lexical items frequently referred to as determiners (e.g. in Bennett (1974), Barwise & Cooper (1981), Scha (1984), among others) may actually be adjectives. For example, Hoeskema (1983) argues that the numerals are adjectives, and Link (1986) argues that they may in some cases be postmodifiers of determiners. Besides the numerals, many and few display adjectival properties: they take degree modifiers such as as and so, (compare with almost and nearly before numerals) and frequently occur in superlatives and comparatives, as predicate adjectives or after a definite or demonstrative determiner. Most also occurs after definite, though it doesn't display the other adjectival characteristics. Thus, the distribution of these possibly adjectival elements is in contrast to that of the true determiners, whether quantificational or group, since the latter occur only in SPEC position. Semantically, all but the numerals are non-logical, that is their definition may vary from model to model, and their interpretation is often context dependent, like scalar adjectives such as big or small. The true determiners, on the other hand, are logical, not varying from model to model. In any case, I am not convinced that the group vs. distributive character of these lexical items is contingent on their categorical status, and in what follows, I will assume they are determiners for the sake of simplicity.

Other theories have recognized a split into two types of determiners, but the split is not made in the same way as I propose. For example, May (1977,1985), along with other theorists in the Government and Binding framework, assumes that there are two kinds of NPs, the referential and the quantificational, only the latter undergoing Quantifier Raising at LF. However, he would classify a number of NPs as quantificational which Heim and Kamp (and I) consider individual denoting. For example, some CN is quantificational in his theory. His criteria for what is quantificational are unclear; in May (1985), he claims that even plural pronouns may at times be quantificational.

As we saw in Section 3.1.3, Scha (1981) also has two classes of NPs, the collective and the distributive. But again, his taxonomy, repeated here, does not coincide with mine:

(41) \textbf{SCHA'S TAXONOMY OF DETERMINERS}

<table>
<thead>
<tr>
<th>Distributive</th>
<th>Collective</th>
</tr>
</thead>
<tbody>
<tr>
<td>each</td>
<td>$\phi$</td>
</tr>
<tr>
<td>every</td>
<td>all</td>
</tr>
<tr>
<td>a</td>
<td>some$_{pl}$</td>
</tr>
<tr>
<td>both</td>
<td>no$_{pl}$</td>
</tr>
<tr>
<td>$\phi$</td>
<td>2,3,4 \ldots</td>
</tr>
<tr>
<td>all</td>
<td>the$_{pl}$</td>
</tr>
<tr>
<td>some$_{sing/pl}$</td>
<td>2,3,4 \ldots</td>
</tr>
<tr>
<td>no$_{sing/pl}$</td>
<td></td>
</tr>
<tr>
<td>the$_{sing}$</td>
<td></td>
</tr>
</tbody>
</table>

For example, the term collective suggests plurality, so only determiners which take plural CNs are included in this class. Singular indefinites and definites are considered unambiguously distributive.

Since my treatment of the singular determiners has been argued elsewhere (see, especially, Heim (1982)), I will focus on the classification of the plural determiners. The general rule of thumb in classifying plural determiners as group denoting or
quantificational is that group denoting NPs can appear to have a distributive reading, by virtue of adverbial distributivity (see Section 3.4 below), whereas plural NPs with quantificational determiners do not have group readings. It thus appears that the crucial factor in classifying a particular determiner is whether or not it may have a group reading.

There are a number of other tests for group readings. One of the most important is whether the NP formed by the determiner may serve as an antecedent for discourse anaphora. Quantificational NPs may only bind anaphoric elements within their scope. As we saw in Chapter 1, that scope may in some cases seem to be extended by teleopting; but strictly speaking, an NP only has scope over material in a predicate (syntactic or formed by abstraction) with which it combines. I will discuss this in detail in Chapter 4. For now, it suffices to note that this scope restriction is the basis of the frequently made claim that the scope of an NP is restricted to the sentence in which it occurs.

Individual denoting NPs, in contrast, are treated as variables in the Kamp/Heim theory of discourse, and, if they are not under the scope of another operator, they are bound by an existential operator over the entire discourse in which they occur. Suppose an individual denoting NP maps onto a discourse referent \( x_i \), and later the discourse referent for a pronoun is equated with \( x_i \) as we have seen in numerous examples above. Then both discourse referents will be bound by the same existential operator over the discourse in which they occur, and, since the assignment function will recognize them as the same variable, the NPs themselves will be coreferential. This accounts for the anaphoric potential of individual denoting NPs in discourse, in contrast to that of the quantificational NPs.

Above we have seen a number of examples of the availability of conjoined proper names, definites, and numeral NPs to serve as discourse antecedents. And we have seen that in general NPs with singular universal determiners may not license discourse anaphora. There are, however, complications in demonstrating the unaccessibility of plural quantificational NPs to serve as discourse antecedents. Consider:

\[
\text{(145)} (a) \quad \text{Many men lifted a piano.} \\
\text{(b) They got a crick in their back later.}
\]

Although many was classified above as a quantificational determiner, here it seems to serve as discourse antecedent for they in (145b). However, I believe that many men is not the antecedent of they. Rather, the search for an antecedent for they triggers the accommodation of a plural discourse referent 'the men who lifted a piano', an accommodation licensed by the existential implications of many CN.

As evidence for this, notice that there is no group reading of (145a). Even if the indefinite a piano has wide scope, the truth conditions of the sentence require that each man has the property of having lifted the piano. (Compare many men got together to lift a piano, where there may only have been a group lifting.) If many men were individual denoting, then such a group reading should exist, as it does in (1) and (2):

\[
\text{(1) Four men lifted a piano.} \\
\text{(2) Bill, Pete, Hank and Dan lifted a piano.}
\]

Thus, I take it that many is unambiguously distributive, but its existential implications often lead to accommodation of a group-denoting discourse referent which may serve as antecedent for a plural pronoun in subsequent discourse. See Chapter 5 for some discussion of the phenomenon of accommodation in plural anaphora.

There is another test for whether an NP is group denoting which is also based on the differential anaphoric potential of the two types of NPs. This is the NP's behavior in c-command anaphoric constructions. Recall that in Chapter 2 I agreed with Reinhart that the sloppy reading of such examples is due to c-command anaphora, and argued that the nonsloppy reading arises when the object in the first conjunct is discourse bound to its subject. Consider:

\[
\text{(146) Many girls love their mother, and Bob does too.}
\]

(146) is ambiguous between the sloppy reading, where Bob loves his own mother, and the nonsloppy reading, where he loves the mother of Mary, Susan and Kathy. But (146) has only the sloppy reading. These examples seem to support the classification of many as quantificational. However, I find the nonsloppy reading harder to get with an indefinite subject in the first conjunct:

\[
\text{(147)} (a) \quad \text{Some girls like their teacher, and Bernie does too.} \\
\text{(b) Four girls like their teacher, and Bernie does too.}
\]

I believe I can get a nonsloppy reading of (147a) if some is unstressed.\(^{38}\) But it

\(^{38}\)In fact, in general, some CN seems more likely to have a distributed reading when the determiner is stressed. This may be due to the fact that a contrast is being made between two classes of entities in the denotation of CN, on the basis of whether or not they have some property. Also, stressed some in particular has a scalar implication not every. Since every is purely quantificational, the implicature may introduce distributivity by association.
Chapter 3. Distributivity

is harder with (147b), and I am not sure of the source of the problem. Note that (147b) may occur with an unambiguously group reading:

(148) Four girls like their teacher. She encourages them.

If *four girls* were quantificational and bound *their*, then *their teacher*, which is referentially dependent on *their* and hence on *four girls*, would not be accessible as a discourse antecedent. Since it is in (148), we can conclude that the first sentence has a group reading.

Frey & Kamp (1986) note that in general indefinites are more likely to have distributed readings than definites. This seems to be true, and, whatever the explanation, it may be the reason why (147b) does not seem to have a nonsloppy reading. But generally, this test reinforces the results of the anaphoric potential test and the others below, so that in most cases all and only those examples with individual denoting NPs may have the nonsloppy reading.

Another test, discussed in Section 3.3.1, is whether an NP occurs felicitously as the subject of a predicate with a floated quantifier:39

(149) (a) The students all left.  
(b) # Few students all left.

As discussed there, it is not clear that examples such as (149b) are ungrammatical — the contrast in (150) is not one of grammaticality:

(150) (a) The students were all wet.  
(b) Few students were all wet.

In (150), only in the (a) sentence is ambiguous, meaning either that each student was wet through and through or that each student was wet. But (b) is unambiguous, meaning only that each of a small number of students was wet through and through. The unacceptability of (149b), then, seems to arise from the lack of a plausible 'through and through' meaning addition to the verb to leave.

Bennett (1974) points out that the prepositions among and between take a

---

This is related to Dowty & Brodie's (1984) pair in (i), showing that control structures display the same properties:

(i) (a) The students appeared to have all left.  
(b) # Many students appeared to have all left.

---

13. Determiners and Distributivity

group denoting complement:

(151) (a) Ellen found a thistle among some roses.  
(b) # Ellen found a thistle among few roses.

(152) (a) Jonathon found a poem stuck between (the) pages of logical formulae.  
(b) # Jonathon found a poem between both pages of logical formulae.

Finally, we may test the behavior of an NP as one conjunct of a coordinate NP, to see whether the conjoined whole may have the nonintersective reading which only group NPs license. (96), repeated here, has only an intersective reading, whereas (153) may have a group reading:

(96) Both apartments and most of the mobile homes have a fire extinguisher in the kitchen.

(153) Two movers and some neighbors carried the piano into the house.

Another kind of test which has frequently been used in the literature to gauge whether an NP is quantificational is the possibility of combining it with a so-called group predicate, such as gather, be numerous, and the like. In particular, the possibility of predicating gather of many CN, as in (154) has been taken as further evidence that that determiner may be ambiguous between an individual denoting and a quantificational reading:

(154) Many people gathered in the square to protest.

This kind of test is questionable. First, with respect to many and few, I argued in Section 3.2.4 that there is a predicative reading of these lexical items which would give (154) the meaning roughly 'the cardinality of the maximal group of people who gathered in the square to protest is many'. This restricted version of Link's plural quantification thus accounts for (154) without the assumption that the subject is group denoting. Note further, that the group predicates do occur with quantificational NPs, where the CN denotes groups, as in (155) and (156):

(155) (a) All the different species of insects were numerous that summer.
Most church congregations gather to worship on Sundays. (156)

(155) means that each species was numerous, (156) that the individual congregations in question gather to worship. If such group nouns as congregation denote atomic elements in the domain, then the subject of gather need not be a nonatomic i-sum. It is our world knowledge about what it is to be a congregation and what it is to gather which tell us that (156) is plausible, while John gathered is not.

3.4 Adverbial distributivity

I have claimed that the distributivity in examples with a group denoting subject arises due to an adverbial distributivity operator. In this section I will explore the character of adverbial distributivity in a variety of constructions. First, I will consider examples which contain explicit "floated" quantifiers, which Dowty & Brodie (1984) have analyzed as adverbial. This includes examples such as (157), with adverbial each:

(157) The men each lifted a piano.

As Dowty & Brodie point out, such examples seem to be related to partitive constructions, and I will briefly explore this relationship. In both of these constructions some explicit element serves to assure a distributive reading of a group denoting NP.

Then I will turn to consider related examples such as (56) with a definite plural subject and a distributive reading but no explicit adverb:

(56) The women from Boxborough brought a salad.

I will discuss the motivation for positing an implicit adverbial in such cases, I will also consider examples involving what Choe (1985) calls "antiquantifiers", and show how these may be related to adverbial distributivity. Finally, I will examine a type of example which Rooth (1986a) calls "numeral based donkey sentences", as in (158):

(158) (a) Seven fathers with two children send them to Montessori School.
(b) They think it's a good investment.

These sentences display the same structure and anaphoric relation between a complement in the subject and a pronoun outside its scope as the classic universal donkey sentences of Geach (1962), but here the subject is not quantificational, and, further, may itself serve as a discourse antecedent, as in (158b). I will argue that these are also cases involving implicit adverbial distributivity.

3.4.1 "Floated" quantifiers

Dowty & Brodie (1984) discuss "The Semantics of "Floated" Quantifiers in a Transformationless Grammar". These involve examples such as (159a) and (160a), which in earlier accounts were derived transformationally from the (b) examples:

(159) (a) The students each left.
(b) Each of the students left.

(160) (a) The students all left.
(b) All of the students left.

They focus their attention on the examples with all, and develop an account in which these quantifiers are base-generated adverbials. The synonymy between (160a) and (160b) is explained by the relationship between the semantics of constructions with adverbial quantifiers and the semantics of partitive constructions.

In this theory, VP denotations are third-order, where a "third-order" VP denotes a function from generalized quantifiers to truth values (i.e. of type \(((e, t), t))\), instead of a function from individuals to truth values (type \((e, t))\). That is, where in Montague (1973) NPs are functions which take VPs as arguments, here the VPs are the functions, the NPs arguments. This is necessitated in their account by the logical translation of all as a function from third-order VPs to third-order VPs:

---

40See Dougherty (1968, 1969, 1970) for the earliest arguments I am aware of that floated quantifiers are moved from quantified subjects; he posits the transformations of Quantifier Postposition and Quantifier Movement. Chomsky (1971) uses Dougherty's analysis to argue for interpretation from Surface Structure, instead of Deep Structure; Partee (1971) argues against Dougherty and Chomsky in her Section 3.1.1 on 'each-hopping'. I am not sure of the origin of the term 'floated quantifiers'.

41Dowty & Brodie do not directly address examples like (160b); however, they do discuss the partitive construction, and I believe that the comparison I am making here is implicit in their theory.

42Montague (1970), Keenan & Faltz (1978), Bach (1980) and Bach & Partee (1980) also treat VPs as functions which take NPs as their arguments, each offering independent reasons for this move.
(161)  \[ \{\text{all VP}_\text{p}\} = \{p \in \text{D}_{\text{VP}} \cap p \subseteq \{y \mid y \notin \text{D}_{\text{VP}}\}\} \]

(p.76, where \(\cap p\) = the intersection of all the sets in \(p\); \(\text{"a"} = \{X \mid p(X)\} \) (i.e. the maximal filter generated by \(y\); and \(\text{D}_{\text{VP}}\) = the domain of NP-denotations). What this says is that every individual that is a member of the intersection of all the sets in the NP-denotation has the property denoted by the VP. What it means intuitively (and loosely) is that each individual in the extension of the subject has the property denoted by the VP. Dowty and Brodie argue that it isn't possible to have VP's be of the usual type, \((e,t)\), and make \(a\) a function from type \((e,t)\) to type \((\{e,t\}, t, t)\), because of the interaction of all with verbal auxiliaries, including modals. Depending on their order, \(a\) and a modal auxiliary may have different relative scopes. If these elements are to be of uniform type, surely desirable on the grounds of uniform semantic and syntactic contribution, then VPs with and without \(a\) should be of the same type.

As we saw, Link (1986) adopts Dowty & Brodie's general view of such adverbs and incorporates their insight about their semantic contribution without lifting the type of VPs. He uses his \(D\) operator as the logical translation of \(a\):

\[ \text{D}_{\text{VP}} := \lambda x y (\text{atomic-i-part-of}(y, x) \rightarrow \text{VP}(y)) \]

Because plural individuals are of the same type as atomic individuals, he doesn't define \(D\) in terms of set inclusion, but uses instead the two-place i-part relation between individuals of the same type.

Dowty & Brodie also discuss a restriction on the type of NP which can serve as subject in such a construction. They note that this restriction parallels that on the complement NP in partitive constructions, such as the books in each of the boxes. Jackendoff (1972) called this the "Partitive Constraint", claiming that it requires such NPs to be definite. Barwise & Cooper (1981) redefine the partitive constraint as a requirement that a partitive complement be interpreted as a principal filter (this is what constitutes definiteness, in their view). However, since NPs with any of the universal quantifiers are principal filters but are not acceptable as complement NPs in partitives, Barwise & Cooper must add the stipulation that the principal filter be proper, that is, generated by a nonempty set in all worlds. Although this eliminates the universal quantifiers, it also weakens the intuitive value of their version of the constraint.

Ladusaw (1982) notes that there is one remaining counterexample to Barwise & Cooper's constraint: both CN is always a principal filter where defined, and yet such NPs are clearly unacceptable as partitive complements. Ladusaw then provides a characterization of the constraint which rules out both the universal quantifiers and both. (In what follows, I am abstracting away from Ladusaw's particular formulation.) He observes that if we adopt the idea that a group may be semantically an individual, then, roughly, group NPs denote individuals, or, more precisely, they denote principal filters generated by an individual. Then we can state the partitive constraint as a requirement that the complement NP denote such a filter. Since both and the universal quantifiers are unambiguously distributive, never group denoting, Ladusaw assumes that their denotations are not generated by an individual, and hence that they are ruled out.43

It is appealing to claim, as Dowty & Brodie do, that this generalization applies to the subjects of floated quantifiers as well: they must denote a principal filter generated by an individual, where groups are individuals in the model. There are, however, a couple of ways in which the requirements on partitive complements differ from those on the subjects of predicates with floated quantifiers.

One way in which they differ is that in general definiteness, in the sense adopted by Heim (1982) and Kaye (1982), seems to be a requirement on partitives, but not on the subjects of predicates with floated quantifiers. Compare the unacceptable partitives in (162) with the floated quantifier constructions in (163):

(162) (a) # all of some men
(b) # most of four girls
(c) # few of sixty diplomats

43J. Hoeksema (1985) presents an interesting compositional analysis of the syntax of partitives in which he claims that they are headless NPs with prepositional complements. Among other things, he claims that they are related to comparative and superlative NPs of the form exhibited in (i) and (ii), where the complement NP may contain all:

(i) the most beautiful of all
(ii) the best of all pupils

I have no idea at present of how to analyze these, or what they say about the analysis of partitives Ladusaw proposes.

44In some cases, the requirement that partitive complements be definite in this sense seems too strong. Ladusaw notes the existence of examples like the following, with "specific" indefinites:

(i) one of three people
(ii) one of several students who arrived late

However, though such examples cast doubt on the generalization that partitive complements must be definite, the issue of "specific" indefinites is a much broader problem, and its resolution may make it possible to retain such generalizations about definiteness, or recast them in a form which accommodates these examples.

43
Some men each lifted a piano.

Four girls each presented a science project.

Sixty diplomats all presented the President with letters from their heads of state.

(162a–c) are unacceptable due to the indefinite nature of their complements, but these same indefinites are acceptable in (163a–c).

Besides the definiteness requirement, it seems that conjoined NPs are not generally acceptable as the complement NP in a partitive, although they are acceptable in cases involving floated quantifiers:

(164a) # All of John, Mary and Susan ate pizza.

(b) John, Mary and Susan all ate pizza.

(165a) # Each of the bicycle, the tool kit, and the oven sold for $10.

(b) The bicycle, the tool kit, and the toaster oven each sold for $10.

Summarizing the discussion of partitives so far, we may conclude that they require their complement NP to denote an individual, in the extended sense of Link where we have plural individuals as well as singular, and also that they require their complement to be definite in the sense which concerns Heim (1982) and Kadmon (1986a,b). However, these requirements do not suffice to rule out conjoined NPs, and thus it seems that some further constraint remains to be discovered. In contrast, subjects in floated quantifier constructions do not obey the definiteness restriction and they permit conjoined NPs.

Ladasaw also notes the existence of examples with singular complement NPs, such as the following:

(166) some of the book

(167) most of the oatmeal

First note that these are all definite, individual denoting NPs, and thus obey Ladasaw’s partitive constraint—they each denote the principal filter generated by an individual. But here they are singular. These mass partitives demonstrate an unexpected bonus of couching Ladasaw’s insight in the context of a theory such as Link’s. Since the class of individuals in Link’s structured domain includes not only objects, but also individual portions of matter, we can make sense of mass partitives by giving the complement NP its marked, but readily available, mass interpretation. Link (1983) argues at length for the existence of a mass denotation of CNs and NPs which are generally count, through examples such as:

(170) There is apple in this salad.

(171) The apple in this salad is mealy.

where the portions of apple in (171) need not come from or constitute one piece of fruit. In the mass partitives above, the book may be viewed as its contents, the oatmeal as a mass, me as my physical mass or, more abstractly, the mass of my consciousness.

Recall that the elements of the atomic subdomain D of individual portions of matter are also ordered by a relation of material part/whole. The lattice which is formed by this relationship is nonatomic, but in other respects it is like the lattice on the entire domain E. Suppose that in (167) we consider the material correlate of the oatmeal, say the individual portion of matter a_m in the model. We then find a way of characterizing a level of homogeneous i-parts of a_m. For example, in most of the oatmeal got wet, we might consider a partition of the mass of oatmeal into equal portions of matter. These will also be atomic individual portions of matter on the count lattice which are related to the entire mass of the oatmeal by the material part relation on the nonatomic mass lattice. Then we compare the proportion of those portions of oatmeal which are wet to those which are dry. Again, as we saw with the plural quantification examples, the quantification here is not over all parts of the denotation of the oatmeal (in fact, in a nonatomic lattice, there may be infinitely many parts of a given individual), but over parts homogeneously characterized in some appropriate fashion.
CHAPTER 3. DISTRIBUTIVITY

Note also that there are floated quantifier constructions which are counterparts to the mass partitives:

(172) John was all tired out.

(173) The dog was all wet.

Again, if we partition the mass correlate of the dog (or its surface) into relatively homogeneous portions, and then determine whether each is wet, we seem to get the right truth conditions for (173).

These mass partitives and floated quantifier constructions support Ladusaw's characterization of the constraint on these constructions in terms of groups as individuals, especially in a framework such as Link's which has a built-in relationship between individuals in the count and mass domains. The individual denoted by a partitive complement or by the subject of a predicate with adverbal all need not be nonatomic. In the cases examined above where the individual is atomic on the count domain, we simply shift to its counterpart in the mass domain, and the distributivity over parts of the individual proceeds in analogous fashion in the two domains.

What does the partitive-like constraint on the subjects of adverbially distributive predicates amount to? Is it a formal prohibition of non-individual denoting subjects in such cases? Dowty & Brodie argue that there is no syntactic or semantic prohibition against combining a quantificational subject with a predicate modified by a floated quantifier. For example, they claim that the use of adverbial all with the subject all (of the) CN, as in (175) below, is felicitous due to "the repetition of a homophonic word which contributes nothing new to the meaning of the sentence" (p.81). One reason they argue for the lack of a formal prohibition of such combinations is that they get better with length. Compare their examples:

(174) ? All (of) the students have all left.

(175) All of the students in the phonology class that I taught at the 1973 Linguistic Institute have all gone on to become well-known linguists.

(175) does seem to be more acceptable than (174), but this may be due to the "forgetful speaker" effect — in other words, we may accept (175) as a case where by the time she got to the predicate, the speaker forgot that the subject was a partitive, instead of just a definite, and repeated the distributive operator. Contrast (175) with (176):

(176) # Most of the students in the phonology class that I taught at the 1973 Linguistic Institute have all gone on to become well-known linguists.

I'm not sure what (176) could mean. It seems to me to make conflicting claims about the proportion of the speaker's students who are successful, given that the use of most generally implicates not all. The illformedness of (176) is in contrast to (177):

(177) Most of the people who came in from the storm were all wet.

I think we need to explain not only why cases such as (174) are unacceptable, but why (177), unlike (176), is acceptable.

Dowty (1986) offers an discussion of the nature of all, which sheds light on this problem. He points out that it is difficult to ascertain whether this determiner is distributive or collective. (Recall that for Scha (1981) it was ambiguous.) Examples such as Dowty's (178), with collective predicates, seem to argue that all is collective and that it only has what Dowty & Brodie (1984) called a "maximizing effect", so that the CN and all the CN are truth conditionally equivalent:

(178) (a) John, Mary, and Bill are all alike.

(b) All (the) students gathered in the hall.

But (179), due to Bill Ladusaw, and (180), due to Barbara Partee, show that all can change truth conditions:

(179) (a) The students voted to accept the proposals.

(b) All the students voted to accept the proposals.

(180) (a) The trees are denser in the middle of the forest.

(b) All the trees are denser in the middle of the forest.

Thus, all the CN does not seem synonymous with group denoting the CN. And Dowty's (181b), with another collective predicate, is unacceptable, in contrast to the version in (a) without all:
CHAPTER 3. DISTRIBUTIVITY

34. ADVERBIAL DISTRIBUTIVITY

(181) (a) The students are numerous.
(b) # All the students are numerous.

As a point of departure in explaining these facts, Dowty argues that although the so-called collective predicates, such as gather, can only be true of a group qua group, many also have entailments about the individual members of their group subjects. He calls these the distributive subentailments of such verbs. For example, in order for a group to gather, most of its members must undergo a change in location, just as when a group disperses all must be in some central location to begin with and then leave; in order to disagree, the members of a group must each have opinions on some issue, etc. There are some predicates, he notes, which do not seem to have such subentailments, verbs like numerous which refer only to some characteristic of the group as an individual, but these are just the predicates which cannot take a subject with the determiner all, as in (180), unless, of course, the subject is a group of groups, as in (182), and the collective predicate numerous is distributed over each of the member groups:

(182) (a) The different species of insects were all numerous that summer.
(b) All the different species of insects were numerous that summer.

Given this fact about collective predicates, Dowty then characterizes the effect of all with such predicates, as follows:

Dowty’s Hypothesis: the effect of all on a collective predicate is to fully distribute the predicate’s sub-entailments to every member of the group argument: Instead of merely holding of some (proper) subset of these members, as required by the predicate by itself, all requires that these sub-entailments hold of every member of the group.

To see how this works, consider again Ladusaw’s (179). Dowty notes that in order for a group to vote for a proposal, it is necessary that a certain percentage of the members of the group have the property of voting for the proposal. In general, it isn’t necessary that all members have the property, but when all modifies the subject in (b), then the subentailment must be distributed over the entire group; each member must have the property. This explains the different truth conditions of (179a) and (b).

Just as many “collective” predicates appear in this light to have a distributive aspect, Dowty points out that often the “distributive” predicates have a reading in which they do not distribute over all the members of a group denoted by their subject. He offers the following examples:

- At the end of the press conference, the reporters asked the President questions.
- What was that noise?
- Oh, I’m sure it was only the children getting up to watch cartoons. Go back to sleep.

Dowty argues that these examples may be literally true even though not all the reporters asked questions or one of the children is still in bed. He points out that, as with G. Carlson’s battalion examples cited in (78) above, the number or percentage of the members of the group denoted by the subject which must have the property denoted by the predicate varies according to the sense of each. So, it may be that (183) is true although only a few of the reporters at the press conference, say one in four, got to ask questions, but if only one of four children is watching television, it seems false or inappropriate to say The children are watching tv.

These predicates, then, seem to have an extended sense in which they may be true of a group. But if we modify (183) with all, either on the subject or on the predicate, then the truth conditions are stronger, and the sentence is only true if each reporter asked questions. Thus, all also seems to strengthen the distributivity of these “distributive” predicates. Dowty suggests that if we regard the distributive entailments of ordinary distributive predicates as a special case of the distributive sub-entailments mentioned in (182), then the hypothesis may be generalized to hold over all predicates, and not just the collective ones.

Dowty does not provide an explanation of Partee’s (180), but I believe that in Link’s framework this may be viewed as an extension of the use of all to the mass domain. Suppose that be dense means something like ‘the material correlate of the subject has the mass property of density’. The trees is a group denoting NP. In this context, it may be taken to refer to the i-sum of all the trees in the forest. Then, the (a) example means that ‘the mass of the trees in the forest is denser in the middle’. But all in (b) distributes this mass property over the atomic i-parts of the group denoting subject, that is, it requires that each individual tree be such that its material correlate has this mass property. So the truth conditions of the (b) example are considerably different from those of (a): ‘the mass of each tree in the forest is denser in the middle’.

Summarizing Dowty’s account of all, it seems to distribute the lexical subentailments of a verb over the members of a group denoting subject. Where the

(184) A: What was that noise?
B: Oh, I’m sure it was only the children getting up to watch cartoons. Go back to sleep.

The boys in the eleventh grade are cheating.

where it is not necessary for all the boys to be cheating in order for the sentence to be true.

---

46 Frey & Kamp (1986) also point out examples of this phenomenon, such as (i):
verb is itself basically distributive (i.e., involving activities or states which we generally attribute in the strict sense only to individual agents with a single will and consciousness) or a mass predicate, then all "maximizes" the distributivity of the predicate, so that it holds of each atomic i-part of the subject. Dowty points out that there are difficult cases involving what he calls "cooperative group endeavors". These are illustrated by Link's (1983) (185):

(185) (a) The children built a raft.
(b) All the children built a raft.
(c) The children all built a raft.

If all were translated as the D operator, then we might expect that (185b) and (c) would mean only that each child built a raft, but this is not the only reading of these sentences. Link proposes that all be translated as a 'participates in' operator, so that the difference between (185a) and the examples in (b) and (c) is that, although the group of children as a unit may have accomplished the raft building in each case, in (b) and (c) each individual child participated in the building, while in (a) some may not have taken part actively. Dowty shows that this may be just a subcase of all distributing the lexical subentailments of such a verb over the members of the group. For example, there are a number of different tasks involved in building a raft. The predicate itself only requires that its subject accomplish all those tasks, but if all is involved, either as a determiner (or modifier of a determiner) or as an adverbial operator, then each of the children must satisfy some of the lexical subentailments of the verb. That is, each child must actively participate.

With this perspective, we may return to the question of whether, and when, floated quantifiers are acceptable with distributive subjects. Consider the unacceptable acceptability of floated quantifiers in examples such as (174). All (of) the students is not a group denoting NP; it has a distributive element. Thus, the sentence could only mean that each of the students has the property of 'having all left'. But what can it mean for a single individual to have 'all left'? This is quite strange, for it implies that the individual could have partly left. Similarly, (176) should mean that each of a large number of the students has the property of all being successful, but this is quite an odd property to predicate of an atomic individual.

By contrast, in (177) we have a mass predicate, wet, and it seems quite reasonable to say of a number of people who came in from the storm that they each have the property of being all wet. When used in conjunction with mass predicates, all seems to be able to split atoms, by shifting to consideration of homogeneous material parts of the material correlate of the subject. This is clearly related to its use in the mass participles. These facts about all suggest that it is not adequately translated by Link's D operator, (83) above, which requires distributivity of the full sense of the predicate over all and only the atomic i-parts of an individual

denoted by its subject. All is flexibly distributive: according to context, it may distribute either the whole sense of the predicate or its subentailments, and it may also select the level of individuation of individuals over which it distributes such properties, either atomic or subatomic (mass parts).

Floated each and all are different in this respect, and I believe these differences (which are also reflected in their use as determiners) may be captured by translating each as the D operator. First, notice that there are no mass participles with each, suggesting that it doesn't have the flexibility to distribute over mass parts, as does all:

(186) * John is each wet.

Now consider the following examples:

(187) (a) The children all built a raft.
(b) The children each built a raft.
(188) (a) Every group of children built a raft.
(b) Every group of children all built a raft.
(c) # Every group of children each built a raft.
(189) (a) That group of children all built a raft.
(b) # That group of children each built a raft.

The use of each in (187b), in contrast to the use of all in (a), seems to require that each child built a raft by herself. That is, each does not seem to be able to distribute subentailments in general, but only the full sense of the predicate. (187a) may also have the reading of (187b), so that it is ambiguous. In (188), all seems felicitous in (b), at least for some speakers, with a reading where each child in each group participated in the group project of building a raft. But (c) is unacceptable. And in (189), the (a) example may only mean that all the children participated in building a single raft, while the (b) example is felicitous.

If we assume that each does translate as the D operator, then we can explain these differences. The difference between (187a) and (187b) falls out automatically, since D distributes the full sense of the predicate over atomic i-part of the denotation of the subject. With respect to (188), since every is distributive, then (188c) means that any given group has the property of each building a raft. If each translates as the D operator, and if the denotation of group is nonatomic, then applying an adverbially distributive predicate to this subject would mean that each atomic i-part in each group, each child, built a raft. But (188c) doesn't
have such a reading (and in fact neither does (188b), although in other examples, such as (187a) all may lead to a reading which is synonymous with the use of D.

However, recall that I discussed above the possibility that a CN such as group may denote an atomic individual. This example supports this analysis of group nouns. Group seems to put a lower bound on the distributive potential of each, since none of the i-parts of a group is also a group. And it similarly prevents all (188b) from distributing the property of building a raft over each of the children in each group. Hence, even though the denotation of the subject in (188) is a group in the pretheoretic sense, distributivity over the individual members does not take place, as it does when the children or the children in the groups is substituted in the subject. The examples in (189), with a non-distributive subject, support this analysis of group as denoting atomic individuals, since the (a) example may only mean that all the children participated, and the (b) example is infelicitous.

Throughout this discussion of examples with nongroup denoting subjects and adverbial distributivity, I have used the symbol for infelicity, $\#$, instead of the star marking ungrammaticality. This is because, like Dowty & Brodie, I do not think the floated quantifier counterpart to the partitive constraint is a requirement on grammaticality. Examples (177) and (188a) are acceptable, though they have non-group denoting, distributive subjects. And we could understand the unacceptable of examples (174) and (176) as a consequence of the pragmatic oddness of the resulting translation.4 If I think the unacceptable of (190) is parallel to that of other examples where a predicate with floated each has a subject which denotes an atomic individual:

(190) $\#$ John each built a raft.

If we translate each as D, the truth conditions for (190) are simply that each atomic i-part of the subject has the property of building a raft. Now, John is the only atomic i-part of the denotation of John, so the sentence should mean the same as John built a raft. The function of adverbial each is to distribute in predicate over the i-parts of the denotation of the subject. It may be that it presupposes or implies that the subject has more than one i-part. Then, when a subject has none, the presupposition is unsatisfied, or the cancellation of the implicature violates Gricean cooperative maxims. In any case, the result were

4Angelika Kratzer (p.c.) points out that (i) seems unacceptable:

(i) $\#$ The committee all sing.

I am not sure why this is so. I find (ii) alright:

(ii) The committee all sang Christmas carols at the last meeting.

This suggests that the generic mood of (i) may be the source of the problem.

4.4 Adverbial Distributivity

In the odd, and not ungrammatical.

In summary, while each and all are both adverbial distributivity operators, they have slightly different properties, the former translatable in terms of Link's D operator, the latter able to distribute below the atomic level (as with the group examples) or to shift into the mass domain. And while I cannot comment here on the status of the partitive constraint, the parallel constraint on floated quantifiers does not appear to be a condition on grammaticality.

3.4.2 The D operator

I argued above that the source of the distributivity in examples such as (56) and (58) is an implicit adverbial operator, which we may translate as Link's D operator:

(56) The women from Buxborough brought a salad.

(58) John gave a pumpkin pie to two girls.

Let me review the evidence for this claim. Both the women and two girls are individual denoting: we know that they are not quantificational, or distributive, but because of the general character of the determiners the and two, discussed in Section 3.3, and, in these particular examples, because they both may serve as discourse antecedents. The properties of bringing a salad and of being an x such that John gave x a pumpkin pie are both mixed predicates in Link's sense, that is, they may be true of a group (a non-atomic i-sum) or of an individual (an atom). But if the sentences have a group reading, then a salad and a pumpkin pie are accessible to serve as discourse antecedents, while if the sentences have distributive readings these NPs are not accessible for discourse anaphora. We need to explain how, if these singular indefinites are not under the scope of a quantificational NP, they are masked with respect to anaphora.

My proposal is that they are under the scope of a quantificational operator, an implicit adverbial on the predicate. This will explain both the distributive reading of the sentence and the anaphoric masking, and, since the subjects (syntactic in (56), by abstraction in (58)) are not themselves under the scope of the adverbial operator, we can understand why they are accessible for discourse anaphora even though their sentences have a distributive reading.

There is another type of example which I believe involves the use of the D operator: Choe (1985) has noted the existence in Korean of what he calls an "anti-quantifier", the postnominal -sul, -sul generally occurs on NPs in the syntactic VP, though it may occasionally occur on the syntactic subject. It seems to require that some other NP distributes over a predicate in which it occurs (either the syntactic VP or one formed by lambda abstraction). This happens even when
the (syntactic or quantified in) subject is group denoting, such as a plural definite
description or plural pronoun. (I am not sure if the subject may be quantificational
itself.) Link (1987) discusses the German particle je, which seems to have similar
properties. He proposes that je triggers the application of the D operator on a
predicate in which it occurs. And there seem to be a number of other languages
which have such anti-quantifiers. For instance, David Peeters (p.c.) suggests
that Russian and Polish jo may be an anti-quantifier.

What about English? There is a use of each in ditransitive predicates which
differs from the floated quantifiers considered by Dowty & Brodie (1984). The
floating quantifiers occur generally after the first auxiliary of the predicate (though
they may also occur before a modal, indicating that they have wider scope). Dy-
transitive each is exemplified in (191) – (194):

(191) The students gave the guardmen a flower each.
(192) The students gave a flower each to the guardmen.
(193) The students gave the guardmen each a flower.
(194) * The students gave a flower to the guardmen each.

The occurrence of this each seems to be restricted to adjacency to the direct object
of a ditransitive verb. 144 45 The requirement of adjacency to the direct object is
illustrated by the unacceptability of (194). In (191) and (192), where each follows
the direct object, there are two readings available, one where each of the students
gave a flower to the guardmen, the other where the students as a group gave a
flower to each of the guardmen. (193) has only one reading, where each of the
guardmen received a flower from the students.

The generalisation seems to be that ditransitive each requires that the direct
object occur in a distributive predicate. If each follows the direct object, as in
(191) and (192), then either the syntactic predicate or one formed by abstraction

144 I take the direct object to be the NP which occurs immediately following the verb
when passive to marks the indirect object. This may not be an uncontroversial asser-
tion (see Bach (1979,1989b), Dowty(1983)). Alternatively, in these examples, we might
say that ditransitive each must follow the theme of give.

45 David Peeters (p.c.) also notes that this each appears to be unacceptable in
examples involving wh-movement:

(i) What kind of flower ( * each) did the students give the guardmen ( * each).

The indirect object may be distributed, in the first case over the syntactic
subject, in the second, over the indirect object. But if each precedes the direct
subject, as in (193), then it is only the predicate formed by abstraction on the
indirect object which is to be distributed, over the direct object.

My proposal is that, like Cho’s Korean antquantiier, ditransitive each trig-
gets the use of the adverbial D operator on some predicate in which it occurs. Its
moving position may put additional constraints on what predicate is distributed
and what argument. Now consider Partee’s (1985) – (197), which receive the type
of reading which was not available in (188), with floated each, a reading where
each man receives $5:

(195) That group of men received $5 each.
(196) Every group of men received $5 each.
(197) Every group of delegates had the same number of votes each.

Consider also:

(198) The group received $5 each.
(199) Four boys received $5 each.
(200) Alison gave the boys $5 each.
(201) Jim sold all the ginseng plants he had collected. He received $5 each.

Like ditransitive each in (191) and (192), this use of each follows the direct
object. And like ditransitive each, it seems to amount to the requirement that
a predicate containing the object should be distributed over some other range of
individuals. This range may be explicit, as the subject in (199) (which, on the
most likely reading, is equivalent to The boys each received $5), or it may be
inferred via a part-whole relation, as in Partee’s examples and (198), even where
the parts are part of an atomic individual such as a group. The predicate may be
the syntactic predicate, as in (185) – (199), or a predicate derived by abstraction,
such as (201) (given in (200)).

(201) is interesting because here the range of individuals over which the predi-
cate containing $5 is to distribute is given only in the context, by the ginseng.
plants of the preceding sentence. It is as if receive in this example requires an implicit, contextually given argument for its interpretation, present in the discourse representation, which is referentially dependent on the ginseng plants. Note that this same kind of reading is available for other examples, given a suitable context. For example, consider (198) after John, Joe, and Jose collected several ginseng plants. (198) may then mean that the group of boys received $5 for each of the plants, a reading which would not be available for the boys each received $5. So postnominal each seems to permit distributivity over the elements of a contextually given group, unlike auxiliary each, which is restricted to i-parts of the subject.

In these examples and in the examples with floated each the D operator appears to give the correct truth conditions. There are, however, a number of important unanswered questions about adverbial distributivity. Consider the range of quantificational determiners in English — there are quite a few of them, which differ considerably in meaning. Does adverbial quantification in English display universal force? There are a number of other examples which it is tempting to analyze as involving nonuniversal adverbial quantification, as in (202):

(202) The children in my class mostly have a computer at home.

The most readily available reading of this example is that most of the children have a computer. But, although mostly here doesn’t appear to quantify over times or situations, it often does so in other examples, and is regarded as one of Lewis’ (1975) adverbs of quantification. This raises the question of the relation of adverbial distributivity in general to adverbs of quantification.

Another open question is whether adverbial distributivity is generally atomic, as suggested in Link’s translation of D, or whether there are more cases which are like those with all in being flexible both in what property is being distributed, and in the level of individuation of parts of the subject.

In sum, though I believe the D operator is adequate to derive the proper truth conditions for a variety of examples, its character and the extent of its use need to be investigated further, as well as its relation to more general processes of adverbial modification.

### 3.4.3 Numeral based donkey sentences

Rooth (1986a) discusses examples which he calls “numeral based donkey sentences”.50 In these, the plural subject with a numeral determiner may serve as

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(203) (a)</td>
<td>Every father with two children sends them to Montessori school.</td>
</tr>
<tr>
<td>(b)</td>
<td># They both love it.</td>
</tr>
<tr>
<td>(c)</td>
<td># He thinks it's a good investment.</td>
</tr>
<tr>
<td>(143) (a)</td>
<td>Seven fathers with two children send them (both) to Montessori school.</td>
</tr>
<tr>
<td>(c)</td>
<td># They both love it.</td>
</tr>
<tr>
<td>(b)</td>
<td>They think it's a good investment.</td>
</tr>
</tbody>
</table>

The relevant reading of (143a) is the one where each father has two children and sends both of them to the school. Two children in (143a) binds the pronoun them in the same way as in (203a). And as with the narrow scope indefinites in the donkey sentence two children cannot license discourse anaphora, as shown in (143b) and (203b). But seven fathers with two children can serve as discourse antecedent for the pronoun they in (143c), while the subject of (204a) may not license the subject of (205c).

Rooth shows that a grammar which incorporates a variation on Heim’s (1982) file change semantics can provide an analysis of these examples. He points out that the essential idea of file change semantics is that a sentence meaning is a relation between information states (Heim’s ‘file change potential’). He then points out that what Barwise (1985) calls ‘dynamic interpretation’ may be viewed as another formal realization of this same idea, in an extensional fragment.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>Each of the farmers with a donkey beats it.</td>
</tr>
</tbody>
</table>

Whatever Rooth’s reasons for omitting them in his later manuscript, I think the numeral based donkey sentences have the relevant readings, as in (205), and illustrate an important fact about distributivity.

Rooth focuses on the following example:

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>Three researchers, with two microscopes, use them both.</td>
</tr>
</tbody>
</table>

While I agree that this example may have the reading indicated, I find that these examples improve when the relation between the head and the complement NP is more likely pragmatically to be distributed, and especially when the head is relational, as in (205). My informants concur. There is no difference here in structure or the potential for discourse anaphora.

---

50Rooth (1986a) is an earlier version of Rooth (1986b). The two manuscripts differ considerably; among other things, examples like (205) are only discussed in the earlier version; Rooth (1986b) discusses instead related examples with partitive subjects, such as (i):
Barwise’s proposal is different from Heim’s in three principal respects. He has a slightly different conception of the structure of files than that of Heim, called dynamic interpretation; he extends dynamic interpretation to constituents below the sentential level; and he introduces the use of parameterized sets. Heim views the extensional file change induced by a sentence as an ordered pair: the Domain, which is a set of numerals which corresponds to the numbers on the file cards required in processing the sentence, and the Satisfaction Set, a set of assignment functions such that for any assignment function, the conditions on the files whose indices are in the domain are true of the individuals denoted by the correspondingly indexed variables under that assignment function. A file change potential is a function from such ordered pairs, or files, to other ordered pairs, or files. Barwise’s treatment essentially eliminates the Domain, and makes the denotation of a sentence a function from partial assignment functions to partial assignment functions, showing, thus, that the Domain is inessential in an extensional fragment. Barwise also takes his ‘dynamic interpretation’, or file change, above the sentential level. That is, not only are the denotations of sentences functions from assignment functions to assignment functions, but smaller constituents involve such functions, as well. For example, a VP denotation is not simply a set of individuals, but a set of triples of the sort defined for the intransitive walk by Rooth’s (204), a relation between individuals and pairs of assignment functions:

\[ g, x, [\text{walk}]y \text{ if } x \in F(\text{walk}) \text{ and } y' = y \]

The input function, \( g \), and the output function, \( y' \), are the same because the predicate contains no indefinite NPs which might set up antecedents for NPs later in discourse. Any individuals \( x \) are required to be in the extension of \( \text{walk} \), \( F(\text{walk}) \). The denotation of an indefinite NP, on the other hand, has an output function which differs from the input function just in the value assigned to the variable which has the same index as the NP:

\[ g, x, [\text{a man}, n]y \text{ if } x \in F(\text{man}) \text{ and } y'^* = y'^* \]

Here, the output function \( y' \) may differ from the input function \( y \) in the value for the \( n \)-th variable, which is specified to be that of \( x \), an element of the extension of \( \text{man} \); this is the meaning of \( y'^* = y'^* \).

Finally, Barwise introduces the parameterized sets analysis of quantification. In generalized quantifier theory the denotation of a quantificational determiner such as every is taken to be a relation between the denotations of two predicates, that denoted by the CN and that denoted by the VP, i.e. a relation between two sets. But parameterized sets contain as elements ordered pairs of individuals and assignment functions. In universal donkey sentences such as (203), there is a subject internal NP, two children, whose value varies with that of the whole subject (a given father). The CN of the subject father with two children, denotes a parameterized set, where in each ordered pair the individual is a father and the \( x \)-th variable of the assignment function is an individual whose i-parts are two children. Recall that a VP, for example send them, is Montessori school, denotes a set of triples, where the second element is an individual with the relevant property. Since \( g \) is pronounal, and hence requires a prior antecedent, the input and output assignment functions will be the same for all the triples in the denotation of send them, as they were for walk. In general, a nonquantificational subject of such a predicate must denote an individual who is the second element in one of the triples in this set. The quantifier every in (203) is a relation between a parameterized set, that denoted by its CN, and another parameterized set, \( \{x, y : x \in F(\text{send them})\} \), i.e. a parameterized set drawn from the denotation of the predicate. The relation between the two parameterized sets is the subset relation, i.e. the parameterized CN denotation must be a subset of the parameterized set drawn from the VP denotation. This permits the binding of an object pronoun them, by two children, since the assignment functions in the ordered pairs of the CN denotation will now be the same as those in the VP.

Rooth develops a small fragment of English which incorporates the relational semantics and parameterized set constructions of Barwise (1985). He then compares this with a Montague grammar treatment of the same fragment, and shows that apart from the inability of the latter to treat donkey anaphora, there is a natural mapping between the two fragments. This gives insight into the way in which indefinite NPs may be interpreted as variables a la Heim (1982) and yet be equivalent to Montague’s quantificational indefinites.

Besides this theoretical result, Rooth claims that his fragment shows two applications of Barwise’s ideas. The first lies in the treatment of a problem with examples such as (206), a problem which both Partee (1983) pointed out and is discussed in Bauerle & Egli (1985); Rooth calls this ‘Farmer/Donkey Asymmetry’ and Kedmon (1986) calls it ‘the Proportion Problem’:

\[ g, x, [\text{a donkey} n]y \text{ if } x \in F(\text{a donkey}) \text{ and } y'^* = y'^* \]

Most farmers that own a donkey beat it.

Basically, people feel that (205) is false in situations where nine farmers own one donkey each and don’t beat it, but one farmer owns thirty donkeys and beats them all. Yet, since both Kamp (1981) and Heim (1982) quantify over farmer/donkey pairs without distinguishing the two elements of the pairs, they predict that (206) is true in such a situation.

Rooth (1985) proposes to handle this problem within the DR framework. She suggests that we consider not just all farmer/donkey pairs, or, technically, all embedding functions such that the discourse referent for the whole subject corre-
sponds to a farmer who owns a donkey which is the denotation of the discourse referent for a donkey; but rather that we consider equivalence classes of such embedding functions, where all embedding functions which pick the same farmer but different donkeys are in the same equivalence class. Before, we interpreted (206) as meaning that most of the embedding functions where there was a farmer and a donkey in the owning relation could be extended so that the farmer beat the donkey. But under Root's proposal, (206) means that most equivalence classes of embedding functions can be so extended. Now (206) won't be true in the problematic situation described above. Rooto points out that with the distinction between the farmers and the donkeys which is built into the parameterized set approach, one can incorporate Root's proposal very naturally.25

The other application which Rooto considers is more directly relevant to our main topic in this chapter, distributivity. He proposes to account for examples such as (143) by treating both the subject and the predicate as involving parameterized individuals, even though the subject is not quantifiational.

He does this by means of a rule which changes ordinary set-denoting CNs into parameterized set-denoting CNs, prior to pluralization (the latter along the lines of Link (1985)). Any rule which forms an NP out of a (not necessarily quantifiational) determiner and such a parameterized CN will end up denoting triples defined by a relation between a parameterized individual and a pair of assignment functions. A separate parameterization rule for VPs makes them denote triples, each of which is defined by a relation between a parameterized individual and a pair of assignment functions. The subject-quantifier combination rule then has the effect of permitting only pronouns in the VP to have access to the indirectly complement in the subject (two children in (143)), and not any subsequent pronouns in the discourse. The entire subject, however, since it is not inherently quantifiational, may induce a change in the output assignment functions, so that it may serve as a discourse antecedent:

(143) (a) [Seven fathers with two children]j, send them (both) to Montessori school.
(b) # They, both love it.
(b) They, think it's a good investment.

While this way of treating the numeral donkey sentences is descriptively adequate, it misses an important generalization. In all the examples where parameterization is appropriate, distributivity is involved. In quantified donkey sentences like (203), the distributivity is introduced by the quantifiational determiner only in that example. In the numeral based donkey sentences such as (143), the possibility of the subject serving as a discourse antecedent suggests that the NP

25See also Kadmon (1986) for extended discussion of the Proportion Problem in the context of Discourse Representation Theory.

3The acceptability of DPs and whether or not they are required in such examples seems to be language-specific, and even to vary from person to person within English.

3.5 Dependent plurals and global agreement

Dependent plurals (DPs) are exemplified in the following examples. (208) is from Chomsky (1975); (209) and (210) are from deMey (1981).
We may understand the speaker of (208) as claiming that each unicycle has one wheel, (209) may mean that each boy brought his own father along, and (210) may be taken to inform us that in each of some regularly spaced temporal intervals, a single train leaves for Amsterdam. There is a distributive sense in each case: yet wheels, their fathers and trains are syntactically plural. And in each case, there is another plural element, whether the explicit plural subjects unicycles or all the boys, or the adverb regularly, which seems to suggest a number of regular intervals. It is generally the case that DPs must be in the scope of such a plural element; hence the name.

In this section I will consider the consequences of the dependent plural phenomenon for a theory of plurality and distributivity. I will argue that we should distinguish whatever it is that licenses dependent plurals from their effect on interpretation. With respect to interpretation, I believe that they may be interpreted as regular plural NPs, and so have little truth conditional effect, though in some cases they may be pragmatically useful in avoiding misunderstanding. And although I cannot offer a theory of what licenses dependent plurals, I will present some data which suggests the appropriate line of approach. First, however, I want to review some earlier suggestions in the literature regarding dependent plurals. In particular, some authors have taken this phenomenon as evidence for what Barbara Partee (p.c.) calls a rule of global pluralization. For example, Chomsky (1975) uses (208) to argue that:

(208) Unicycles have wheels.

(209) All the boys have brought their fathers along.

(210) From here, trains leave regularly for Amsterdam.

...a principle of compositionality is suspect. Global properties of the sentence, which may be quite involved, seem to play a role. We cannot simply assign a meaning to the subject and a meaning to the predicate (or to a sentence form with a variable standing for the subject), and then combine the two. Rather, the meaning assigned to each phrase depends on the form of the phrase with which it is paired ...Plurality is, in some sense, a semantic property of the sentence ...
(212) (a) denotes:
\[ \text{[invited-to}(j, \text{ey parents-of}(y, j), \text{ey place-of}(v, j @ m)) & \text{[invited-to}(m, \text{ey parents-of}(y, m), \text{ey place-of}(v, j @ m))] \]
Hence, 'John invited his parents to their place and Mary invited her parents to their place'. We will return to a discussion of the treatment of analogs which this example implies in Chapter 5. Here, we will focus on the feasibility of a rule of pluralization such as this as a general approach to the problem of DPs.

Notice that what Link's pluralization does is make their an identity function on elements in the restricted domain of the distributivity induced by *, i.e. on the 1-part of the 1-sum Johnmary. Since the head parents is a function from elements denoted by the possessive NP to a male and a female person, this makes the whole NP their parents a function from Johnmary. Imagine the following extension of this treatment to a general rule of global pluralization: Suppose that in general DPs were functional, although not always identity functions, and that they contain an implicit variable for the argument of that function. To illustrate what I mean, consider the following derivation of The unicyles have wheels.

(213) (a) [the unicyles] : APP(ey unicyles(y))
(b) [have wheels]: Au[3w(week-of(x, u) & have(u, x))]
(c) VP pluralization applies to (b), yielding:
\[ *\{Au[3w(week-of(x, u) & have(u, x))] \}
(d) [the unicyles have wheels]:
\[ *\{Au[3w(week-of(x, u) & have(u, x))]\} \}
In (a), the unicyles denotes the s-sum of all unicyles; in (b) have a wheel denotes not just the property of having any old wheel, but of being an individual who has that individual's wheel. This may seem a bit redundant with have, but it isn't if wheel-of is a singular DistirP, and hence contains only atomic individuals in its extension. Now let us assume that when a predicate such as (b) is pluralized as in (c), this triggers syntactic pluralization of the affected NPs, those whose translation contains the bound variable. So a wheel becomes wheels, just as it became theirs in (212). Actually, if Link had wanted to derive the same kind of reading as (211) for the very similar John and Mary invited their mothers to their place, he would have had to percolate the plurality of the pronoun up to the level there, as well.

Chomsky seems to have had a different kind of account of DPs in mind, one which operated syntactically on VPs, affecting morphology only. Notice that if we are to entertain a role of global pluralization or agreement, it must be at least as sophisticated in its characterization of what it is to be a predicate as the one I have sketched, following Link. This is because DPs do not always "distribute" over the syntactic subject, as we see in (214) and (215).

(214) Jane gave the kids toys for Christmas.
(215) German civil defense workers spotted two of our planes.

Types in (214) can have a DP interpretation, where each kid received one toy. In (215) when two of our planes has wider scope than the subject we may understand its instance to mean that one worker spotted each plane. My extension of Link's approach, utilizing lambda-abstraction, can develop the appropriate abstracted predicate for such cases. Also, this approach belies Chomsky's claim that the DP phenomenon requires an uncompositional treatment. Since bare plural DPs (plural CNs with no determiner) are treated as involving an implicit bound variable, pluralizing them in the fashion described seems no less compositional than pronominal agreement.

However, even though such an approach might be initially appealing, it is ultimately inadequate to deal with the full range of DP data. Barbara Partee (p.c.) has noted examples like the following:

(216) Those men married wives who are similar.

Here, although we easily interpret the direct object as dependent on those men in the sense that we understand that each man has one wife, the relative clause contains a symmetrical group-denoting predicate, one which isn't true of atomic individuals in the extension of wife. We cannot provide a derivation of (216) along the lines of (213), since there is no singular DistirP wife which is similar, with only atomic individuals in its extension, \textsuperscript{23} Thus, there are at least some examples with DPs which would not be accounted for on the extended-Link approach.

Partee also points out that a rule of global pluralization which operates to pluralize the predicate of a plural subject cannot account for the examples of DPs

\textsuperscript{23} In her reply to Chomsky (1975), Partee (1976) notes the following example:

(0) The boys gave the girls nickels.

In this example, not only the bare plural nickels can receive a DP interpretation, but the distal NP the girls. Chomsky's inherent vs. pluralized reading dichotomy would predict only two readings of the predicate here, 'gave the girls nickels', the inherent reading, and 'gave one of the girls a nickel', the pluralized reading. But two other readings are possible, 'gave the girls nickels', and 'gave one of the girls nickels'. But the extended-Link approach would permit us to derive any of the readings available in (0).

\textsuperscript{23} There is, of course, a functional use of similar which could be treated as containing a discourse bound variable, as in (ib) and (ib):
with a non-NP antecedent, such as deMey's (210) above, and her (217):

(217) John often wears loud neckties.

In this instance, we may naturally assume that John wears these neckties one at a time.

DeMey (1981) offers an analysis of the DP phenomena which does not involve a global pluralization rule. He notes the following general characteristics of DPs:

A) We can't account for them "by just introducing a rule that replaces, under suitable circumstances, a singular by a plural". I take this to be an argument against global pluralization of the sort just discussed. As evidence, he points out the dependent plurals which are not licensed by NPs.

B) He notes that all DPs must have a plural "antecedent" which has wider scope, whether an NP or a temporal adverbalial. (He also claims that only bare plurals and plural possessives may have DP readings.)

C) "Dependent readings can arise only in cases where there is additional lexical or pragmatic information that makes such a reading probable or even mandatory". He cites the DP reading of (208) as evidence. Since we know that a unicycle can only have one wheel, we interpret wheels as a DP. In further support of this, I note that the structurally identical (218) does not normally receive a DP reading.

(l) (a) Allen's wife is tall and dark.
   (b) Steve has a wife who is similar.
   (b') Steve and Jerry have wives who are similar.

The relative clause on the DP in (5) could be treated by the extended-Link proposal, since it is not symmetrical, as shown by the felicity of (ib).

Besides these characteristics which DPs have in common, deMey claims that there are really two kinds of DP, those with an NP antecedent and those with a temporal antecedent. His evidence for this claim stems solely from the difference in number of the possessive pronouns in (160) and (l):

(160) All the boys have brought their fathers along.
(1) He always takes his girlfriend to such parties.

However, I think this difference has nothing to do with the DP phenomenon, but stems solely from a strict requirement of number agreement on anaphors and their antecedents. In (160) the plural subject happens to be both the anaphoric antecedent, so that the pronoun must be syntactically plural, and the "antecedent" for the DP fathers. In (1), always is the DP "antecedent" for girlfriends, and it is the singular subject which is anaphoric antecedent for his. I will argue for this view of the number of pronouns in Chapter 6.

DeMey suggests that we analyze examples with DPs semantically in terms of Schä's (1981) collective — collective readings of NPs. Both the DP and its "antecedent" receive a collective reading, where if the "antecedent" is a temporal adverbalial it must be possible to analyze it as denoting a collectivity of times or periods. I will not discuss his technical proposal in detail, since it lacks some of the advantages of Link's approach to plurality and distributivity. Essentially, in proposing that the proper treatment of DPs is as collective-collective readings, he claims that the apparent one-to-one relation between the members of the group denoted by a DP and the members of the group denoted by its antecedent is an additional constraint on the truth conditions above and beyond the logical form of the sentence, a constraint which is generally motivated by pragmatic factors and in some cases partly lexically motivated as well.

As an example of what he has in mind, he offers what he calls a "meaning rule" (presumably optional), which would apply to the interpretation of (209).

This rule has to the effect that, in the case of a functional CN such as father-of, if two collectivities are related via the father-of relation and some other relation given by the verb (here bring along), then there is a function from one group to the other which takes, for example, a boy, and gives as output the individual who is that boy's father and was brought along by the boy.
Let me characterize deMey’s collective-collective account of DPs as follows: When one group denoting plural NP a, is under the scope of another group denoting plural b, then a may be considered the range of a function f on b. b may be an NP, or it may be an adverbial element denoting a group of temporal periods, places, etc. The function f is at least partly given by the denotation of the predicate of which a (and b, if it’s an NP) are arguments. In the case where a has a functional head CN, that may contribute to a complex function.

The characterization I have given of deMey’s proposal makes a prediction that not only bare plurals and pronouns, but other group denoting terms can also have DP readings, since there is no principled reason that they should not be able to enter into the same collective-collective interpretations as deMey’s examples. deMey denies this, but I think (221) argues that a DP reading of plural definite descriptions is possible:

(221) Those men married the ex-wives of their neighbors.

It isn’t necessary to interpret the ex-wives of their neighbors as a DP to make sense of the sentence, but we tend to do so.

As deMey acknowledges, the meaning rule he offers for (209) to add the distributive sense which underlies the DP reading is ad hoc in several respects. Just as we saw with the interpretation of sentences containing reciprocals and in Scha’s examples of collective quantification, it is very difficult to develop a single characterization of the nature of the relations between the individuals in the two groups which is adequate to all cases. deMey himself offers (222) and (223), where weaker relations between the two groups may be denoted than in (209):

(222) The boys surprised their fathers with a school play.

(223) The boys helped their fathers build the new school.

It seems like two factors play a role in the DP reading of an example like (209). First, the noun father is functional — there is one for each child. Second, the verb bring along triggers certain pragmatic expectations: where did they bring the fathers along from? Perhaps from home to some event. But which boys would be in a position to bring along any particular father from home? Most likely his own son or sons. Hence, we may interpret a sentence more strictly than its truth conditions require.

The similar contrast between the most natural readings of (208) and (218) undoubtedly lies in similar facts about the denotations of unicycles and bicycles. An even looser relation is suggested in (224):

(224) Junkyards have wheels.

(225) is an example involving another kind of DP licensing adverbial:

(225) In all the rooms, there were smoke detectors.

Here, the locative adverbial licenses the DP reading of smoke detectors. The same pragmatic considerations apply; the DP reading is much less likely in:

(226) In all the rooms, there were books.

The variation in the relationships between two groups is just the sort we found in our consideration of reciprocals and the cumulative reading in Section 3.1.3.2. This suggests that the collective-collective reading may be an appropriate interpretation for many DP examples, but that, as with the reciprocals, it would be futile to attempt to give any but a very weak unified characterization of the relations which may hold between the sets. And this in turn suggests that the functions between collectively interpreted NPs which deMey proposes should not be a part of the truth conditions of sentences such as (209) and (225). They are pragmatically given, as with the character of the relations in cumulative readings and reciprocals more generally.

Though I think deMey is correct in pointing out the importance of pragmatic information (world knowledge, particular lexical items, etc.) and convention in the use and interpretation of DPs, his account in terms of collective-collective relations is not sufficiently general to cover the full range of dependent plural phenomena. There are some technical problems with the treatment of some examples of the DP in this fashion, and, more importantly, if DPs are instances of the more general type of collective-collective readings, we need to explain why we think of them as a separate phenomenon. If we simply treat them as collective-collective readings, like the many others we considered in Section 3.1.3.2, then this blurs the distinction between DPs and other examples.

Technically, examples such as (208), (218) and (224) present a problem because both their subjects and their objects are bare plurals, and the sentences have a generic interpretation. We do not seem to be talking about some indefinite set of unicycles or bicycles or junk yards, but about natural kinds (or, in the case of junkyards nominal kinds — see G. Carlson (1983) for discussion of the distinction). G. Carlson (1977) presents an analysis of bare plurals in which they always denote kinds. It is only certain non-generic ("stage-level") predicates which “lower” them
CHAPTER 3. DISTRIBUTIVITY

to apparently denote individual exemplars of the kind they denote. But if these NPs denote kinds, not groups like the plural indefinite some unicyles, then it doesn’t seem that we can count for the relevant examples by means of a collective, collective interpretation.

The view of bare plurals as kind-denoting is not uncontroversial. See, for example, G. Carlson (1990), which explores a variety of problems with his earlier view, and Wilkinson (1986), who explores the parallels between singular indefinite generics and bare plurals. But it remains to be seen whether an adequate analysis of the bare plural can support an analysis of DP such as deMey’s.

Another kind of problem involves examples with quantificational subjects. Consider:

(227) Few bicycles have horns.

In this example, a distributive plural subject appears to license a DP reading, since we may readily interpret it as meaning that each bicycle under consideration has only one horn. Assuming that the subject does not denote a group, as argued in Section 3.3, there can be no question of the reading arising by virtue of a pragmatically suggested function between the members of two groups.

Edmund Gettier (p.c. to Barbara Partee, 1975) has noticed another kind of DP example with quantificational “antecedent”:

(228) No students wore neckties that were similar.

Like Partee’s (216), the direct object here contains a restrictive relative clause with a symmetric, group predicate. The determiner no makes the subject distributive, but the most natural reading of (228) isn’t that there are few students each of whom wore some set of similar neckties. Rather, we want a reading where there were few groups of students who stood in the wearing relation to a group of similar neckties. The denotation of neckties, including both atomic individuals and i-sums, must be intersected with the extension of the predicate be similar, giving a denotation whose elements are all i-sums whose i-parts bear the similarity relation.

I think one of the problems in offering an adequate analysis of DPs is that given a semantic theory of plurality and distributivity such as I have outlined here, there are in fact too many ways to assign the correct truth conditions to some examples of the DP. For examples with a group denoting subject, deMey’s suggestion about the collective-collective reading, on the purely pragmatic basis I suggest, is plausible. We saw above, in Section 3.1.3.2, that such readings

15. DEPENDENT PLURALS AND GLOBAL AGREEMENT

are quite common, and often seem to introduce some suggestion of a function between individual members of the groups involved, in both reciprocal sentences and those with two plural group denoting NPs. But also, recall that for Link, the extension of a plural CN includes the atomic elements in the extension of its singular counterpart. Hence, strictly speaking, the unicyle has wheels is true if the unicyle has only one wheel. And few bicycles have horns is true if the number of bicycles which have one or more horns is few. We might claim that in the unicyles have wheels the predicate is adverbially distributive, so that the property of ‘having wheels’ is predicated of each individual member of the group of unicyles. Since something may have this property by virtue of having exactly one wheel, this would also give us the proper truth conditions both for the examples with group denoting subjects, and those with quantificational subjects.

With respect to Gettier’s example, (228), the DP reading is just the reading we get through plural quantification, few quantifying perhaps over i-sums of pairs of students instead of atomic individuals. No such pair has the property of wearing similar neckties. Then because neckties are items of personal apparel generally wore one at a time, we understand the sentence to suggest that there is a function from each student in a given group in the domain to the single necktie he wore. Here we see the collective-collective interpretation suggested by deMey under the scope of a plural quantifier.

I think it is important to distinguish our formal treatment of DPs from speculation about the motivation for using DPs instead of their singular counterparts. I suspect that the motivation derives from something like agreement. One kind of case which argues for this view is Link’s (1986) example:

(229) Strange voices were to be heard everywhere.

On the reading where everywhere has wide scope, strange voices may have a DP reading, even though everywhere is not itself syntactically plural. But of course, this adverb seems to presuppose that there are several salient locations. So the plural implicatures here license the DP. This argues that DPs do not arise from some sort of syntactic process per se.

But there is evidence to suggest that they respect constituency in function/argument structures: There are structures which may involve more than one DP, and Barbara Partee (p.c.) has brought to my attention the facts that in such cases there are restrictions on the order of combinations of DPs and singular NPs. Consider her examples in (230):

---

69Note the singular kind in this sentence; of course, it could have the DP kinds instead. This illustrates well the flexibility in our use of DPs and their singular counterparts.
(320) (a) The boys bought cars that had a steering wheel with a leather cover.
(b) The boys bought cars that had steering wheels with leather covers.
(c) The boys bought cars that had steering wheels with a leather cover.
(d) # The boys bought cars that had a steering wheel with leather covers.

In these examples, the whole NP (a) steering wheel(s) with (a) leather cover(s) is referentially dependent on its complement (a) leather cover(s). The generalization seems to be that if this complement is a DP, then the matrix NP itself must be. If the complement is not a DP, the matrix itself may or may not be.

Now consider examples with more than one NP in a VP:

(321) (a) All my neighbors put their dog in a kennel.
(b) All my neighbors put their dogs in kennels.
(c) All my neighbors put their dogs in a kennel.
(d) # All my neighbors put their dog in kennels.

Assuming that in each case where there is a plural in the predicate it has the DP interpretation, the generalization here about which NPs may be plural, which singular seems to be that the direct object must be plural if the locative object is, but not vice versa. This is similar to the facts with ditransitives:

(322) (a) All my neighbors gave their dogs furcoats.
(b) All my neighbors gave their dog a furcoat.
(c) All my neighbors gave their dogs a furcoat.
(d) # All my neighbors gave their dog furcoats.
(e) # All my neighbors gave furcoats to their dog.
(f) All my neighbors gave a furcoat to their dogs.

Again, (322d) seems not to have the intended reading, where there is one furcoat per dog. The contrast between (e) and (f) shows that the generalization cannot have to do with the linear order of the arguments. The direct object must be plural if the indirect object is. It may be that the reason has to do with the order of combination of arguments with the predicate semantically. Putting-in-a-kennel and giving-a-furcoat-to seem to be semantic constituents, while putting-a-dog and giving-to-a-dog do not.10

And consider:

(323) The men carried a spear in their hands.
(324) # The men carried spears in their hand.

Here, the PP is an manner adverbial on carried spears. It seems that if the VP it modifies has a DP, then it too must be a DP.

Link’s (229) argues that the DP phenomenon isn’t only licensed by syntactically plural “antecedents”, but by NPs which implicate the existence of a group (or set of times, etc.). The DPs in the cases we have just considered seem to be involved in a successive dependency relation, which shows that the occurrence of the phenomenon is restricted by something like function/argument structure: once agreement has begun in a constituent, it must spread to nominal elements in more inclusive constituents. I use the term “agreement” because it seems to share some features of verbal agreement. It is syntactically constrained (the verbal dependencies), but there is a certain discretion on the part of the speaker about when to use DPs: DPs are most often optional, they are idiolectal, and they are in some cases conventional (or idiomatic).

I think that functionally, one use of DPs is to avoid confusion about scope and distributivity. Consider:

(324) The men lifted pianos.

Unlike the similar examples in (1) and (2) with singular definite direct objects, one is not likely to give (324) a reading where the group of men together lifted one piano. In fact, it may mean that as a group they lifted a bunch of pianos, but the DP use suggests that there was more than one piano, and, in line with the collective-collective reading suggested by deMey, that there was some function between the lifters and the pianos. Whether the two NPs are in fact both group denoting or the VP is interpreted as adverbially distributive, the use of the DP seems to the same effect in this example, and a range of others, as a quantified subject or floated quantifier.

Though my arguments here are inconclusive about the nature of the dependency relation involving DPs, I believe that they generally tend to support treating DPs, including bare plural DPs, semantically as if they were ordinary group denoting NPs. Whatever their syntactic motivation and manner of percolation, it is

10See Bach (1979, 1981b) for discussion of the treatment of discontinuous constituency in a categorial grammar.
3.6 Conclusions

Now I will briefly review the proposals made in this chapter about the general nature of distributivity. In a predication, where some argument is said to have a property, a distributive reading arises from the presence of a quantificational operator, which is introduced either by a determiner in the NP corresponding to the argument or by an adverbial modifier of the predicate. The predicate in question may correspond to the meaning of a syntactic VP, or it may be derived by abstraction on a non-subject argument. Adverbial distributivity in English may involve an overt floated quantifier such as each or all, or it may arise implicitly. In the latter case, when distributivity is over the count (non-mass) domain, the effect of distributivity is generally well-captured by the D operator of Link. The D operator asserts that the property denoted by the predicate holds of every atomic i-part of the denotation of the subject. There are cases, especially involving plural quantification and distributivity over the mass domain, where the parts over which the property is distributed are not necessarily atomic; in such cases, these parts seem to be homogeneous in character—parts at a certain level of the lattice involved. The evidence does not seem to support readings where distributivity is ever over all i-parts of the subject.

Distributivity is not taken to be due to meaning postulates over a special class of predicates, nor is a "collective" reading due to the nature of a class of group predicates; it is argued that "distributive predicates" do indeed have group readings, and, following Schu, the fact that a predicate takes only group/collective subjects is taken to follow from selectional restrictions on the predicate in question. The existence of a third class of readings, "cumulative" readings, is rejected, and the relevant examples are argued to involve a relation between two groups instead; the relation between such "group/group" relations and Langendoen's Elementary Reciprocal Sentences and Elementary Plural Relational Sentences is explored, and a further weakening of his characterization of Weak Reciprocity, and the corresponding logical form for the Plural Relational Sentences, is recommended.

Several auxiliary hypotheses have been explored, as well. Two classes of determiners are proposed, quantificational and individual denoting, and tests are proposed to determine the correct classification of individual determiners. It is argued that subject-verb agreement is syntactic (though pragmatically influenced) and not semantic in character. However, number in CNs is semantic, in that a singular CN denotes a subset of A, the set of atomic elements in the domain of discourse, whereas the denotation of a plural CN is the set of all elements in the lattice generated by its singular denotation.

Without further stipulation, this analysis accounts for the fact that NPs within the scope of the quantificational operator responsible for distributivity are marked anaphorically, since it is generally assumed that (apart from cases involving modal
Chapter 4

The Representation of Scope

The goal of this chapter will be to suggest a representation for quantifier scope and distributivity which is compatible with the theory of anaphora I outlined in Chapters 1 and 2, and with the discussion of plural anaphora in Chapter 5.

I argued in Chapter 2 that the Binding Theory applies at an S-Structure which contains anaphoric indices. Coincidence at that level will ultimately be interpreted as bound anaphora. S-Structure is mapped directly onto a Discourse Representation, and it is at that level that Discourse Binding is represented, via the equation of discourse referents. The theory of anaphora, then, does not require a level of Logical Form, or LF, in the sense of May (1977, 1985).

In order to make the approach I am suggesting convincing, it is necessary to show how we might represent quantifier scope without LF. There are two ways one might approach this problem. One would be to disambiguate quantifier scope at the level of the DRs. However, May has explored in detail a number of structures where the possibilities of quantifier scope, and in some cases anaphoric binding dependent on scope, are constrained by configurational properties of the sentence in which the relevant NPs occur. Some of the most interesting cases along these lines involve the phenomenon of inverse linking. Because of such cases, the other approach to the disambiguation of quantifier scope without LF seems preferable: the representation of scope at S-Structure. In Section 4.1, I will propose an extension of Williams' (1986) proposal to indicate scope at S-Structure by means of scope indices, showing how we may represent phenomena such as inverse linking by the use of complex indices.

I will then turn to the question of how to represent distributivity. The general hypotheses about distributivity which I proposed in Chapter 3 might be implemented in various ways. Again, the implementation I propose here is intended to facilitate integration with the material on anaphora in Chapters 1, 2, and 5. Recall that although distributivity is distinct from quantifier scope, they interact crucially: the scope of a NP which is interpreted distributively is the predicate which distributes over it. Thus, the representation of this phenomenon must be integrated with that of scope. In Section 4.2 I will discuss an aspect of S-Structure which is important for distributivity: the representation of some cases of adverbial
CHAPTER 4. THE REPRESENTATION OF SCOPE

distributivity with the D operator.

In Section 4.3 I will propose how the representation of quantifier scope and adversative distributivity at S-Structure influence the mapping onto a DR. The idea is simple: NPs with widest scope are mapped first onto the DR. Distributivity, whether induced by a quantified NP or by an adversative operator, involves the introduction of subordinate DAs, or “box-splitting,” which we saw earlier in Kamp’s treatment of universal and conditional donkey sentences, as well as as in any proposed representation of epistemic modal subordination. Since box-splitting puts constraints on the anaphoric potential of material in the subordinate boxes, via the accessibility relation, the anaphoric phenomena associated with distributivity will fall out automatically from this treatment.

As in Chapter 1, I assume that interpretation of a DR is model theoretic. In many respects, the model I assume and the relation between syntactic categories and semantic types are of the sort which is standard in recent extensions of Montague’s “The Proper Treatment of Quantification in English” (1973).4 There is one important departure from the PTQ models: Following the discussion in Chapter 3, Section 3.2, these models will have a lattice structured domain of the sort proposed by Link (1983). Thus, both single objects and groups are treated as individuals, the latter non-atomic, or 1-sums. Discourse referents will not be distinguished according to whether the NPs which correspond to them are singular or plural; the type of individual onto which they are mapped in the model, whether atomic or 1-sum, will depend on any conditions in the DR which constrain their interpretation. However, I will postpone discussion of this aspect of the theory until Chapter 5, in the discussion of plural anaphora.

4.1 The representation of quantifier scope

In the 1960s both George Lakoff (see references) and Richard Montague (see Thomason 1974) independently developed systems for the representation of quantifier scope: Quantifier Lowering within the Generative Semantics framework, and Quantifying in which has come to be called Montague Grammar. Cooper & Patonis (1976) showed various ways in which Montague’s treatment of quantification could be integrated into a version of the Extended Standard Theory. Chemnitz (1975,1976), Sag (1976) and Williams (1977) all argued for the establishment of a level of Logical Form, or LF, within grammars in the Extended Standard Theory, and May (1977) developed an influential treatment of quantification at LF. May (1985) proposes radical departures from his earlier work, permitting him to address more complex issues within the Government and Binding framework of Chemnitz (1981). In frameworks closely related to Government and Binding,

4This is not to be taken as a rejection of recent proposals which introduce radical changes in Montague’s models, e.g. Chierchia (1984) and Landsma (1986), among others. Rather, the issues involved in such discussions do not bear directly on the issues under consideration.

4.1.1 Quantifying in and quantifier raising

Montague (1973) interprets NPs as generalized quantifiers, constituents which take a VP argument to form a sentence. The type of NPs is uniform, whether or not the NP contains a quantificational determiner. In this basically categorial grammar, NPs may either be generated in place or they may be quantified in at the S, VP or CN level.5 When an NP is quantified into an S, in the corresponding translation into intensional logic a lambda operator abstracts on a variable in the S, any \( e_s \) in the argument position of the NP; it also binds any pronominal elements with the same index. The constituent derived by this abstraction is of type \( \langle e_s, \gamma \rangle \); the same semantic type as VPs. The translation of the NP then takes the interpretation of this abstracted predicate as an argument, as in standard subject/predicate combination. This has the same effect as coindexing the NP with the variables \( \gamma \) in the Binding Theory — the result is bound anaphora, as in Bird’s (1983) sense. Versions of Montague’s (1973) syntactic rules for Quantifying in, S14 – 16, are given below, with minor terminological changes, as well as in translation rules T4 – 16, translating the syntactic categories derived in S14 – 16 respectively into expressions of intensional logic (whose model theoretic interpretation is straightforward). \( P_e \) denotes the set of all phrases of the category \( e \); the rules are schemata, so that for any rule \( P_{Fe} \) there are an infinite number of rules \( P_{Fe}, P_{Fe}, P_{Fe}, \ldots \), etc., depending on the index of the variable which is abstracted over; \( \epsilon \) in the following rules denotes a variable, although Montague’s original rules used subscripted pronouns instead of \( \epsilon \). I have given purely extensional versions of the translation rules, in the interest of simplicity:

\[S14.\]
If \( \rho \epsilon P_{Fe} \) and \( \rho \epsilon P_{Fe} \), then \( P_{Fe} \epsilon P_{Fe} \), where either (i) \( \rho \) does not have the form \( e_1 \), and \( P_{Fe} \epsilon P_{Fe} \) comes from \( P_{Fe} \) by replacing the first occurrence of \( e_1 \) by \( \gamma \) and all other occurrences of \( e_1 \) by the appropriate pronominal form (with respect to gender, number, and case), or (ii) \( \gamma \) is \( e_1 \), and \( P_{Fe} \epsilon P_{Fe} \) comes from \( P_{Fe} \) by replacing all occurrences of \( e_1 \) by the appropriate pronominal form.

5The syntactic categories S, VP and CN in Montague (1972) are called 1, IV (1/ε), and CN (1/P), respectively. I use S and VP to facilitate comparison with other theories.
CHAPTER 4. THE REPRESENTATION OF SCOPE

S15. If \( axP \) and \( bxCN \), then \( F(a,x,P,b,CN) \).

S16. If \( axP \) and \( bxP \), then \( F(a,x,P,x) \).

T14. If \( axP \), \( pxP \), and \( a, p \) translate into \( a' \), \( p' \) respectively, then \( F(a,x,p) \) translates into \( a'(x)p' \).

T15. If \( axP \), \( bxCN \), and \( a \) translate into \( a' \), \( x \) respectively, then \( F(a,x,p) \) translates into \( \lambda p'(x)[\lambda p(p)] \).

T16. If \( axP \), \( bxP \), and \( a \) translate into \( a' \), \( x \) respectively, then \( F(a,x,p) \) translates into \( \lambda p'(x)[\lambda p(p)] \).

We see an example of the use of S14, T14 in (1), where the sentence is formed by the application of S14 to the NP and S in (a) and (b) (shown with their rough translations); the translation rule T14 applies simultaneously to give (c):

(1) A lightning bug flew to every child.
    (a) [\( \forall \) every child] \( \lambda P(y)[P(y)] \)
    (b) [a lightning bug flew to \( x \)] \: \text{flew-to}(u \text{-} \text{bug'}, x)
    (c) [A lightning bug flew to every child.] \( \lambda P(y)[P(y)] \) \( \lambda x[\text{flew-to}(u \text{-} \text{bug'}, x)] \)

Quantifying into VP or CN works in a similar fashion, and permits the derivation of the truth conditions shown for example (2), showing the use of S16, T16 for Quantifying In to VP, and Joan Brennan’s (3) (cited in Partee (1972)), showing the use of S15, T15 for Quantifying In to CN:

(2) Every child, has a lightning bug, in her jar and plans to release it.
    (a) [\( \forall \) every child] \( \lambda P(y)[P(y)] \)
    (b) [a lightning bug] \: \text{flew-to}(u \text{-} \text{bug'}, y)

---

Note that the same rule number, 20, is involved as in S14 — the syntactic operation therefore has the same effect on the surface form of the resulting CN as we saw on the resulting S in S14.
(d) \[\lambda PVe[\langle w,y \rangle \text{women's college}(y) \& \text{girl}(w) \& \text{attended}(w,y) \& \text{gave-to}(w,y)](x) \rightarrow P(x)\]

= (by lambda conversion)

\[\lambda PVe[\langle x,y \rangle \text{women's college}(y) \& \text{girl}(x) \& \text{attended}(x,y) \& \text{gave-to}(x,y)] \rightarrow P(x)\]

(e) (4) applied to was included in the list:

\[\forall y[\langle y \rangle \text{women's college}(y) \& \text{girl}(y)]
\&
\text{attended}(x,y) \& \text{gave-to}(x,y)]
\rightarrow \text{was-included}(x)\]

Bennett (1973) retains Montague's essential approach to quantification, extending Montague's fragment to include a variety of other determiners, and making other changes which do not concern us here. As we saw in Section 3.2 of Chapter 3, one of the advantages of this approach is that since any NP can be QI'd, even what I have called the nonquantificational, group denoting NPs, Bennett can give an adequate account of examples where a group denoting NP can have wide scope over an inherently distributive, or quantificational NP. It would be relatively straightforward to add an adverbial distributivity operator to the Montague fragment, as Dowty & Brodie (1984) have done (see Section 3.4.1). Then, abstracting away from problems of discourse binding such as the donkey sentences and adding further determiners along lines suggested by Bennett, Montague's approach would essentially yield the proper predictions regarding distributivity. Of course, the discourse problems are non-trivial, and it is phenomena such as the donkey sentences and modal subordination more generally which force us to reconsider some of Montague's assumptions about the form of a grammar. Hence, we need to consider how we may retain the advantages of Montague's approach in a theory of discourse.6

May (1977) begins from a different perspective: Lakoff (1965,1970,etc.) had noted a strong parallel between the constraints on wh-elements and those on quantifier scope. May explores these parallels in the Extended Standard Theory of Chomsky and associates (see, e.g., Chomsky (1976,1977)) by comparing the characteristics of wh-movement and those of a rule of Quantifier Raising, or QR. QR is taken to map S-Structures onto LF's by moving certain NPs to adjective to a dominating S, in the same way that wh-movement maps D-Structures onto S-Structures by moving wh-elements into the COMP of a dominating S. The NPs

6Note (1986a) contains some very interesting discussion of the relationship between the NP types in Montague (1972) and in Heim (1982) and Barwise (1985). His results there suggest that the differences between them should not be overestimated.

May (1985) offers a very different system for the treatment of quantifier scope, one whose primary goal is not to fully disambiguate a given sentence in this respect, but rather to provide a partially disambiguated representation which allows the incorporation of an account of the subject/object asymmetries studied by Kayne.

Kayne (1981a,1981b) addressed asymmetries involving multiple sub-constructions, that-trace phenomena, and relative quantifier scopes in French and English. He proposed to account for these by means of the Empty Category Principle, or ECP; this principle is taken to apply at LF, in order to handle the quantifier scope cases. The ECP requires that all traces be properly governed (see Chapter 2 example (4), and (9) below for two definitions of government), which means that they are either lexically governed or, in the case of a subject trace (which has no lexical governor) are governed in a specially stipulated fashion by a raised constituent, either a wh-element in COMP or a QR'd NP adjointed to S. In the latter type of government, the wh-element or NP must be the first clausal, so that nothing intervenes between its raised position and the S immediately dominating the subject trace. With multiple sub-constructions, this predicts the so-called Superiority Effects, whereby only the subject wh-element may be ob-served at 2-Structure, as shown by the contrast between (4) and (5):

(4) [who]COMP [s if, likes whom]

(5) *[whom]COMP [s who, like e]
but also discusses a number of examples where interaction of a wh-raised element in COMP and the scope of a quantified NP provides further evidence for some principle along the lines of the ECP. Consider, for example, his (6) and (7).

(6) What did everyone buy εi for Max?

(7) Who, εi bought everything for Max?

(6) is ambiguous, either asking for the identification of the one thing which everyone together bought for Max, or asking, for each person, what that person bought for Max. (7), on the other hand, is unambiguous, asking only for the identity of the person (or group) which bought all of Max’s presents. This type of asymmetry, May points out, appears to be related to the Superiority effects we saw in (4) – (5). He suggests that this shows that the quantified NP everything in (7) cannot be adjoined to S by QR at LF, since that would bring about an ECP violation — who, in COMP would no longer properly govern its trace in the subject position of (7), since everything, adjoined to S, would intervene. How is this related to the relative quantifier scopes which are available for the two examples?

May argues that any two operators which stand in a certain relation at LF can be interpreted with either relative scope, in accord with his Scope Principle, given in (11) below. These are operators which both form part of what he calls a “Sigma Sequence;” the definition of Sigma Sequence, given in (10), in turn depends upon the definitions of c-command (8) and government (9), which May (1985) adopts from Aoun & Sportiche (1981):

(8) a c-commands b =⇒ every maximal projection dominating a dominates b, and a does not dominate b.

(9) a governs b =⇒ a c-commands b and b c-commands a, and there are no maximal projection boundaries between a and b.

(10) A sigma sequence is any class of operators Ψ, such that for any Oi, Oj which are elements of Ψ, Oi governs Oj. By “operator” here is meant phrases in A’ positions at LF.

(11) Scope Principle:
Sigma Sequences are arbitrarily interpreted.

There is only one LF for (6), where, in addition to the wh-element what in COMP, the subject everyone has been adjoined to S. In this configuration, everyone properly governs its trace in subject position, since nothing intervenes. What and everyone here form a Sigma Sequence — as shown in (12), both are in A’ positions, and the only maximal projection which governs either, S’, governs both, so that they mutually c-command each other and, hence, govern each other:

(12) S’
    COMP
    S
    what, everyone
    S
    NPε
    VP
    buy εi for Max

Since the two operators form a Sigma Sequence, they can take either relative scope, so that the two different readings of (6) can both be derived from this single LF.

Examples of this sort had been discussed previously by a number of other authors, including Keenan & Hull (1973) and Karttunen (1977).

Fred Landman (p.c.) argues that although this reading of (7) is very dominant over the blue, the other reading, where the direct object has wide scope over the wh-element, is also possible. He offers the following parallel example, in (i), which does seem to have the relevant, ECP-violating reading in the context (8):

(i) Let’s check to see who fired every employee.

(ii) Here’s a list of the employees fired in the last two years. If you suspect fraud, we have to be very careful. So, let’s check to see who fired every employee.

We will see the importance of the distinction between a maximal projection and its boundary below.
In (7), on the other hand, raising everything to adjoin to S at LF would bring about an ECP violation, since then who would not properly govern its trace in subject position. everything must undergo QR; however, May argues that an object NP may be adjoined to VP, instead of S. He offers various independent arguments for this possibility, to which we may add the example involving anafraphs between conjoined VPs under the scope of the subject, as in (2) above, which motivated Montague's Quantifying In at VP. However, if everything in (7) is adjoined to VP, then it will not form a Sigma Sequence with who in COMP, and thus their relative scopes will be fixed, with the sub-element wider than the quantified object.10

In addition to May's discussion of the subject/object asymmetry, he makes another contribution to the question of the representation of relative quantifier scope in his examination of the phenomenon of inverse linking, exemplified by (13)–(15):

(13) Everyone in [some Italian city] met John.
(14) The head of [every public authority in New York] is a crook.
(15) Someone from [every city] despises it.

May (1977) considered examples such as (13) unambiguous, with some Italian city taking wide scope over the entire subject, and claimed that QR of such quantified complement sentences to S was obligatory. However, May (1985) no longer takes examples such as (13) as unambiguous; he considers (14), where it is clear that the sentence may be true if there is a different person heading each public authority and each is a crook, the inverse reading, or on a non-inverse reading, where there is a single man (e.g., Robert Moses) who heads all the public authorities and is a crook. Examples such as (15) are taken by May (1985) to argue that the complement NP in such inverse examples must have scope over the entire S in order to bind the object pronoun it. This also presupposes that anafraphic binding...

---

8May argues that a subject NP may not be adjoined to VP, since in that case a maximal projection (VP) would intervene between the QR & NP and its trace in subject position. Thus, the subject trace would not be properly governed, and an ECP violation would result.

9For reasons which are not directly relevant here, May (1985, Chapter Five) later argues that the ECP as Kayne (1981b) formulates it is inadequate, and that the proper account of the subject/object asymmetry should be based on Pesetsky's (1982) theory of Paths, a theory which is closely related to Kayne's (1983) theory of Connectives. In each of these theories, certain overall characteristics of a configuration, including multiple operator/variable relations, are considered in determining the well-formulatedness of the representation. In this revision, the adjunction possibilities which I have just described, leading to the contrast between (6) and (7), remain the same.
inability to represent readings where a group-denoting NP has wide scope over a distributively interpreted NP, remains, since May (1985) retains the view of referential and quantificational NPs that only the latter undergo QR. And there are other problems which arise in the attempt to broaden the coverage of the theory.

Williams (1986) points out some problems with accounts of the subject/object asymmetries in terms of the ECP or Patha. He notes that May uses a very narrow range of quantifiers in his examples, principally everybody and everyone, and points out that the asymmetry is not manifest with all quantifiers. For example, each may have wide scope over the wh-element in both Williams' (17) and (18):

(17) Who, does each boy dance with ti?

(18) Who, ti danced with each boy

It has been widely noted that each must take wide scope over some other element in the sentence in which it occurs. Thus, whatever causes the subject/object asymmetries, each is able to overcome it in (18).

Williams also claims that "multiplicity of questions" readings are available in examples involving group-denoting NPs, such as (19) and (20), and not (as May would predict) only in examples involving quantified NPs, such as (8) and (17):

(19) Who, did they dance with ti?

(20) Who, ti danced with them

But Williams denies that this reading is available in all cases where a quantified NP is subject in a question, claiming that the multiplicity of questions reading is not available for (21):

(21) Who, did every girl dance with ti?

I think the data pertaining to the asymmetries are as yet unclear (witness the frequent use of question marks and the difficulty of many judgments in the any detail. De Cearco (1983), following Stowell, suggests the rule:

(i) Adjoin Q to X*.

and offers arguments involving opaque verbs. She does not cite the earlier Montague grammar literature on this subject.

ECP and parasitic gap literature), and that it would be difficult at present even to make a clear descriptive generalization about the extent and character of the subject/object asymmetries, although I do agree that they exist in a wide range of cases, and that they probably involve configurational properties of representations, such as Connectedness, Paths or Koster's (1984) Global Harmony. However, it may well be that other factors are at play as well.

Another factor which complicates judgments about possible answers to questions such as Williams' (18) is the relationship of scope in a question to the possible answers which would satisfy the question. This is a very complex matter, addressed by Engdahl (1979,1980), and Groenendijk & Stokhof (1984), among others. For example, suppose that (18) does not have a reading where there takes wide scope over the wh-moved who. Does this mean that it cannot be satisfied by what Groenendijk & Stokhof call the "pair-list" answer, where we name for each person in the group denoted by them the person who danced with that person? I myself am not sure, and in any case this would require more of an argument than Williams has given.

I also seem to get the subject/object asymmetry in examples such as (22) and (23):

(22) What, did the kids see ti?

(23) Who, ti saw the kids

I think that the kids can have wide scope over the wh-element in (22), but not in (23). What this would suggest is that there is a general subject/object asymmetry phenomenon, but that it involves all NPs, and not just quantificational NPs. This is an area which I think requires a great deal more work.

Another problem with May (1985) arises in cases where an NP is Chomsky adjointed to VP via QR. Consider the type of structure involved:
In this structure, NP1 has been adjoined to the VP. May claims that in this position, NP1 c-commands both the subject and any operators in A' position, but that it does not govern elements in either position. He reasons as follows. As we saw in example (16) above, where an inversely linked NP was QRd to S, the nodes created by adjunction constitute a single projection. In (24), the two VP nodes are a single maximal projection. But, by May's stipulation, a constituent dominated by only one of the nodes in such a multi-node projection is not dominated by the projection. This is relevant for the Aoun & Sporlic definition of c-command given in (8). In particular, in (24) NP is not dominated by the VP projection, so the only maximal projection which dominates it is S, as is the case with A' operators and the subject; hence these elements all mutually c-command each other. But because there is a maximal projection boundary, the top VP node, between NP, and the higher A and A' NP positions, there is no mutual government between these positions. The lack of mutual government has various consequences. Among them, NP cannot form a Sigma Sequence with any operators in COMP or adjoined to S.

The problem arises from the fact that NP, c-commands the subject and COMP positions. On the assumption required for his treatment of inverse linking and crossover, that binding principles apply at I,F, this should permit a quantificational object to bind pronouns in these positions. For some examples, May capitalizes on this possibility, since it allows the backwards bound anaphors required for an account of the crossing confluence sentences, as in (25). However, he does not note that, since such structures do not involve crossover, this feature of his theory also permits the unacceptable readings of examples (26) and (27):

(25) Every pilot that shot at [hit some Mig that chased him.
(26) * His mother loves every boy.
(27) * Which of his friends loves every man.

I take it that this does not demonstrate the undesirability of permitting quantifiers to take VP scope, but rather that, as I argued in Chapter 2, binding in general, and the crossover phenomena in particular, are to be characterized in terms of S-structure A positions, and not at LF.

The remaining argument for applying the Binding Theory at LF is the inverse linking phenomenon. I will show in 4.1.2.1 how this might be accounted for at S-structure as well. There is an empirical problem with May's treatment at LF. Given the mutual governance of the adjoined NP, and its S-adjoined matrix NP, in (16), either relative scope order should be possible. Of course, if NP, with its unbound trace of NP, were taken to have wide scope over NP, then the result would be uninterpretable, and would be ruled out on these grounds. (Since NP, would still c-command its trace in NP, the configuration would satisfy the Binding Theory at LF even where NP, took wider scope in the ultimate interpretation. Thus, the unacceptability is not syntactic under May's account.) The problem arises when there is another quantified NP in the same sentence. After this NP is QRd at LF, it will form a Sigma Sequence with NP, and NP. Given May's Scope Principle, this predicts that any order of scope for the three NPs should be possible (apart from the interpretive requirement just discussed, that NP, have wider scope than NP). But this doesn't seem to be the case. Rather, as Larson (1985) points out, the scope of NP, is tied to that of NP, and no other NP's scope may intervene, as we see in (28):

(28) Two politicians spy on someone from every city.

No reading in which the relative scopes are *every city — two politicians — someone from x* is available for this or related examples, even though in such a reading every city would have wider scope than its trace in the matrix NP. Thus, May's theory predicts unacceptable readings for such examples.12

Note that it is not the functional reading of Engdahl (1980) which is at issue in this example.

12Actually, for reasons which do not concern us here, it would be Chomsky-adjoined to NP, rather than to S.

13Burkern Parlette (p.a.) points out that in examples such as (1), the relevant order of scope does seem to be available, as indicated:
Larson (1985) develops an account of Inverse Linking in a modified Montague grammar fragment which uses Cooper Store, (cf. Cooper (1983)). The central proposal there is that inverse linking arises when an NP internal complement has scope over the whole NP. Larson suggests revisions and extensions of Cooper's system to incorporate this possibility. He is able to predict the correct scope and anaphoric possibilities for a range of examples (not including sub-operators), and he recognizes the piggyback relationship between the scope of the inversely linked complement NP and that of its matrix NP, formally ruling out the undesirable V-2-3 reading of (29). However, the usual problems which the donkey sentences and other examples of discourse anaphors present for Montague grammar remain.

Neither of the two systems for the representation of quantifier scope which we have considered here, Montague's Quantifying In or May's Quantifier Raising at LF, permits us to develop the kind of mapping from S-Structures directly to DRSs which we found desirable on anaphoric grounds in Chapter 2. In the following section I will propose a theory of quantifier indexing at S-Structure, modifying and extending ideas from Cooper & Parsons (1976) and Williams (1986), which will permit such a mapping.

4.1.2 Scope indexing at S-Structure

Cooper & Parsons (1976) propose a grammar which, when certain filters on indexing are added, is equivalent in power to that of Montague's (1973) fragment

\[(\text{i}) \quad \text{Every detective joined in a search for a man with red hair.}\]

The reading which interests us is that where there is a man with red hair such that each detective joined in some one (of possibly several) searches for that man. While I agree that this reading does seem to be available for (i), it is not clear what is at issue here. First, note that the intensionality of search does not seem to be a factor in the availability of this reading, as shown by the availability of a reading with the same scope orders for (ii):

\[(\text{ii}) \quad \text{Every woman made one of the gifts for a new baby in our building.}\]

In (ii), the wide scope of a new baby in our building seems to arise because the NP has a specific flavor, however that is to be analyzed (see Fodor & Sag (1982) for some discussion). But more importantly, I have not been able to find an example where such a scope order is possible with inherently quantificational NPs in the inversely linked complement of the object:

\[(\text{iii}) \quad \text{Every detective participated in a search for many men with red hair.}\]

Many is not generally regarded as a "referential" indefinite determiner. And I cannot get a reading of (iii) along the lines of 'there are many men with red hair such that every detective participated in one of (possibly many) search.'

The English, while using a syntax which consists of a Deep Structure and a Surface structure which are transformationally related, rather than a categorial syntax of the sort proposed by Montague. Quantifier scope is represented by indexing at Deep Structure; NPs are not in general indexed in this system, but only receive a index when they are affected by the rules of Abstraction Marking (for relative clauses) or Quantification Marking (for quantifier scope). The rule of Quantification Marking coincides a node of the category S, VP, or Nom (the categories for which Montague (1973) provided Quantifying In rules) with a) some nonpronominal NP which it dominates, and b) optionally one or more pronouns which it dominates and which are preceded by the NP in (a). Interpretation is off of the Deep Structure, and includes interpretations of Quantification Marked trees which parallel the interpretations the Quantified In structures of Montague which we saw in the previous section. In the general schema for interpretation which follows, I have changed Cooper & Parsons' notation slightly, but essentially, to clarify the relationship of their proposal to others under consideration here, and, again, have given a purely extensional version.\[14\]

\[(9)\]

\[\square\]

\[\text{NP,}\]

\[\text{Det,}\]

\[\text{Nom}\]

\[\text{translates as } \lambda p (x)\]

\[(30)\]

\[\square\]

\[S,\]

\[\cdots\]

\[\text{NP,}\]

\[\cdots\]

\[\text{translates as } \lambda x (S')\]

\[14\text{Recall that the prime notation, e.g. } \alpha', \text{ in the translations here means 'the translation of } \alpha, \text{ and is not the same as its use in } X' \text{ notation.}\]
(31) VP Scope:

\[
\begin{array}{c}
\text{VP}_1 \\
\ldots \text{NP}_1 \ldots \\
\end{array}
\]

translates as \( \lambda y [\text{NP}'(\lambda z_i [\text{VP}'(y)])] \)

(32) Nom Scope:

\[
\begin{array}{c}
\text{Nom}_n \\
\ldots \text{NP}_1 \ldots \\
\end{array}
\]

translates as \( \lambda y [\text{NP}'(\lambda z_i [\text{Nom}'(y)])] \)

(29) tells us that any NP which is indexed is translated as a variable with the same index. This rule enters into the compositional interpretation of the S, VP or Nom constituent whose top node is coindexed with this NP, so that the interpretation of the larger constituent contains a variable \( z_i \) in the position of NP, but then acts as a lambda abstracted argument for the 'true' denotation of the indexed NP, NP'. This gives the same interpretations as Montague's for similar constructions.

Cooper & Parson's fragment is very limited, and, as we saw in Chapter 2, D-Structure as it is currently conceived in Government and Binding theory is not the appropriate level for interpretation. However, Williams (1986) argues for the representation of quantifier scope via scope indexing at S-Structure, and his discussion suggests that the interpretation he has in mind is compatible with their schema for interpretation. I will adopt Williams' idea, extending it to include the representation of inverse linking and showing how its interpretation in terms of DRe parallels the lambda abstractions of Montague (1973) and Cooper & Parsons (1976).

Williams (1986) points out that in general there are four elements of a quantificational structure, or “Q-Structure,” as given in (33):

(33) Q-Structures:
(a) a Quantifier
(b) a restriction on the range of the variable
(c) a variable
(d) a scope

In general, when an NP is fronted by wh-movement, the resulting structure is as in (34), with the elements of (33) as shown:

(34) Which car did John see
[Det car], [John saw t]s
\( a = \text{Det of NP in A' position} \)
\( b = \text{NP in A' position} \)
\( c = \text{trace t} \)
\( d = S \)

(35) In Situ Q-Structure Schema:
[...[Q N']...]
\( a = Q \)
\( b = N' \)
\( c = \text{the A-position with the index 'i'} \)
\( d = \text{the phrase bearing the index 'i'} \)

Here, although the NP has not been moved, it is its A-position which acts as the variable corresponding to (c) in (33), and thus the in situ schema displays the same elements as the adjunction schema, shown in its general form in (36):
(36) Adjunction Schemas:  
\[
[[Q N][\ldots(z)]]_{a}
\]
\[a = Q
\]
\[b = N
\]
\[c = \text{the}\ A\text{-position with the index }'i'\]
\[d = \text{the phrase bearing the index }'i'
\]

By pointing out that (35) and (36) are analogous in both containing the basic elements given in (35), Williams shows how it possible to capture the analogy between quantifier scope and wh-movement which May has emphasized, without requiring actual movement or a distinct LF level of the grammar. We may regard S-Structure as the sole grammatical level which serves as input to interpretation, while avoiding the problems with reconstruction which we discussed in Chapter 2, some of which Williams (1986) discusses as well.

Williams proposes that scope indexing of the sort that builds structures in (35) takes place in the mapping from his NP-Structure to S-Structure, at the same time as the wh-movement which results in adjointed structures like (36). Here, I will regard them as introduced either in a mapping from a D-Structure onto S-Structure or in the construction of a base-generated S-Structure. Since I do not think the issue of whether or not S-Structure is transformational derived is directly relevant to our topic here, I will not attempt to argue for one or the other.

Williams further argues, on the basis of evidence from Slicing (see his p. 269), that it is the scope index \(i\), and not the quantifier itself, which is the operating binding the variable. He does not suggest a formal semantic interpretation of such structures, but if we accept the suggestion that it is \(i\) which binds the variable and treat it as a lambda operator, then we can readily see how Williams’ structures can be given a truth conditional interpretation along the lines suggested in Cooper & Parsons (1979). Besides the S-scope which Williams discusses, I also provide for VP-scope and N-scope:

\(^{13}\text{Williams (1986) himself does not do so. For him, anaphoric relations are represented as NP-Structure. He does not consider interpretation into a discourse level; however, in mapping derivations in his grammar to such a level, we would presumably need input from both NP-Structure and S-Structure, the former for anaphoric information, the latter for operator scope.}

---

(37) Interpretation of in situ Q-Structures:

(a) S SCOPE:  
Interpret a structure of the form  
\[
[\ldots [[Q N][\ldots(z)]]_{a}
\]

as follows:  
\[
\lambda \phi[[N][[[\lambda \zeta \phi(\lambda \zeta)]\lambda[\lambda \zeta\phi]](\phi(z))]]
\]

where \(\phi\) means "the translation of \(N\) with \(\lambda \zeta \phi\) substituted for \(\phi\)."

(b) VP SCOPE:
Interpret a structure of the form  
\[
[\ldots [[Q N][\ldots(z)]]_{a}
\]

as follows:  
\[
\lambda \phi[[V][[[\lambda \zeta \phi(\lambda \zeta)]\lambda[\lambda \zeta\phi]](\phi(z))]]
\]

where \(\phi\) means "the translation of \(V\) with \(\lambda \zeta \phi\) substituted for \(\phi\)."

(c) N SCOPE:
Interpret a structure of the form  
\[
[\ldots [[Q N][\ldots(z)]]_{a}
\]

as follows:  
\[
\lambda \phi[[N][[[\lambda \zeta \phi(\lambda \zeta)]\lambda[\lambda \zeta\phi]](\phi(z))]]
\]

where \(\phi\) means "the translation of \(N\) with \(\lambda \zeta \phi\) substituted for \(\phi\)."

The interpretation of NP, with S scope in (a) corresponds directly with Cooper & Parsons’ (and Montague’s) interpretation for the same constituent, and (b) and (c) extend Williams’ proposal in a natural fashion to yield interpretations which also parallel those of Cooper & Parsons.

In the general theory I am developing, however, S-Structures are not interpreted in intensional logic, but are mapped onto a DR. We will see how quantifier scope indexing affects this mapping in Section 4.3.2.

So far, we have seen arguments that an NP may take scope over an S, VP or N which dominates it. In the following section, I will propose a further extension of scope indexing to permit an NP to have a scope index at a dominating NP. This will permit an account of inverse linking and the scope of possessive NPs. I do not know of good arguments that scope indexing should take place at other constituents, and so I tentatively restrict it here to S, VP, N and NP.\(^{17}\) But \(\lambda \phi\) is an extensional version of Montague’s schema for pronoun translation.\(^{18}\) As Williams (p.e.) has pointed out to me, there is a sense in which the conservative assumption is that quantifier scope indexing may be at any dominating constituent,
before discussing the issue of scope indexing at NP, there are some more general questions to address.

Williams does not discuss the representation of the relative scopes of NPs in sentences with multiple NPs, as in (38):

(38) [Everyone in this room], speaks [two languages].

Suppose that we represent structures where two NPs have sentential scope in the following fashion: instead of simple S1, we have S1/fj or S1/lj, using the slash indices of Halle (1984), though in a way which differs considerably from her intentions. Since most people are familiar with the representation of relative scope in prenex form in the predicate calculus, let us use this familiarity as a mnemonic device and stipulate that the NP corresponding to the first index gets wide scope over any following indexes. As we will see in Section 4.3, in mapping a quantifier-indexed structure onto a DR, the order of scope, widest first, indicates the order of mapping of the NPs onto the DR.

Another issue of considerable interest is the relationship of an NP’s scope to that of wh-moved elements in COMP, and constraints such as the subject-object asymmetries. While May’s data is important, I feel that more research on this topic will be required before a clear picture emerges, and I will have nothing further to say about it here. I believe that any scope orders and constraints which can be represented at LF can be represented with quantifier indexing at S-Structure; in particular, it may be that the relative scope of a wh-element in COMP could be indicated in the same series of indices at S as that of NPs in situ, including unmoved wh-elements. Since the scope indices stand in relation to A positions, just as do the raised NPs in May’s LF, constraints based on subject/object asymmetries should be expressible in terms of the relation of a scope index to the corresponding A position.

Also, although I will interpret indexed S-Structures as indicating fixed scope relations among the NPs involved, this is not a necessary feature of using scope indices at S-Structure, as opposed to QB at LF. One could just as well specify that the series of scope indices at a given node form something like May’s Sigma Sequences, so that they would not fully disambiguate scope for a given S-Structure. The system which resulted would be very similar to that of May (1985).

In the system I envision, scope indexing helps to guide the order of interpretation of NPs in the mapping onto DRs. Each NP must have a scope; however, since its scope may be any S, VP, CN, or NP node which dominates it, 18 including since I have offered no reason why NPs should be constrained to take scope only over S, VP, CN, and NP, and over no other types of constituents. His point is well taken; however, I am being conservative here from a descriptively point of view.

As we will see in the following section, the scope of a possessive NP is obligatorily given as that of its matrix NP.

4.1. The representation of quantifier scope

the NP node itself, the system has essentially the freedom of Montague’s (1973) approach to quantification, where an NP may either be generated in place or quantified in. Any given S-structure will be unambiguous with respect to quantifier scope.

Compare Heim’s (1982) Logical Form, a level which is disambiguated with respect to quantifier scope and simpler than May’s (1985) LF, from which her F-s are derived. There, Quantifier Construal is obligatory for all NPs except pronouns, and involves adjoining the NP to a dominating S. The quantificational NPs, e.g. those with universal determiners, induce a tripartite structure in LF, the three parts being the determiner (or operator) of the NP, its CN (the restrictive clause), and the remainder of the sentence with a variable in place of the NP (the nuclear scope of the operator). Nonquantificational, or individual denoting, NPs induce a bipartite structure, the raised NP and its nuclear scope, the remainder of the sentence.

The distinction between the two types of NPs in Heim’s LF is paralleled in Kamp’s (1981) DRs by the way in which quantificational NPs induce box-splitting in DRs while other NPs do not. Box-splitting in effect puts the material in Heim’s restrictive clause into the left-hand, or antecedent box, and the material in Heim’s nuclear scope into the right-hand, or consequent box. The operator is then syntactic-gerative in DRs, causing the splitting and the consequent differences in mapping from the DR onto a model. Kamp appears to treat the syntactic representations from which his DRs are depictions of with respect to scope; he has suggested (class lectures, 1983, University of Massachusetts at Amherst) operations on DRs which have the essential effect of May’s (1977) Quantifier Raising, changing the relative scopes of the NPs represented. Landman (1986a) and Heim (p.c.) each point out that his system thus corresponds in this sense not to her F-s, but to her LF. But the present proposal differs from Kamp in this respect. DRs here are derived from S-Structures which are fully disambiguated with respect to scope. And so my proposal differs from Heim’s (1982) in lacking an LF.

Williams’ (35) above shows that all the essential elements of her tripartite structures are still available, the operator and its restrictive range in situ but possibly interpreted ‘out of turn,’ for example when indexed at a dominating S, and the A-position itself interpreted as a variable along the lines I suggest in the schemas in (37) above. The system is intended to cover a broader range of data than Heim considers, and it is more flexible than Heim’s in the scopes it permits.

One remaining question is whether pronominal NPs should be assigned scope. Montague (1973) permitted Quantifying In of pronouns, and in general in the system proposed here, this would appear to be relatively innocuous. If a pronoun is discourse bound, the order in which we process it in mapping the sentence onto a DR does not affect its interpretation, so long as an accessible discourse antecedent has already been interpreted, in Kamp’s sense of accessible discussed in Chapter 1. If this is not the case, then an ill-formed reading, with a free variable, results. On the other hand, if a pronoun is C-command bound, then by the indexing algorithm proposed in Chapter 2, it is already coindexed with its binding antecedent when
the mapping to a DR takes place, and hence, correctly, they will have the same scope.

I will leave one final question unanswered. Roger Higgins (p.c.) has pointed out that in general the use of indices may actually be a way of overlaying one structure on another. In the present case, this would amount to S-Structure representing two structures at once, one the surface order of the constituents (more or less), and the other, via various indices, a sort of covert LF, giving binding and scope relations. Similarly, Cooper & Parsons (1976:344) note that in their system, "nothing hinges on the fact that we mark the indexing in the trees. It would be possible, though less perspicuous, to represent the indexing as a separate object, not unlike Jackendoff's (1972) tables of reference, and then define the translation procedure on a tree and its indexing in a way strictly analogous to the present proposal." I concur with these observations. It may well be that what we have here is two structures in one. Yet, as we saw in the discussion of the Binding Theory in Chapter 2, the information which is thus available at the combined S-Structure and index structure to guide interpretation is not the same as that available in a transformationally derived LF of the type proposed by May (1977) or (1985). Because non-un-moved NPs remain in situ, we retain aspects of the underlying S-Structure which are crucial for the understanding of anaphora and its relation to quantifier scope, aspects of S-Structure which LF with Reconstruction and Weak Crossover does not adequately represent.

4.1.2 Inverse linking and the scope of possessive NPs

In our discussion of May's (15), repeated below, we noted, following Larson (1985), that the inverse linking in these examples indicates that NP-internal complements may take quantificational scope over the whole matrix NP, and in addition that this scope appears to ride piggyback on that of the matrix, so that the scope of other NPs in A positions in the same sentence could not intervene between that of the inversely linked complement and that of its matrix. Of importance is developing an adequate account of inverse linking is the fact, also illustrated in (15), that the wide scope of the complement NP, here every city, licenses it to bind a pronoun c-commanded by the matrix, someone from every city.

(15) Someone from [every city], despises it.

We may observe a closely related phenomenon in (39), involving a possessive NP.

(39) [Everyone's mother], loves him,

Here as well, everyone, with wide scope over its matrix NP, binds him, although the pronoun is only c-commanded by the matrix, and not by its binder.

4.1.2.1 Inverse linking and the scope of possessive NPs

Jackendoff (1977) earlier pointed out that an N complement (but not an N complement) could take scope "out of an NP dominating it," and supported this claim with examples such as (40), where the N complement in the subject NP, few children, licenses a negative polarity item, any, in the direct object. This example contrasts with (41), where an N complement cannot license the negative polarity item.

(40) Fathers of few children have any fun.

(41) * Fathers with few children have any fun.

We will discuss the N'/N" contrast below. At this point, (40) may be taken as further evidence that it is the wide scope of the subject's complement NP which is the central feature of the inverse linking phenomenon. And the possibility of negative polarity items in the similar (42) argues that wide scope is central in cases with possessive NPs such as (39) as well:

(42) Few children's fathers have any fun.

Two other kinds of examples support the idea that inverse linking licenses what I have called c-command anaphora, as opposed to discourse anaphora. First, in all the cases where an inversely linked complement or a possessive NP binds a pronoun, such as (15) or (39), the matrix NP c-commands the bound pronoun. (43) shows that the matrix NP need not be the subject, while (44) shows that binding is not possible when the matrix doesn't c-command the pronoun:

(43) (a) The FBI warned the secretary of every expected spy to keep an eye on him.
(b) The FBI warned every spy's secretary to keep an eye on him.

(44) (a) * He is loved by the mother of every boy in Amherst.
(b) * He is loved by every boy's mother.

The other type of example supporting the c-command requirement on inversely linked and possessive NP binding involves sloppy identity. The following are

35See Ladefoged (1978) for arguments that a negative polarity item must be in the scope of a downward entailing operator, such as the monotone decreasing determiner few.

36The examples which I am using are drawn from unpublished experimental psycholinguistic work which I carried out with the help of Chuck Clifton and Lys Frazier.
CHAPTER 4. THE REPRESENTATION OF SCOPE

generally conceded to be quite acceptable on the sloppy reading:

(45) The owner of every cat taught its to be well behaved, and the owner of every dog did too.

(46) Many girls’ fathers are concerned about their schooling, and many boys’ fathers are too.

There is only one type of example which I am familiar with which convincingly argues that inversely linked complement NPs may not have the same general potential to serve as binders as do NPs which directly c-command anaphors. (47a) and (48b) are taken from Reinhart (1977):

(47) (a) Every organisation suffers some setbacks in its early years.
(b) In its early years, every organisation suffers some setbacks.

(48) (a) Members of every organisation suffer in its early years.
(b) * In its early years, members of every organisation suffer.

Here, the matrix subject containing the inversely linked NP every organisation, interpreted with wide scope, c-commands the trace of the proposed constituent which contains it, so that under the characterization of c-command anaphors given in Chapter 2, every organization should be able to bind it; however, the sentence is unacceptable on the binding relations indicated, in contrast to the well formed (47b), where every organisation directly c-commands the trace. I have no account of why this is so. Several possibilities arise: one is that the system of binding I proposed in Chapter 2 is incorrect; another is that binding by inversely linked NPs is the fall of 1985. The experiment was an attempt to ascertain the availability of sloppy identity in sentences where the binders are inversely linked or possessive NPs. The judgments on some examples with inverse linking were mixed, but preliminary results indicate that the likelihood of a sloppy reading increases when the head noun of the matrix NP which contains the inversely linked complement is functional, especially when it involves “inalienable” objects or relationships such as body parts or kin. In the same study other examples testing for the availability of the sloppy identity reading without inverse linking or possessive NPs also seem to show an increase in probability of the sloppy reading when alienability is involved, so this may be a characteristic of sloppy identity, rather than a symptom of a weaker binding relation when the binder is inversely linked.

Also, in general, possessive NPs seem to be slightly more likely to act as binders in sloppy identity than inversely linked NPs. I would attribute this to the obligatory wide scope of possessive NPs, as discussed below.

is not c-command binding; a third is that there are extra constraints on binding into preposed constituents, besides c-command of their trace by the binder. Since there seem to be other problem examples involving preposed constituents (cf. the discussion of Lakoff’s Near John he saw a snake in Chapter 2), I will assume here that the explanation lies in that area.

Under the assumption that binding by an inversely linked or possessive NP is c-command binding, I propose to account for these examples by giving the inversely linked or possessive NP scope over the matrix NP in which it occurs. This may be represented by adding a slash index to the matrix NP, as shown in the indexing schemas in (49) and (50):

(49) Inversely Linked Complement NP (optional):

\[ NP_{ij} \]

(50) Possessive NP (obligatory):

\[ NP_{ij} \]

In both schemas, the index of the NP-internal NP, is added to that of the matrix, taking wide scope over it in the same way as when there are two or more scope indices at S.21 There are two principal advantages of this approach to wide scope of complement or possessive over its matrix NP. First, inverse scope as represented by Kayne (1981a) argues that QR can’t adjoin to NP, on the basis of the fact that (i) cannot be synonymous with the reading of (ii) where the scope of nobody is confined to the complement sentence:

(i) John is bemoaning nobody’s presence.

(ii) John is bemoaning that nobody is present.

But under the present proposal, this is exactly what we would expect — nobody in (i) is adjoined to the direct object NP, which may have either ‘in situ’ VP or S scope. But, on any of these options, since John is not distributive, the same reading will result, where there is nobody whose presence John bemoans. More important, consider (iii):

(iii) John is bemoaning nobody’s presence.
in this fashion will have the piggyback characteristic we found to be desirable when the matrix NP is given scope, let's say at \( S \), then its entire index will be copied, \( k \). No other scope index will intervene between \( i \) and \( j \) because the scope indexing procedure simply adds one (possibly complex) index at a time to the front of any indices which are already present at the given node. Second, we may now stipulate that the matrix NP which results from (49) or (50) may c-command bind any pronominal indexed either \( i \) or \( j \). Thus, scope indexing at NP also serves as anaphoric indexing. This explains the binding in examples (13), (49), and (43). In examples (45) and (46), as in the sloppy identity examples discussed in Chapter 2, the sloppy identity is licensed by c-command binding in the first conjunct, with the indices of the corresponding constituents in the second conjunct substituted at the DR level.

In keeping with Jackendoff's claim that only N complements may take wide scope out of NP, (49) requires that NP; be N-internal. Jackendoff argued that N complements, and only N complements are subcategorized arguments of the head noun. He offered three tests to distinguish between N and NP complements. One is the possibility of inverse linking (though he didn't use this term). A second is order, under the assumption that an N complement may not be separated from its head. The third is behavior with one anaphora; under the assumption that one is an N anaphora, N complements, but not N complements may serve as complements of one. However, Jackendoff points out that one anaphora is an adequate test because it is valid only for NP complements, not for other PPs in N, it seems that some other PPs which are N complements by the first two tests also occur with one, giving mixed results.

I believe that Jackendoff's claim that it is only arguments of the head which may be inversely linked is plausible, though I find the suggested syntactic tests rather weak. Here, I merely make the restriction of inverse linking to N complements a stipulation, although it is possible that this might be made to follow from a requirement that the head govern its arguments.

However, if only N complements are true arguments of the head noun, then what about the head's relation to possessive NPs? This brings us to a discussion of the difference between the two kinds of cases, a difference encoded above in the fact that the indexing schema for inverse linking in (49) is optional, while that for possessive NPs in (50) is obligatory.

(iii) Everyone bemoans nobody's presence at the dance.

I find this example ambiguous: either 'there is no one whose presence everyone bemoans' (the most popular missing person interpretation) or 'everyone bemoans nobody's presence at the dance last night' (e.g., since the committee had worked so hard).

See rule (60) below giving the proposed intentional logic translation for 's.

*Here again my debt to Balk (1984) should be obvious, although I use 'indirect binding' in a way very different from her intentions. She does not seriously consider the inextricably linked examples in her paper.*

11 Consider the following:

(51) (a) John's picture
    (b) John's picture of John
    (c) John's picture of John's
    (d) John's picture of John's picture
    (e) John's picture of John's picture of John

In (51a), the possessively marked John may be construed as filling one of three relations to the head, picture: owner, painter, or subject of the study. However, in (b) and (d), the complement John may only be the thematic patient of picture, that is, the subject of the study. In (c), the possessively marked John may be either the owner or the painter of the picture, and possessive Al in (d) may bear these same roles. But (c) shows that two possessively marked NPs may not occur in the same sentence. I suggest that this is explained as follows: In an unpublished work, Barbara Partee (p.c.) suggests that the relation between a possessive NP and its head is introduced by a free variable \( R \) over relations in the translation of 's; let us suppose that we implement this idea in an intensional logic translation of 's, as follows:

\[
\lambda \alpha \lambda \beta \lambda p \lambda x. R(x,y) \cdot (\forall x, y \cdot z = y \land R(x,y) \land P(y))
\]

What this formula means is that 's first takes an NP (\( P \), as in John's above, to make a determine type, that is, a function from a CN (\( Q \)) to a function from one place predicates (the VP-type variable \( P \)) to truth values. The constituent which results after combining 's with John is thus the usual type of determiners — a relation between two one-place predicates, the denotations of the CN and the VP.

The formula specifies that there is some two-place relation \( R \) which holds between the possessive NP and the unique element in the extension of the CN (the uniqueness clause is underlined here for clarity); the translation thus builds in the often noted definiteness of possessive NPs. As Partee has suggested, we will consider the value of \( R \) to be contextually given. This explains the variability of the relations between the possessive and the head in (51a) — since there are various possible relations that a person might bear to a picture and since we have no context here to suggest that one is more salient than the others, we may suppose that John bears any of these relations to the picture under discussion. This is

39 Here is a variable over NP types, \( (\alpha, \gamma, \lambda) \) in an extensional system; \( P \) and \( Q \) are variables over the type \( (\alpha, \gamma), \) i.e., the type of CNs and VP's; \( z, y, \) and \( z \) are individual type variables. \( R \) is a variable over relations between individuals.
in contrast to the subcategorized complement John in (51b), since arguments are generally taken to have conventional thematic roles with respect to the predicate which subcategorizes for them. In (51c), I follow Stockwell, Schachter & Pacific (1973), in considering the possessive to be extraposed; we may assume that the impossibility of interpreting John’s here as the subject of the picture arises out of a sort of functional efficiency: since the true complement in (51b) bears this relation to the head in the same position, the extraposed possessive NP lacks this relation by contrast. This is similar to the possibilities for the interpretation of R in (51d), where the subcategorized John already has taken the role of theme. Finally, (51e) supports the extraposition account of (51c), showing that an NP with two possessive NPs is ill-formed.

Note that (52) also automatically gives the possessive NP wide scope over other elements of the NP. Of course, since we now have the possibility of quantifying in at NP (or its equivalent in DR terms), an N complement might still take wider scope than the possessive NP.

Now consider the following examples, parallel to those in (51), but with the inherently distributive few CN, instead of the referential John, permitting us to test for the possibilities of inverse linking:

\[(53)\]  
(a) Few people’s pictures  
(b) (the) pictures of few people  
(c) * (the) pictures of few people’s  
(d) Ali’s pictures of few people  
(e) * Ali’s pictures of few people’s  

(53a) has only one interpretation, where few people has wide scope over the entire NP; we are not talking about pictures whose subjects are few, but about the range of a function whose domain contains few people, each mapped onto his or her picture. The inverse character of the possessive licenses anaphoric binding, as when (53a) takes a predicate such as look like them, binding them; and because few is monotone decreasing, the predicate may contain a negative polarity item, as in the predicate are attractive at all. As in (51a), the type of relation which each of the few people in the domain may bear to his or her picture is open: owner, painter or subject; but since my intuition is that they must all bear the same relation to their pictures, I have given R wider scope than P in (53).

We already saw that the inversely linked reading of complement NPs is optional in May’s (14), repeated here:

\[(14)\]  
The head of [every public authority in New York] is a crook.

(53b) as well may have either the inversely linked reading, where the truth conditions are similar to those of (53a), or “relational reading,” where each of the pictures under consideration (however many there may be) has few subjects.24 But on either reading, the relation which the argument bears to the head is not free, as with the possessive NP cases, but is given by the subcategorization; here, as in (51b), it is a patient, the subject of the painting.

I cannot say why the extraposed quantificational NP seems to make (53c) unacceptable. It seems likely that it is a scope problem, since the parallel (51c) was fine. I can only speculate that extraposed elements behave as adjuncts (N or, more likely, N complement); as Jackendoff points out, these may not take scope over the matrix NP. Since the extraposed NP in (51c) was nonquantificational, this caused no problem for interpretation. For some reason here, the narrow scope of adjuncts is not acceptable in a possessive NP, which we are accustomed to giving wide scope.

(53d) is of interest with respect to the prediction that quantifying in at NP would permit the QJD NP to have wide scope over a possessive NP in the matrix. (53d) doesn’t have an inversely linked reading, where the predicates look like them (then bound by few people) or are attractive at all would be acceptable.25 One way to handle this problem would be to stipulate that only one NP index may “propagate” up to the matrix NP. There are two types of examples which show this:

I believe this terminology is unfortunate, since on this reading few behaves more like a one-place predicate; cf. the predicative few of Section 3.2.4.26

May (1985) claims that the parallel (i) does have an inversely linked reading, as opposed to (ii):

(i) John’s pictures of everyone are hanging on the wall.  
(ii) John’s picture of everyone is hanging on the wall.

May attributes this difference in scope possibilities to the Specificity Constraint of Feag & Higginbotham (1981), under the assumption that “singular (as opposed to plural) NPs are specific.” However, I know of no independent support for such an assumption. Further, the closely related (iii) is not acceptable, so that the examples seem to fail the anaphoric binding test for inverse linking:

(iii) John’s pictures of every woman are hanging on her wall.

Williams (1986) points out that everyone (as opposed to every CN) may have a group reading. (iv), with a collective predicate, provides evidence for this claim:

(iv) Everyone gathered in the square at 6pm.

If everyone in (i) has such a group reading, then we might have a group-group relation between the denotation of the whole subject and that of its complement. As discussed in conjunction with cumulative readings in Section 3.1.3.2, there might then be an extra-
that this is not the correct generalization. First, there are cases of inverse linking
where the inversely linked NP itself has an inversely linked complement, as in the
following example, after Lanson (1985):

\[(54)\]

The National Enquirer has been looking for a gossipy friend of
every debutante in an obscure midwestern city.

Suppose that the National Enquirer’s readers love scandals about small-town
America. What the magazine wants to find is some obscure midwestern city
where every debutante has a gossipy friend who will tell all. The fact that the
most deeply embedded NP, an obscure midwestern city has the de dicto reading,
so that it is under the scope of the opaque verb looking for, shows that it does
not receive Fodor & Sag’s (1982) specific reading, which would give it wide scope
without inverse linking. The reading in question (which I think is available) may
only be obtained if an obscure midwestern city has inverse scope over its matrix,
NP, with the head debutante, and every debutante in an obscure midwestern city
in turn has inverse scope over its matrix, NP, with the head friend. Hence, the
direct object should end up with the index i/j/k.\(^{26}\)

One might modify the constraint proposed to handle (53d) by saying that only
one application of (49) per NP node was permissible, since then cyclic application
would permit the intended reading of (54). Or, one might simply say that the
presence of a possessive NP blocks the application of (49).\(^{27}\) However, the other
type of example where an NP receives more than two indices argues against either
of these approaches. Consider (55):

\[(55)\]

[[Some superpower]]’s destruction of [[every city]]’, killed all
of its inhabitants.

where both NP, and NP, have wide scope over the matrix NP. Some superpower
may also bind a pronoun, as in the extension with it’s advanced technological
superiority. Another possible reading of the subject is where some superpower has
widest scope. At first one might be tempted to represent this by indexing the

\(^{26}\)One test for whether the indexing suggested is correct, in view of the binding poten-
tial of such indices, is to test whether all of the ‘stacked’ NPs in (62) can bind a pronoun
with the full matrix NP in commands. For example, assuming that the beds are
all male, can the various indices on the full matrix bind the pronoun is the following:
to tell everything he knows about her contempt for it’s morals? I’m not sure. However,
not that May’s system would make the same predictions as the present proposal in this
case.

\(^{27}\)See, for example, Fiengo & Higginbotham (1981). They might claim that (61d)
is specific because of the proper name in SPEC, and hence violates their specificity
constraint.

4.2 The D operator

Here, I will briefly consider the S-Structure representation of the adverbial distri-
butive operator D which was discussed in Chapter 3, Section 3.4.2. The in-
clusion of D into S-Structures will permit us to map onto the discourse level
from a level which is fully disambiguated with respect to both NP scope and

distributivity.

Recall that D may operate on either the syntactic VP or on a predicate de-

drived by lambda abstraction, so that, for example, a group denoting object NP

may distribute over a lambda abstraction of the dominating S, the lambda binding

a variable in place of the NP. The representation of distributivity at S-Structure

must thus take the more abstract possibility into account. For the sake of sim-
plicity, the representation I propose introduces D only in conjunction with scope

indices at an S node. This index on D will be used in the SS to DR mapping.

The schema for introducing D at SS is as follows:

\[(57)\]

For any NP index i, S \ldots (i(D) \ldots) indicates that the D operator applies
to the predicate which is the argument of NP.

The proposed representation rules out implicit distributivity for some representa-
tions of NP scope, for example where the subject (or some other NP) is interpreted
in situ; however, I see no harm in this so long as there is some representation
available of the distributed reading, i.e. where the subject is scope indexed at S.
In order to support the claim that (57) is adequate for the representation of distributivity, we might consider the so-called "small clauses," which, following Williams (1983), I regard as nonconstituents. Distributivity in (58) and (59) is induced by the determiners of the underlined NPs. In (60) it is suggested by the meaning of tarred and feathered, while in (61), it is induced by D:

(58) I want every oyster raw in its shell.

(59) I left many oysters raw in their shell.

(60) I saw the gamblers tarred and feathered.

(61) I need the dining room chairs glued back together.

Although (63) could mean that the speaker wants the chairs all glued into one unit (the group interpretation), the more likely interpretation is that she wants the joints of each chair reghed. Under the proposals for scope indexing and the representation of D, this reading may be represented by the S-Structure in (62), with an interpretation along the lines of (63):

\[
\Gamma(\lambda x [\text{the chairs' need}(x, \text{glued}(y))])
\]

Under the interpretation of D discussed in Chapter 3, its application to \(\Gamma(\lambda x [\text{the chairs' need}(x, \text{glued}(y))])\) in (63) means that each member of the group denoted by the chairs will be glued, i.e. each chair will be glued. Thus (57) is adequate for the treatment of distributivity in the "small clauses."

\[\lambda y \lambda z \lambda x [\text{need}(x, y) & \lambda x \lambda y [\text{glued}(y)]]\]

D operates on VP type constituents, of type \(\langle e, r \rangle\) in Montague's system, so it would not apply to (52), which is of type \(\langle \{e, r\}, \{e, r\}, \{e, r\}, \{e, r\} \rangle\). The portion of the formula where the possessive NP is the subject of a VP type lambda expression is shown in (66):

\[\Gamma(\lambda x \lambda y [\text{need}(x, y) & \lambda x \lambda y [\text{glued}(y)]]\]

The girls, substituting for \(P\) in (66), have the property of being a (nonatomic) individual \(x\) whose unique father-group (here the relation \(R\) is suggested by the fractional character of the head noun father) bought them cars (the VP's translation substituting for \(P\)). If the property is modified adverbially by \(D\), then each atomic \(i\)-part of the nonatomic individual denoted by the girls would have the property that her unique father bought her cars. (Recall that the denotation of cars includes the denotation of car.)

Distributivity of the girls over its matrix would work similarly in (65). Sup-
pose that in a Montague grammar framework the translation rule for Quantifying in an NP would be something like (67), parallel to the Montague’s (1973) rules T14 – T16 given in modified form in Section 4.1.1.36

(67) If \(P_{\text{fam}}\), \(P_{\text{fam}}\) and \(a, b\) translate into \(a', b'\) respectively, then \(P_{\text{fam}}(a, b)\) translates into \(\lambda x.\lambda y.\lambda z.(P(x, y, z))\).

Here, the distributive operator \(D\) would operate over the derived predicate \(\lambda x.\lambda y.\lambda z.(P(x, y, z))\) so that each atomic \(i\)-part of the quantified in NP (the girls) would have the property ("being an \(x\) such that \(x\)’s bought \(x\) a car").

The question of whether \(D\) may apply in the way just discussed (or in the truth conditionally equivalent DR interpretation to be proposed in the following section) is obscured by two factors. First, it seems that the head of an NP with a plural possessive or an inversely linked NP complement NP must itself be morphologically plural. Consider the following examples, where the inherently distributive character of the determiner few rules out the possibility that it discourse binds the pronoun they:

(68) (a) The owner of few cats thinks they’ve got fleas.
    (b) The owners of few cats think they’ve got fleas.
    (c) Few cats’ owner thinks they’ve got fleas.
    (d) Few cats’ owners think they’ve got fleas.

(69) (a) A resident of few cities hates them.
    (b) Residents of few cities hate them.
    (c) Few cities’ resident hates them.
    (d) Few cities’ residents hate them.

The (a) examples, where the bound anaphora indicated by underlining requires that the \(N\) complement be inversely linked, are unacceptable because their head nouns are singular. This is in contrast to the (b) examples, which are identical except that the head is plural. The (c) examples are completely ungrammatical like the (a) examples, they have singular heads, but, as predicted by (50), they can only have scope over the entire NP. The cause of the ungrammaticality seems to be a clash between the singularity of the head and the NP scope of a plural NP. The full acceptability of the (d) examples confirms this generalization.

My guess is that this phenomenon is related to the dependent plural phenomenon discussed in Chapter 3, Section 3.5. When the matrix has scope under

---

36This rule is quite similar to that given in Larson (1985), and amounts to the same thing truth conditionally.

4.3 Scope and distributivity in discourse representations

In the grammar I am arguing for here, S-Structures are mapped onto DRSs, rather than formulas of intensional logic. This mapping is not bottom up, as in Montague Grammar, but top down, as in Kamp (1981) and Heim (1982). A DRS, with discourse binding introduced consistent with the accessibility of (possibly accommodated) antecedents in its hierarchical structure, is then interpreted in a model. In Chapter 5 I will discuss the representation of number in DRSs. Here, I am concerned with the representation of distributivity, whether introduced by explicit or implicit adverbial distributivity in a predicate or by the determiner of its subject NP. In general, it will be seen that both kinds of distributivity bring about box-splitting in DRSs. And this, without further stipulation, will explain the unaphor constraints on NPs under the scope of a distributive operator (determiner or adverb), just as it does in Kamp’s (1981) and Heim’s (1982) treatments of the classic donkey sentences.

4.3.1 Determiners and the mapping onto discourse representations

As discussed above, Kamp (1981) and Heim (1982) both claim that there are two kinds of NPs, those, such as singular indefinites, which are interpreted as variables and those, such as universally quantified NPs, whose determiner sets up a relationship between the denotation of the CN and that of the predicate of which the NP is subject. The first type of NP is what I called in Chapter 3 an
individual-denoting NP. Recall that the group-denoting NPs are a subset of the individual-denoting NPs, those whose denotations include nonatomic elements of the lattice-structured domain. The second type of NP is what I called the quantificalational NPs. In Chapters 1 and 2, I simply assumed that the two types of NPs behaved differently in the mapping onto a DR. Here, I will discuss in more detail how the discourse representation of NPs is related to their individual-denoting or quantificalational character.

In DRs, all individual-denoting NPs are represented in much the same way as singular indefinites, with the possible addition of further conditions, such as the cardinality condition in the representation of an NP with a numeral specifier, or the anaphoric condition on definites (see Chapter 5). Recall that in Heim's theory, LFs involving such NPs have a bipartite structure; the NP is prefixed to the S in which it occurs, and the S itself, with a variable in place of the NP, is called the nuclear scope of the NP. Such an LF maps onto a file where the predicate denoted by the CN and the predicate denoted by the nuclear scope of the NP (as if it were a lambda abstraction on the NP variable) are conditions on the discourse referent introduced by the NP in the file. In DRs, similarly, for any individual-denoting NP we simply enter a discourse referent, with the CN and the nuclear scope of the NP acting as conditions on the discourse referent. The following is a general characterization of the mapping of nonquantificalational NPs with sentential scope onto a DR:

(70) Mapping individual-denoting NPs onto a DR:

To map a constituent of the form [[...[DET CN]....]] onto a DR, where NP, is individual-denoting, enter a discourse referent \( x_1 \) into the DR, along with the condition \( CN(x_1) \). Then enter the nuclear scope of the NP into the DR, where the nuclear scope is \( S \) with the variable \( x_1 \) in place of \( N_1 \).

Although the mapping is exemplified in (71), where the direct object in the predicate has been ignored for simplicity:

(71) [The man]; lifted a piano.

\[
\begin{array}{c}
x_1 \\
\text{man}(x_1) \\
x_1 = x_2 \\
x_1 \text{ lifted a piano}
\end{array}
\]

here, the equation \( x_1 \) with \( x_2 \) is intended to satisfy the anaphoric requirement on the definite NP the man; \( x_2 \) must be a pre-existing discourse referent accessible to \( x_1 \) in a larger DR, as in other cases of discourse anaphora we have examined. The truth conditions for (71) will be 'the (already salient) individual which is a man has the property of having lifted a piano.' The nuclear scope may also be represented as 'lifted a piano(\( x_2 \)).'

The discourse representation of all quantificalational NPs involves box-splitting. As I noted above, this reflects Heim's tripartite structures at LF: the left hand or antecedent box represents the CN, Heim's restrictive term, and maps onto the set which it denotes, while the right hand box represents Heim's nuclear scope, the sentence with the NP replaced by a coindexed variable. Kemp's (1981) treatment of the universal operator is syncategorematic: the arrow between the two boxes tells us that the set denoted by the left hand box must be a subset of that denoted by the right hand box, so that, as in generalized quantifier theory, the determiner is essentially a relation between two sets. Other quantificalational determiners may also be treated in DR theory as relations between two sets, so that box-splitting occurs but the relation between the sets denoted by the boxes differs from determiner to determiner.\(^{[72]}\) The following is a general characterization of the mapping of such NPs:

(72) Mapping Quantificalational NPs onto a DR:

To map a constituent of the form [[...[DET CN]....]] onto a DR, where NP, is quantificalational, form two subordinate boxes, the left accessible to the right; enter a discourse referent \( x_1 \) into the left hand box, along with the condition \( CN(x_1) \). Enter the nuclear scope of NP, into the right hand box of the DR. DET serves to characterize the relation between the two boxes in the embedding into a model.

In line with (72), we might represent a sentence such as (73) roughly as in (74) (to be revised), again ignoring for the time being the singular direct object:

(73) [[Few men]; lifted a piano.

\(^{[72]}\) I understand that Root (1986) works out something along these lines for the treatment of determiners such as many and few in Discourse Representation Theory, but I have not had the opportunity to read her account.
Thus, the distributivity in a quantificational determiner leads us to consider all and only the atomic elements in a given set.

However, recall that atoms need not be individuals in the pretheoretic sense. This is crucial for the representation of examples such as Link's (76), discussed above in Chapter 3, Section 3.2.4:

(76) All competing companies have common interests.

I argued in that section that these examples do not involve Link's plural quantification. Instead all quantifies over units each of whose members are competing companies. If we are willing to concede that each such unit is conceived of as an atom (possibly impure, in Link's sense), then the representation for (76) will be as in (91), parallel to that in (75):

(77) Few men lifted a piano.

Here, each of the atomic elements in the set denoted by *competing-companies involves more than one company, with truth conditions something like: 'each unit which consists of competing companies has common interests.'

When NPs do not have sentential scope, they behave in a fashion similar to the examples shown above. Individual-denoting NPs simply introduce a coindexed discourse referent, and the remainder of the sentence is reduced by replacing the NP with a coindexed variable. Quantificational NPs always induce box-splitting, introducing a discourse referent over atomic individuals in the lefthand box, with the remainder of the sentence entered in the righthand box. We will see further examples of this in the following section.

4.3.2 Scope indexing, $D$, and the mapping from S-Structure to discourse representations

The mapping from an S-Structure to a Discourse Representation proceeds in a top-down and left-to-right manner. At any given node, we first consider any NPs whose indices appear on that node. At an S, for example, we first consider NPs whose scope indices appear on the S (if there are any), widest scope first. We
then consider its daughters left to right, first the subject NP and its daughters, then the VP, and so on down the tree.

Suppose that first we encounter a node marked Si/}. The indexing instructs us to first map NP, onto the DR, in a fashion dictated by its determiner. The remainder of the sentence, with the variable z, in the place of NP, and the scope index S', is the nuclear scope of NP. Its representation will be entered into the DR in the position for NP, 's nuclear scope. This begins with the mapping of NP, into the DR in a fashion appropriate to its determiner, and the substitution of the variable z for NP, in the remainder of the sentence. This remainder then is the nuclear scope of NP, and is entered in the appropriate position. Then the mapping proceeds to consider any daughters of S, top to bottom and left to right, until all the NPs have been treated. In this way, the full DR is derived.

This mapping procedure is given in general form in (78), which works in conjunction with the mapping of NPs given in (70) and (72) above:

(78) Mapping algorithm for DRs:
To map a constituent with root node C onto a DR,
(a) if C is indexed : i . . . , map the first (moving from top to bottom, left to right) constituent NPj dominated by C into the DR. Then remove the index i from C.
(b) if C has no indices, map in turn its daughter constituents, left to right.

Now let us see how (78) works, in conjunction with (70) and (72), for the two different readings of (79) in (80) and (81):

(79) Everyone in this room speaks two languages.

(a) \[ \text{everyone}, \text{speaks [two languages],} \text{Si/}} \]

(b) \[ \text{z, speaks [two languages],} \text{Si/}} \]

(c) \[ \text{z, speaks x,} \text{Si/}} \]

In (a), the mapping of the NP with widest scope, everyone, into the DR induces box-splitting, with the introduction of the discourse referent z, in the antecedent box, along with a condition reflecting the restrictive term. In this case, since the CN is singular, the condition requiring that z be atomic is superfluous. In accord with the algorithms suggested above, NP, is replaced in the original sentence by
the corresponding variable, and its index is removed from $S$. The remainder of the sentence is then treated as the nuclear scope of $x_i$ by entering it in the right-hand box of (a). In (b) the NP two languages, is treated as an indefinite, individual-denoting NP with cardinality specification, as argued in Kadmon (1985): $|x_j| = 2$ means ‘the number of atomic i-parts of $x_j$ is 2.’ $NP_j$ is then replaced with a variable in the remaining sentence, and its index is removed from $S$. The result is then irreducible in DR terms, since all the NPs have been treated, and it is added as the nuclear scope of $NP_j$ in the right-hand box, in (c). When embedded in the model, this DR will yield the proper truth conditions for the reading indicated.

In (81), we see the mapping onto a DR from the other S-Structure for (79), where the direct object has wider scope than the subject:

\[
\text{(81)} \quad \text{[[Everyone], speaks [two languages]]} \end{array} \text{[[Everyone], speaks $x_j$]}
\]

(a) \[
\begin{array}{c}
\text{x}_j \\
\text{languages}(x_j) \\
|x_j| = 2
\end{array}
\]
remainder: [Everyone], speaks $x_j$

(b) \[
\begin{array}{c}
\text{x}_j \\
\text{languages}(x_j) \\
|x_j| = 2
\end{array}
\begin{array}{c}
\text{x}_1 \\
\text{person}(x_1) \\
\text{atomic}(x_1)
\end{array}
\Rightarrow \text{x}_1 \text{ speaks } x_j
\]
remainder of sentence: [x; speaks x]$

In (a), the NP with widest scope, $NP_j$ is treated, an indefinite acting as a variable with conditions induced by its ON and cardinality specifier, as it did in (80b). The original sentence and its index are reduced accordingly and treated as the nuclear scope of $NP_j$. In (b) the universally quantified NP, induces box-splitting, as it did in (80a). The irreducible remainder of the sentence after this step is then entered as the nuclear scope on $NP_j$ in the right-hand box of the final conditional.

Now consider an NP, say $NP_i$, which is itself group-denoting and hence does not induce box-splitting, but has a predicate which is adverbially distributive, i.e., to which the $D$ operator has applied. NP itself is first entered into the DR in a manner appropriate to its determiner, with the discourse referent $x_i$. Then the distributivity induces box-splitting, with the left-hand box containing a new discourse referent which is specified as an atomic i-part of the discourse referent for $NP_i$. The right-hand box of the conditional contains the discourse representation for the (possibly complex) predicate itself.

\[
\text{(82) Algorithm for the Treatment of Nuclear Scopes with D:}
\]
To map a structure of the form $[\ldots x_1 \ldots]_{[D]}$ onto a DR, where the structure is the nuclear scope of $NP_i$, introduce a conditional structure, with a new discourse referent $x_i$, in the left-hand box, along with the conditions i-part($x_i, x_i$) and atomic($x_i$). Remove (D) from the index on $S$, replace $x_i$ throughout with $x_i$, and introduce this remainder in the right-hand box of the conditional.

We see an example of this in (83):
4.4. Scope in Discourse Representations

Reduced as we saw in earlier examples. In (b), D introduces box-splitting and a discourse referent in the left-hand box; conditions are put on the new discourse referent so that it must be an atomic i-part of that of the subject. \[\text{(83) Some men, lifted a piano.} \]

<table>
<thead>
<tr>
<th>x₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>man(x₁)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>i-part(x₂, x₁)</td>
</tr>
<tr>
<td>atomic(x₂)</td>
</tr>
</tbody>
</table>

**Remainder of S:** [x₂ lifted [a piano] as]

(b)  

```
x₁
man(x₁)
```

```
x₂
i-part(x₂, x₁)
```

```
atomic(x₂)
```

**Remainder of S:** [x₂ lifted [a piano] as]

(c)  

```
x₁
man(x₁)
```

```
x₂
i-part(x₂, x₁)
```

```
atomic(x₂)
```

\[\text{piano(x₂) \rightarrow lifted(x₂, x₁)}\]

Only the subject's scope is indexed at S; the object is interpreted in situ, though this is not essential to the interpretation of D. The D operator is indicated following the scope index of the subject, meaning that the predicate which remains after the treatment of the subject is to be modified by adverbial D. In (a) the group-denoting subject is treated in the usual way for indefinites, introducing a new discourse referent with the same index as the NP. The original sentence is

\[\text{Those people broke their leg learning to ski.}\]

The highly preferred distributive reading of (84) may be paraphrased as 'Each of those people broke his or her leg learning to ski.' Since a demonstrative plural NP is unambiguously group-denoting, the distributivity here must be introduced via D. And though the pronoun is bound by the group-denoting subject at S-structure, in the resulting DR it should be bound by the variable over atomic i-parts which is introduced in processing adverbial D. The construction of a DR for this sentence proceeds as follows, where I have ignored the demonstrative nature of the subject:

\[\text{The index of the discourse referent to be introduced by D might be introduced at S-structure, as follows:}\]

\[\text{D) Some men, lifted a heavy rock.} \]

With the stipulation that the index of D has not been used in the discourse up to that point, I avoided adding such an index in the schema (65) for the sake of simplicity.
(85) [[Those people], broke [theirs, leg]], \(x_i\) \(a_{OH}i\)

remainder of S: \([x_i, \text{broke [theirs, leg]}], a_{OH}i\)

(b) \(x_i\)

*person\((x_i)\)

\(x_k\)

\(i\)-part\((x_k, x_i)\)

atomic\((x_k)\)

remainder of S: \([x_k, \text{broke [theirs, leg]}], a_{ij}\)

(c) \(x_i\)

*person\((x_i)\)

\(x_k\)

\(i\)-part\((x_k, x_i)\)

atomic\((x_k)\)

\(x_j\)

\(x_k\)'s leg\((x_j)\)

broke\((x_k, x_i)\)

In (85b), the index \(k\) was substituted for \(i\) throughout the remainder of (a), and hence in (c) the pronoun will be bound by the atomic \(i\)-part discourse referent, and not by the discourse referent for the group-denoting NP itself.

Mapping structures with possessive or inversely linked NPs requires no further stipulations. The NPs are simply processed in the order of their indices at any particular node, so that the first processed will have widest scope in the DR and the subsequent truth conditional embedding in a model. Consider the construction of a DR for one of May’s examples in (86):

(86) [[Someone in [every city], despies \(x_i\)], \(b_{ij}\)]

remainder: [[Someone in \(x_i\), despies \(x_i\)], \(b_{ij}\)]

(b) \(x_i\)

city\((x_i)\)

\(x_j\)

person\((x_j)\)

\(in(x_j, x_i)\)

remainder: \([x_j, \text{despies } x_i], b_{ij}\)

(c) \(x_i\)

city\((x_i)\)

\(x_j\)

person\((x_j)\)

\(in(x_j, x_i)\)

\(\text{despies}(x_j, x_i)\)

Since the index \(i\) of the inversely linked NP every city has widest scope at \(S\), it is processed first, in (a). Since it is a universally quantified NP, it induces box-splitting. Then the scope index for NP, is removed from \(S\). (When we turn to its daughter NP, because NP, has already been processed and a variable entered in its place, treatment of NP, at that point will be vacuous.) In (b) and (c) the remainder of the sentence is entered in the right-hand box of the conditional, as a
predicate on NP. Note that the same DR would result if both the subject and object NP's were interpreted in situ, instead of having S scope, since the inversely indexed subject would be processed first in that case, and hence the inversely linked complement.

What of Rooth's (1986a) numeral based donkey sentences? As noted in Chapter 3, Section 3.4.3, these all have group-denoting subjects (or other c-commanding NPs) with adverbially distributive predicates. An example is repeated below:

(87) [The "fathers with [two 'children']," send them, (both) to Montessori school.]

Intuitively, the anaphoric relation indicated is licensed in the following way: a) the star operator, *_, tells us that any atomic i-part of the nonatomic individual denoted by NP will be a father and will have two children. b) The distributive operator D tells us that each of these atomic i-parts has the property denoted by 'sends them both to Montessori school.' c) Since any atomic i-part has two children and also has this property with an unbound plural pronoun, we conclude that it is the discourse referent for the NP two children which binds the pronoun.

Suppose that (87) maps onto a DR of the following form:

(88) 

```
   x2
   "father with two children(x2)"
   i-part(x2, x1)         x1
   atomic(x2)             send-to-M-S(x2, x1)
```

In mapping from (87) to (88), the algorithms for the treatment of numerically specified indefinites and D were used. But here there is no discourse referent accessible to the discourse referent for them, x1, and so the representation is infelicitous.

Suppose that treatment of * in DRs involved universal quantification over atomic i-parts of the denotation of the plural CN. This would be consistent with Link's meaning postulates on *, discussed in Chapter 3, Section 3.2. Under this assumption, both fathers with two children and children would induce box-splitting, a DR for (87) would look like (89):

(89) 

```
   x2
   "father with two children(x2)"
   i-part(x2, x1)         x1
   atomic(x2)             father(x2)
```

Even though intuitions (a) and (b) are expressed in (89), the discourse referent x1 in the final conditional is still unbound, so that the DR is ill-formed. Since the two variables x1 are under the scope of different operators, the conditions on one are not relevant for the interpretation of the other.

* and D would have the same effect on DR construction on this approach. But I believe this is misleading, and that since they are not the same operator, they should not receive identical treatment. * is an operation which builds structure, in a sense—it gives a denotation for a plural CN by building the multisemantics generated by the denotation of the singular form. From this we may conclude that all atomic i-parts of the lattice "CN are elements in the singular denotation. Perhaps this fact should be taken to be an entailment of "(father with two children), rather than as an algorithm for deconstructing the constituent in a DR. D, on the other hand, is a distributivity operator, so the introduction of conditions on the atomic i-parts of the subject is its central function.
Chapter 5

Remarks on Plural Anaphora

In Chapters 1 and 2, I outlined a theory of anaphora. My basic assumption has been that English third person pronouns are interpreted as variables. In attempting to defend this assumption, I proposed that there are two types of pronominal binding, subject to different constraints. The constraints on c-command anaphora are expressed in terms of configurational properties of sentences. This is the type of constraint embodied in the principles of the Binding Theory, as discussed and revised in Chapter 2. Another type of constraint is ultimately semantic in nature: an anaphor must be in the scope of its antecedent. In discourse, we find hierarchical structures which are defined by various kinds of operators, including modals, adverbs of quantification, and temporal operators. As in the case of pronouns c-command bound by quantified NPs, a pronoun under the scope of such an operator must be bound within that scope. However, the possibility of accommodating preceding propositions in the discourse to serve as restrictions on the domain of an operator (this possibility itself constrained by mood, tense and the like) licenses the phenomenon of discourse subordination, which seems to extend an operator’s scope.

The examples I considered in Chapters 1 and 2 all involved only singular pronouns. Anaphora involving plural pronouns presents special problems. Here I will focus on one of the most important of these, the relation of their syntactic number to interpretation. This problem raises anew the question of the semantic content of pronouns and how it is related to that of their antecedents.

1Again, I follow Heim (1983) in assuming that declarative propositions are bound to discourse referents which are accommodated due to the salience of the referent in the context of utterance.

2I believe the relationship between these two types of constraints is probably more intimate than their relegation to two different levels of representation would suggest. If one considers the correlates of the Binding Theory principles in a categorial grammar, so that they are couched in terms of fraction-argument structure (as in Bach & Partee (1986)) instead of c-command, the relationship of the configurational constraints to scope becomes more apparent. However, I’m not convinced that the configurational constraints reduce entirely to questions of scope.
In line with the hypothesis that pronouns are variables, completely referentially dependent on their antecedents, I will argue here that the number of pronouns does not contribute directly to its interpretation: the pronoun itself has no content. Rather, any features, including number and gender, serve only to guide the hearer in determining an appropriate antecedent. In general, antecedents and snaphors must agree with respect to these features, but this does not entail coreference.

In the Heim-Kamp theory not only pronouns, but all definite (and indefinite) NPs are treated as variables. Heim (1982, p. 370) argues for what she calls the "Extended Novelty-Familiarity Condition" on the felicity of an utterance in a given context, or File. This condition places two different requirements on the felicitous use of a definite NP. The first is that a definite NP must be anaphoric; that is, in terms of Files or DRs, it is only felicitous in a discourse when it corresponds to a discourse referent which has already been introduced. The second condition is that the context in which it is uttered already presupposes its descriptive content, if it has any. The descriptive content of a definite or indefinite NP is given by its CN, which is treated as a condition on the corresponding variable (discourse referent) in the File (or DR).

One might pose the question about the relation of the syntactic number of a pronoun to its interpretation in terms of Heim's theory of definiteness. Do pronouns have descriptive content? i.e. do pronominal features for number, and perhaps gender as well, also induce conditions on the associated variable? Frey & Kamp (1986) answer in the affirmative. They stipulate that there are two types of discourse referents, those induced by singular pronouns and those induced by plurals; and that only the plural discourse referents map onto sets in the model — the equivalent of Link's nonatomic i-same. However, I will argue that the answer should be negative.4

4Kadmon (1987) argues that there is another condition on the use of definites, the Uniqueness Condition, which she formulates in terms of DR theory. The Uniqueness Condition is a variation of Evans' (1977,1980) requirement that the antecedent of a plural pronoun be the maximal collection determined by the clause containing the antecedent. For example, in his example (5):

\[(1)\]

John owns some sheep. Barry vaccinates them.

them may not refer to just any group of sheep John owns, but must refer to the maximal collection of sheep which John owns. Kadmon reviews Evans' condition in a restricted way, to reflect the fact that the maximality of a set (which is equivalent to saying its uniqueness, particularly in the context of a theory such as Link's) may be guaranteed not only by conditions on the antecedent which are introduced in process of sentence, but by implicit, accommodated and contextually supplied material as well. She also shows that her Uniqueness Condition entails the anaphoric conjunct of Heim's Extended Novelty-Familiarity Condition; she then proposes a condition, the Uniqueness and Familiarity Condition, which combines her Uniqueness Condition and Heim's requirement that descriptive content be presupposed.

In fact, I believe there are not two but three classes of anaphors, that is, of NPs which are interpreted as variables in a DR or File and require an antecedent. There is one class consisting of NPs with descriptive content, that is, definite descriptions. Another class is that of the pure pronouns, exemplified by English third person pronouns, and this is the case which I claim has no inherent content. But there is a third class of nominal anaphors which lack descriptive content and yet have some content apart from agreement. The DPs disjoint anaphors of Saxon (1986) are members of this third class of anaphoric NPs. The disjoint anaphors seem to have the content "other than x," where x is the element bound by the antecedent. I believe that English reciprocals are similar to the disjoint anaphor in this respect, though I cannot develop this idea here.

I think that the number of a pronoun, in contrast to that of a CN (recall the discussion in Section 3.2.3), does not induce any conditions on the discourse referent with which it is associated, and thus has no effect on the embedding of a DR (or File) into a model. The strongest argument for this view centers on the interpretation of plural pronouns which are bound in distributed predicates. Consider (1), from the preceding chapter:
(3) \(\lambda x_2(\text{broke}(x_1, x_2\text{’s leg}))\) (those people)

Here, ‘the property of being \(x_1\) such that \(x_1\) broke \(x_2\text{’s leg}\)’ is predicated distributively of those people.

Then, even though the anaphoric relation between the subject and the plural pronoun in (1) is represented by means of coindexation, as it was in Chapter 4, Section 4.3.2, this coindexation relation is not one of coreference. Again, coindexation is a guide to a mapping onto a DRS, and has no direct model theoretic translation. In this case, the coindexation of the subject and the pronoun in S-Structure has the effect that the subject’s argument position and that of the operator are both bound by the same operator in a logical translation. Here is where the ‘coreference’ lies, and not in a relation between the subject itself and the pronoun. Of course, it is not true coreference, either, since both variables are under the scope of the adverbal distributivity operator. Hence we can only say that their value varies in the same way.

The mapping of adverbially distributive predicate onto DRSs which I proposed in Chapter 4 reflects these facts about the binding of the pronoun in (1). Recall that adverbial \(D\) induces box-splitting and simultaneously introduces in the left-hand box a new variable over i-parts of the subject. Then in the original sentence, all indices identical with that of the subject are changed to that of the new variable over i-parts of the subject. This has the same truth conditional effect as binding by a lambda operator.

The significance of the coindexation of the subject with their in (1), then, is only realized in the relative roles of the two NPs in a predication, and not directly in terms of the ‘reference’ of the NPs in a model. The success of this approach is accounting for the truth conditions of such examples argues that there should be no direct effect of pronominal number on interpretation.

We see further evidence of this in (4):

(4) Most boys think they like themselves.

Here, the A-Pronoun themselves must be bound in its governing category, the complement \(S\), so we may conclude that its antecedent is they, and not most boys. The subject of (4) is quantificational, so that the only reading is distributive — there is no reading where some group has the property of liking itself. (5) is an indexed S-Structure for (4) which yields the proper truth conditions; this has the same truth conditions as the first order formula in (6):

(5) \[[\text{most boys}], \text{think they, like themselves}, x_2\]

(6) \(\lambda x_2(\text{think}(x_1, \text{like}(x_1, x_2)))\) (most boys)

There is no group-denoting antecedent for their in this example, and hence the plural number of the pronouns does not indicate that they are group-denoting. Rather, they only agree with their co-indexed antecedent, a syntactically plural quantificational NP.

(7) is an example where the antecedent of a plural pronoun in a distributive context is not itself the subject:

(7) Lou sends the kids a card on their birthday.

On one reading of this example, the kids may have different birthdays, so that Lou sends each a card on his or her birthday. This reading is derived by coindexing the kids and their as shown in (8), and is equivalent to the formula in (9):

(8) \[\text{Lou sends } \{\text{the kids}, \text{a card on their birthday}\}\]

(9) \(\lambda x_2 (\text{Lou sends } x_1, \text{a card on } x_1\text{’s birthday })\) (the kids)

Example (10), from Link (1986), shows that pronouns in the same distributive predicate may be bound in two ways by the same antecedent, via c-command binding and discourse binding:

(10) John and Mary invited their parents to their house.

The intended reading is one where each of the two individuals John and Mary separately invited his or her parent to the place where they live together. We can account for this reading within the framework of my assumptions by modifying the predicate with adverbial \(D\) and coindexing John and Mary with the first pronoun at S-Structure, while assigning a different index to the second pronoun, as in (11):

(11) \[[\text{John and Mary}, \text{invited } \{\text{their parents}, \text{to } \text{their house}\}]\]

Unlike Link, I believe there is also a reading where both pronouns have a group reading.
A DR for (11) which was constructed in accordance with the procedures in Chapter 4, Section 4.3.1, is shown in (12):

\[ \begin{array}{c}
\text{John(Mary(x_i))} \\
x_i \\
\text{i-part(x_k, x_j)} \\
\text{atomic(x_k)} \\
\end{array} \rightarrow \begin{array}{c}
x_j = x_i \\
x_k = x_i \\
\text{parents-of}(x_2, x_n) \\
\text{place-of}(x_m, x_j) \\
\text{invited-to}(x_2, x_j, x_n) \\
\end{array} \]

In (12), the discourse referent \( x_i \) is first entered for the group-denoting subject. Then, adverbial \( D \) induces universal quantification over atomic i-parts of the subject, replacing the index \( i \) of the subject throughout the reduced sentence with the index \( k \) of the new variable over such i-parts. In particular, this causes the first pronoun to be bound by the i-part discourse referent \( x_k \) instead of by the discourse referent of the subject, \( x_i \) (or, in terms of an equivalent intensional logic translation, by the lambda operator abstracting over the subject position). But the second pronoun was not coindexed with the subject. It must then be discourse bound, and one accessible antecedent is the subject's discourse referent, \( x_i \). Equating \( i \) with \( j \) gives the second pronoun its group reading.

These examples argue that the syntactic number on a pronoun has no direct semantic significance. Syntactic number is an agreement phenomenon. If an NP is singular, it may in general only serve as an antecedent for a singular anaphor. If it is plural, it may in general only be antecedent for a plural anaphor. This assists a hearer who is attempting to locate the proper antecedent for a given anaphor. There are, of course, exceptions to the agreement requirement. One which is increasingly accepted is the use of third person plural pronouns as gender-free anaphors, avoiding the use of the awkward expressions he or she, and his or her, as in (13):

(13) Does everyone have their student ID handy?

Such exceptions tend to confirm the characterization of pronominal number as agreement, since this is the sort of phenomenon we expect with agreement (cf. the discussion of subject/verb agreement in Section 3.2.3.).

Since these restrictions on agreement generally hold across discourse, it is necessary to indicate in some way the syntactic plurality of NPs on their discourse referents. I assume that this could be accomplished via some sort of discrict on the discourse referents. But on this view, contrary to Frey & Kamp's claim, any discricties which distinguish between the discourse referents of singular and plural NPs only serve to restrict potential anaphoric relations between discourse referents, and have no bearing on the embedding of the final DR in a model. That is, discourse referents introduced by singular NPs and those introduced by plural NPs are of the same type in the DR.
Bibliography


Bach, Emmon (1981b) "Discontinuous Constituents in Generalized Categorial Grammar." In Victoria A. Burke & James Pustejovsky (eds.), *Proceedings of NELS XI*, Graduate Linguistic Student Association, University of Massachusetts, Amherst.


Barwise, Jon (1985) "A model of the treatment of anaphors in sition seman-
tics." CSLI Working Paper #1, Stanford, University.

Barwise, Jon & Robin Cooper (1981) "Generalised Quantifiers and Natural Lan-


Brody, Michael (1979) "Infiniteivals, relative clauses and deletion," mimeographed, University of London.


Carlson, Gregory (1986) "The Semantic Composition of English Generic Sente-

Carlson, Lauri (1989) "Plural Quantification," manuscript, MIT.

Carlson, Marcia (1982) "The representation of disjunction," manuscript, University of Massachusetts, Amherst.


Chierchia, Gennaro (1982) "Nominalisation and Montague Grammar: A Seman-


Choong, Jae (1985) "Anti-quantifiers," manuscript, University of Massachusetts, Amherst.

Chomsky, Noam (1971) "Deep Structure, Surface Structure, and Semantic Inter-
pretation." In D.D. Steinberg & L.A. Jakobovits (eds.), Semantics: An Interdis-
ципiplinary Reader in Philosophy, Linguistics and Psychology. Cambridge University Press.

Chomsky, Noam (1975) "Questions of form and interpretation." In Robert Auster-


Cooper, Robin, & Terence Parsons (1976) "Montague Grammar, Generative Se-

Dabek, Nina (1984) "VP Ellipsis in File Change Semantics," manuscript, University of Massachusetts, Amherst.


Fodor, Janet, & Ivan Sag (1983) "Referential and Quantificational Indefinites, manuscript, University of Texas, Austin.


Grommendijk, Jeroen, & Martin Stokhof (1988) "Dynamic Montague Grammar: Going Back to MG and Beyond DRT". Unpublished manuscript, University of Amsterdam and Philips Research Laboratories.

Guéron, Jacqueline (1981) "Logical operators, complete constituents, extraction transformations." In Jan Koster & Robert May (eds.) *Levels of Syntactic Representation*. Foris, Dordrecht.


Heim, Irene (1985) “Presupposition projection and anaphoric relations in modal and propositional attitude contexts,” manuscript, University of Texas, Austin.


Hoeksema, Jack (1985) “Partitives,” manuscript, Ohio State University.


Karttunen, Lauri (1977) “Syntax and semantics of questions,” Linguistics and Philosophy 1:3-44.


Kratzer, Angelika (1977) "What 'must' and 'can' must and can mean," Linguistics and Philosophy 1:337-355.


Landman, Fred (1966a) Towards a Theory of Information. GRASS 6, Foris, Dordrecht.


Roeth, Mats (1986a) "Notes on NP Interpretation," manuscript, CSLI, Stanford University.

Roeth, Mats (1986b) "NP Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics," manuscript, CSLI, Stanford University.


Wilkinson, Karina (1986) "Generic indefinite NPs," manuscript, University of Massachusetts, Amherst.


